A DE FINETTI THEOREM FOR CAPACITIES: AMBIGUITY ABOUT CORRELATION*

Larry G. Epstein Kyoungwon Seo

January 16, 2010

Abstract

The de Finetti Theorem on exchangeability is generalized to a framework where beliefs are represented by belief functions. This is done while extending the scope of the bridge provided by de Finetti between subjectivist and frequentist approaches. The model is shown to accommodate ambiguity about how experiments are "correlated."

1. INTRODUCTION

Let a family of experiments be indexed by the set $\mathbb{N} = \{1, 2, ...\}$. Each experiment yields an outcome in the set S (technical details are suppressed until later). Thus $\Omega = S^\infty$ is the set of all possible sample paths. A probability measure P on Ω is *exchangeable* if

$$(\pi P)\ (A_1 \times A_2 \times .....)\ = P(A_{\pi^{-1}(1)} \times A_{\pi^{-1}(2)} \times .....),$$

for all finite permutations π of $\mathbb{N}$. De Finetti [10, 15] shows that exchangeability is equivalent to the following representation: There exists a (necessarily unique)

*Boston University, lepstein@bu.edu and Department of Managerial Economics and Decision Sciences, Northwestern University, k-seo@kellogg.northwestern.edu. This research was supported by the National Science Foundation (awards SES-0917740 and SES-0918248). The material in this paper was formerly contained in the paper titled "Symmetry, Ambiguity and Frequencies," which no longer exists as a separate paper.

probability measure μ on $\Delta(S)$ such that

$$P(\cdot) = \int_{\Delta(S)} \ell^{\infty}(\cdot)\, d\mu(\ell), \tag{1.1}$$

where, for any probability measure ℓ on S (written $\ell \in \Delta(S)$), ℓ^{∞} denotes the corresponding i.i.d. product measure on Ω. Thus beliefs are "conditionally i.i.d." Kreps [17, Ch. 11] refers to the de Finetti Theorem as "the fundamental theorem of (most) statistics," because of the justification it provides for the analyst to view samples as being independent and identically distributed with unknown distribution function.

Though the de Finetti theorem can be viewed as a result in probability theory alone, it is typically understood in economics as describing the prior in the subjective expected utility model of choice. That is how we view it in this paper. From the choice-theoretic perspective, the subjective expected utility framework precludes ambiguity aversion, as typified by Ellsberg's celebrated experiments. To permit a role for ambiguity, we consider preference on a domain of (Anscombe-Aumann) acts that conforms to Choquet expected utility where the capacity is a belief function - we call this model *belief function utility.*[1] Using the latter as the basic framework, we then impose two further axioms - Symmetry (corresponding to exchangeability) and Weak Orthogonal Independence (relaxing the Independence axiom). These axioms are shown (Theorem 4.1) to characterize the following representation for the belief function ν on Ω (see the noted theorem for the corresponding representation for utility):

$$\nu(\cdot) = \int_{Bel(S)} \theta^{\infty}(\cdot)\, d\mu(\theta), \tag{1.2}$$

where $Bel(S)$ denotes the set of all belief functions on S, μ is a probability measure on $Bel(S)$, and θ^{∞} denotes a suitable "i.i.d. product" of the belief function θ. The de Finetti-Savage model is the special case where (Independence is satisfied and hence) each θ in the support of μ is additive.

In an earlier paper Epstein and Seo [8], we elaborate on the motivation for extending the de Finetti theorem to incorporate ambiguity averse preferences, and we provide such an extension to the class of multiple-priors preferences (Gilboa and Schmeidler [13]). Belief function utility is appealing because it is a special case of both multiple-priors utility and Choquet expected utility (Schmeidler [25]), and

[1] See Section 2 for more on belief functions and the corresponding utility functions.

thus is "close" to the benchmark expected utility model. This closeness permits a sharper representation result. For example, Weak Orthogonal Independence is considerably weaker than the corresponding axiom (called Orthogonal Independence) employed in our other paper. Secondly, the latter employs an axiom called Super-Convexity that is redundant here.

The representation result obtained here is also sharper in another way. The rule for forming the i.i.d. product θ^{∞} is pinned down - it corresponds to that advocated by Hendon *et al* [14]. In contrast, the representation in [8] is much less specific in this regard, reflecting the fact (see Ghirardato [12], in particular) that stochastic independence is more complicated in the nonadditive probability (or multiple-priors) framework in that there is more than one way to form independent products. Ghirardato [12, Theorem 3] shows that the Hendon rule is the only product rule for belief functions such that the product (i) is also a belief function, and (ii) it satisfies a mathematical property called the Fubini property. In our model, it emerges as an implication of assumptions about preference. To our knowledge, our main result (see also Corollary 4.2) is the first choice-theoretic rationale for any particular i.i.d. product rule.

The Hendon product rule permits modeling ambiguity about how experiments are "correlated," which is further motivation for our model. Think, for example, of a cross-sectional empirical model where experiments correspond to the regression errors. The statistical decision-maker faces ambiguity about how these errors are related. For example, she may be concerned that some relevant variables have been omitted from the regressions, though she cannot be specific about which variables and about the consequences of their omission - some omitted variables may influence specific experiments in the same direction and others in opposite directions, while other omitted variables may have different effects. The decision-maker simply does not understand the regression errors in her model well enough to be more specific. But she is sophisticated enough to realize this and wishes to take it into account when making decisions. Our model prescribes behavior for such a decision-maker.

Thus far we have focussed on representation results. However, the importance of the de Finetti Theorem extends beyond the representation to the connection it affords between subjective beliefs and empirical frequencies. One form that the connection takes in the Bayesian framework is to relate subjective beliefs about the unknown but fixed probability law on S (the unknown "parameter"), represented by μ in (1.1), to empirical frequencies. By the law of large numbers (LLN) for exchangeable measures, empirical frequencies converge with probability 1, and thus one can view μ as representing ex ante beliefs about the limiting empirical frequency of Heads, a random variable. Thus a bridge is provided between subjectivist and frequentist theories of probability (see Kreps [17, Ch. 11], for example). Secondly, this connection can help a decision-maker to calibrate her uncertainty about the true parameter. Using a LLN for belief functions due to Maccheroni and Marinacci [18], we show (Section 5) that these aspects of the de Finetti model extend also to our generalization.[2]

Another important aspect of the de Finetti Theorem is the connection between beliefs and observations afforded via Bayesian updating of the prior μ. The combination of the de Finetti Theorem and Bayes' Rule gives the canonical model of learning or inference in economics and statistics. Our generalization of de Finetti's Theorem also admits intuitive (and dynamically consistent) updating in a limited but still interesting class of environments, namely, where an individual first samples and observes the outcomes of some experiments, and then chooses how to bet on the outcomes of remaining experiments. The analysis of updating described in our other paper spells this out and, importantly, it applies virtually verbatim to the present belief-function-based model. Thus we do not elaborate on updating here.

Shafer [27] is the first, to our knowledge, to discuss the use of belief functions within the framework of parametric statistical models analogous to de Finetti's. In particular, he sketches (section 3.3) a de Finetti-style treatment of randomness based on belief functions. His model is not axiomatic or choice-based, but ignoring these differences, one can translate his suggested model into our framework in the following way. Consider the de Finetti representation (1.1), where the probability measure μ models beliefs about ℓ, the unknown 'parameter'. An obvious generalization is to replace μ by a belief function on $\Delta(S)$, or more generally by a set of probability measures on $\Delta(S)$, thus generalizing prior beliefs. Epstein and Seo [8, Theorem 3.2] axiomatize such a model within the multiple-priors framework, and Al-Najjar and De Castro [1] provide a more general model in the same vein. These models intersect the present model only in the classic de Finetti expected utility model. The present paper is most closely related to the second model in [8], which has a counterpart of Theorem 4.1 for the multiple-priors framework. We described above some ways in which the present theorem is much sharper.

[2] Our other paper does not include a version of this result.

2. BELIEF FUNCTIONS

We will deal with two different (compact metric) state spaces - S corresponding to a single experiment, and $\Omega = S^\infty$, describing all possible sample paths. Thus in this section we consider an abstract (compact metric) state space X. It has Borel σ-algebra Σ_X.

A *belief function* on X is a set function $\nu : \Sigma_X \to [0,1]$ such that:[3]

Bel.1 $\nu(\varnothing) = 0$ and $\nu(X) = 1$

Bel.2 $\nu(A) \subset \nu(B)$ for all Borel sets $A \subset B$

Bel.3 $\nu(B_n) \downarrow \nu(B)$ for all sequences of Borel sets $B_n \downarrow B$

Bel.4 $\nu(G) = \sup\{\nu(K) : K \subset G,\ K \text{ compact}\}$, for all open G

Bel.5 ν is totally monotone (or ∞-monotone): for all Borel sets $B_1, .., B_n$,

$$\nu\left(\cup_{j=1}^n B_j\right) \geq \sum_{\varnothing \neq J \subset \{1,\dots,n\}} (-1)^{|J|+1} \nu\left(\cap_{j\in J} B_j\right)$$

The set of all belief functions on X is $Bel(X)$. It is compact metric when endowed with the topology for which $\nu_n \to \nu$ if and only if $\int f d\nu_n \to \int f d\nu$ for every continuous function f on X, where the integral here and throughout is in the sense of Choquet (see Schmeidler [25]).

Denote by $\Delta(X)$ the set of Borel probability measures on X, endowed with the weak convergence topology (generated by continuous functions), and by $\mathcal{K}(X)$ the set of compact subsets of X, endowed with the Hausdorff metric. Both are compact metric. If $m \in \Delta(X)$, then m^∞ denotes the usual i.i.d. product measure on X^∞.[4]

Each belief function defines a preference order or utility function. Interpreting X as a state space, denote by $\mathcal{F}(X)$ the set of all (measurable) acts $f : X \to [0,1]$. For any $\nu \in Bel(X)$, let $U_\nu : \mathcal{F}(X) \to \mathbb{R}$ be defined by

$$U_\nu(f) = \int f d\nu. \tag{2.1}$$

[3] These conditions are adapted from [23], to which we refer the reader for details supporting much of the outline in this section. We point out only that when restricted to probability measures, Bel.4 is the well-known property of regularity.

[4] Throughout product spaces are endowed with the product metric.

Refer to U_ν as a *belief-function utility*, and to the corresponding preference order as a *belief-function preference.*

Since a belief function is also a capacity (defined by Bel.1-Bel.4), belief-function utility is a special case of Choquet expected utility, axiomatized by Schmeidler [25]; and since it is convex (supermodular, or 2-alternating, that is, satisfies the inequalities in Bel.5 restricted to $n = 2$), it is well-known that (2.1) can be expressed alternatively as

$$U_\nu(f) = \min_{P \in core(\nu)} \int f dP,$$

where

$$core(\nu) = \{P \in \Delta(X) : P(\cdot) \geq \nu(\cdot)\}.$$

Thus the current model is also a special case of multiple-priors utility, which has been axiomatized by Gilboa and Schmeidler [13].

Note that acts are taken to be real-valued and to enter linearly into the Choquet integral. This may be justified as follows: Suppose that outcomes lie in an abstract set Z, and that (Anscombe-Aumann) acts map states into $\Delta(Z)$. Suppose also that there exist best and worst outcomes $\overline{z}$ and $\underline{z}$. Then, under weak conditions, for each state ω and act f, there exists a unique probability p, so that the constant act $f(\omega)$ is indifferent to the lottery $(\overline{z}, p; \underline{z}, 1-p)$; refer to such a lottery as (a bet on) the toss of an (objective) *p-coin.* Such calibration renders the util-outcomes of any act observable, and these are the $[0,1]$-valued outcomes we assume herein and that justify writing utility as in (2.1). A further consequence, given (2.1) is that the utility $U_\nu(f)$ is also scaled in probability units - it satisfies

$$f \sim (\overline{z}, U_\nu(f); \underline{z}, 1 - U_\nu(f)). \tag{2.2}$$

Thus f is indifferent to betting on the toss of a $U_\nu(f)$-coin.

Belief functions have been widely studied (see [5, 19, 21], for example) and used (for applications in robust statistics see Huber [16], and for applications in decision theory and economics, see [9, 23, 18], for example). They (and their corresponding utility functions) also admit an intuitive justification due to Dempster [6] and Shafer [26], (see also Mukerji [20] and Ghirardato [11]), that we outline next.

Though there exists a Savage-style state space X, the agent's perceptions are coarse and are modeled through an auxiliary epistemic (compact metric) state space $\widehat{X}$ and a (measurable and nonempty-compact-valued) correspondence Γ

from $\widehat{X}$ into X. There is a Borel probability measure p representing beliefs on $\widehat{X}$.

$$\begin{array}{ccc} (\widehat{X}, p) & \overset{\Gamma}{\rightsquigarrow} & (X, \nu) \\ & \searrow & \downarrow_f \\ & & [0,1] \end{array} \tag{2.3}$$

A Bayesian agent would view each physical action as an act from X to the outcome set $[0,1]$, and would evaluate it via its expected utility (using a probability measure on X). The present agent is aware that while she can assign probabilities on $\widehat{X}$, events there are only imperfect indicators of payoff-relevant events in X. Such awareness and a conservative attitude lead to preference that can be represented by the utility function

$$U^{DS}(f) = \int_{\widehat{X}} \left(\inf_{x \in \Gamma(\widehat{x})} f(x) \right) dp = \int_X f d\nu,$$

where ν is the belief function given by[5]

$$\nu(A) = p\left(\left\{\widehat{x} \in \widehat{X} : \Gamma(\widehat{x}) \subset A\right\}\right), \text{ for every } A \in \Sigma_X.$$

As a foundation for belief function utility, the preceding is suggestive though limited, because $\widehat{X}$, p and Γ are presumably not directly observable. However, Epstein, Marinacci and Seo [7, Section 4.3] describe behavioral foundations for a Dempster-Shafer representation.

A central fact about belief functions is the Choquet Theorem [23, Thm. 2].[6] To state it, note that, by [23, Lemma 1], $\{K \in \mathcal{K}(X) : K \subset A\}$ is universally measurable for every $A \in \Sigma_X$. Further, any Borel probability measure (such as m on $\Sigma_{\mathcal{K}(X)}$) admits a unique extension (also denoted m) to the collection of all universally measurable sets.[7]

[5] We can view Γ as a function from $\widehat{X}$ to $\mathcal{K}(X)$. Then Γ is measurable [2, Thm. 18.10] and induces the measure $p \circ \Gamma^{-1}$ on $\mathcal{K}(X)$. Choquet's theorem (see below) implies that $\nu(\cdot) = p \circ \Gamma^{-1}(\{K : K \subset \cdot\})$ is a belief function.

[6] The final assertion in the theorem stated below relies also on [23, Thm. 3]. See also [19, Thm. 5.1] and [4, Thm. 3.2].

[7] Throughout, given any Borel probability measure, we identify it with its unique extension to the σ-algebra of universally measurable sets.

Theorem 2.1 (Choquet). *The set function $\nu : \Sigma_X \to [0,1]$ is a belief function if and only if there exists a (necessarily unique) Borel probability measure m_ν on $\mathcal{K}(X)$ such that*

$$\nu(A) = m_\nu(\{K \in \mathcal{K}(X) : K \subset A\}), \text{ for every } A \in \Sigma_X. \tag{2.4}$$

Moreover, in that case, for every act f,[8]

$$\begin{aligned} U_\nu(f) &= \int_X f d\nu = \int_{\mathcal{K}(X)} \left(\inf_{P \in \Delta(K)} P \cdot f \right) dm_\nu(K) \\ &= \int_{\mathcal{K}(X)} \left(\inf_{x \in K} f(x) \right) dm_\nu(K). \end{aligned} \tag{2.5}$$

The one-to-one mapping $\nu \longmapsto m_\nu$ is denoted ζ. It constitutes a homeomorphism between $Bel(X)$ and $\Delta(\mathcal{K}(X))$. One perspective on the theorem is that it shows that any belief function has a special Dempster-Shafer representation where

$$\widehat{X} = \mathcal{K}(X), \Gamma(K) = K \subset X \text{ and } p = m_\nu.$$

Conclude this overview of belief functions with a simple example. Let $X = \{H, T\}$ and $[p_m, p_M] \subset [0,1]$, thought of as an interval of probabilities for Heads. Define θ on subsets of X, by $\theta(H) = p_m$, $\theta(T) = 1 - p_M$, and $\theta(X) = 1$. Then θ is a belief function - the measure m from Choquet's Theorem is $m(H) = \theta(H)$, $m(T) = \theta(T)$ and $m(\{H,T\}) = 1 - \theta(T) - \theta(H)$, the length of the interval. An interpretation is that the coin is seen as being drawn from an urn containing many coins, of which the proportion $\theta(H)$ ($\theta(T)$) are sure to yield Heads (Tails), and where there is complete ignorance about the remaining proportion. In particular, for binary state spaces, a belief function can be thought of simply as a probability interval.

[8] $P \cdot f$ is short-hand for $\int_X f dP$.

3. THE MODEL

While the above refers to an abstract state space X, we assume a specific structure here corresponding to the presence of many experiments. Thus let S be a compact metric space thought of as the stage state space, or the set of possible outcomes for any single experiment. The full state space is

$$\Omega = S_1 \times S_2 \times \ \ldots . = S^{\infty}, \text{ where } S_i = S \text{ for all } i.$$

Denote by Σ_i the Borel σ-algebra on S_i, which can be identified with a σ-algebra on Ω, and by Σ_Ω, the product Borel σ-algebra. For the most part, we take the abstract space X above to be Ω, and in that case, when there is no danger of confusion, we abbreviate Σ_Ω, $\mathcal{K}(\Omega)$, $Bel(\Omega)$ and $\mathcal{F}(\Omega)$ by Σ, $\mathcal{K}$, Bel and $\mathcal{F}$. For any $I \subset \mathbb{N}$, Σ_I is the σ-algebra generated by $\{\Sigma_i : i \in I\}$ and $\mathcal{F}_I$ denotes the set of Σ_I-measurable acts. An act is said to be *finitely-based* if it lies in $\cup_{I \text{ finite}} \mathcal{F}_I$. We will have occasion to refer also to $\mathcal{K}(S)$ and $Bel(S)$.

Let Π be the set of (finite) permutations on $\mathbb{N}$. For $\pi \in \Pi$ and $\omega = (s_1, s_2, \ldots) \in \Omega$, let $\pi\omega = \left(s_{\pi(1)}, s_{\pi(2)}, \ldots\right)$. Given an act f, define the permuted act πf by $(\pi f)(s_1, \ldots, s_n, \ldots) = f\left(s_{\pi(1)}, \ldots, s_{\pi(n)}, \ldots\right)$.

We are given a belief-function preference $\succeq$ and the corresponding utility function U. We adopt two axioms for preference (or equivalently, for utility) that describe the individual's perception of how experiments are related.

The first axiom is the preference counterpart of de Finetti's assumption of exchangeability.

Axiom 1 (SYMMETRY). *For all finitely-based acts f and permutations π,*

$$f \sim \pi f.$$

Our second axiom relaxes the Independence axiom.

Axiom 2 (WEAK ORTHOGONAL INDEPENDENCE (WOI)). *For all $0 < \alpha \leq 1$, and acts $f', f \in \mathcal{F}_I$ and $g \in \mathcal{F}_J$, with I and J finite and disjoint,*

$$f' \succeq f \iff \alpha f' + (1-\alpha) g \succeq \alpha f + (1-\alpha) g.$$

It is easy to show that WOI is satisfied if and only if U satisfies: For all $f \in \mathcal{F}_I$ and $g \in \mathcal{F}_J$, where I and J are finite and disjoint, and for all α in $[0,1]$,

$$U(\alpha f + (1-\alpha) g) = \alpha U(f) + (1-\alpha) U(g). \tag{3.1}$$

We use this characterization of WOI frequently in the sequel.

Roughly, the invariance that is imposed by the Independence axiom for all acts is restricted in WOI to mixtures of acts that depend on different experiments. To interpret the axiom, recall the Gilboa-Schmeidler intuition for violation of Independence when ambiguity matters: randomization can be desirable because it smooths out ambiguity, or, adapting finance terminology, because the acts being mixed may "hedge" one another. From this perspective, WOI says, in part, that acts that depend on different experiments do not hedge one another, presumably because the poorly understood factors underlying ambiguity are "unrelated" across experiments. However, as noted in the introduction, "stochastic independence" is multifaceted if there is ambiguity, and WOI leaves a great deal of scope for ambiguity about how experiments are related. This will be confirmed by the representation and its interpretation, and is also illustrated by the following examples.

Let each experiment correspond to a coin toss, $S = \{H, T\}$. It is the same coin being tossed repeatedly, but different tosses are performed by different people. The coin is known to be unbiased (a simplification that can be relaxed, as described below), but the decision-maker believes that outcomes depend also on the way in which the coin is tossed. Her understanding of tossing technique is poor, which leads to ambiguity about the sequence of outcomes she can expect. In particular, there is ambiguity about whether outcomes are "positively correlated" - a specific outcome on the first toss makes it more likely that the same outcome occurs in the second toss - or "negatively correlated", defined analogously.

To see what the latter ambiguity implies, for any event $A \subset \Omega$, let A denote also the bet on A, the act that pays 1 util if $\omega \in A$ and pays 0 otherwise). Then $\{H_1T_2, T_1H_2\}$ and $\{H_1H_2, T_1T_2\}$ denote the bets that the first two outcomes are different or the same respectively. These acts hedge against ambiguity about correlation - the first pays well if correlation is negative and poorly otherwise, while the opposite holds for the second bet. In fact, they hedge each other perfectly in that the mixed bet yields the payoff $\frac{1}{2}$ regardless of the outcomes of the coin tosses. Thus it is intuitive that[9]

$$U\left(\tfrac{1}{2}\{H_1T_2, T_1H_2\} + \tfrac{1}{2}\{H_1H_2, T_1T_2\}\right) > \tfrac{1}{2}U\left(\{H_1T_2, T_1H_2\}\right) + \tfrac{1}{2}U\left(\{H_1H_2, T_1T_2\}\right). \tag{3.2}$$

This is contrary to the Independence axiom, but is consistent with WOI.

[9] By (2.2), this is a statement about preference. It can be restated as: if $\{H_1T_2, T_1H_2\} \sim p$ and $\{H_1H_2, T_1T_2\} \sim q$, then $\frac{1}{2}\{H_1T_2, T_1H_2\} + \frac{1}{2}\{H_1H_2, T_1T_2\} \succ \frac{1}{2}p + \frac{1}{2}q$.

For another example of behavior that can be interpreted similarly and that is consistent with WOI, consider the bets $f = \{H_1H_2, T_1T_2\}$ and $\pi f = \{H_1H_3, T_1T_3\}$, where π is the permutation that switches experiments 2 and 3. Then $f \sim \pi f$ by Symmetry. Suppose the individual is concerned that tosses may follow a pattern - *either* the first two tosses are "negatively correlated", *or* negative correlation exists between the first and third tosses. The first pattern would make f a poor prospect, but not the second, and their roles reverse for πf. Therefore, randomizing between f and πf would hedge the uncertainty about which pattern is valid, and leave a mixture strictly preferable to f. Thus the ranking

$$\tfrac{1}{2}f + \tfrac{1}{2}\pi f \succ f \tag{3.3}$$

seems intuitive even though, unlike in the previous instance, the two bets do not hedge one another perfectly.

There is another side to WOI, because there is another motivation for randomizing that is not derived from ambiguity about correlation. Thus, for example, consider the rankings

$$\tfrac{1}{2}H_1 + \tfrac{1}{2}T_2 \succ H_1 \sim T_2. \tag{3.4}$$

Here we assume for simplicity that Heads and Tails are thought to be equally likely. Modify the description of the coin-tossing experiments by supposing that (i) technique is thought to be irrelevant, but (ii) the coin's bias is unknown, and indeed, is ambiguous. Then the mixed bet $\frac{1}{2}H_1 + \frac{1}{2}T_2$ may be strictly preferable if there is ambiguity about the physical bias of the coin - this is exactly the intuition in Gilboa and Schmeidler [13]. WOI excludes the rankings (3.4). Thus we interpret the axiom as expressing *both* a weak form of stochastic independence across experiments *and* the absence of ambiguity about the factors (such as the bias of the single coin) that are common to all experiments.[10]

For a functional form example, let $K^* = \{(H, H, ...), (T, T, ...)\}$ and

$$\nu(A) = \begin{cases} 1 & A \supset K^* \\ 0 & \text{otherwise.} \end{cases}$$

Then, by the Choquet Theorem, ν is a belief function on Ω. The corresponding utility function, as in (2.1), satisfies Symmetry but not WOI, since (3.4) is easily

[10] See our other paper [8] for further discussion and for a model, in the multiple-priors framework, that captures ambiguity about the bias. We do not, however, have a single model that accommodates both reasons for randomization being valuable.

verified.[11] The reason is that according to ν, there is ambiguity about whether the coin is completely biased towards Heads (corresponding to the sequence $(H, H, ...)$) or completely biased towards Tails (corresponding to the sequence $(T, T, ...)$).

4. THE REPRESENTATION

Given $\theta \in Bel(S)$, we define an "i.i.d. product" θ^∞, a belief function on S^∞, as follows. We have $\zeta(\theta) \in \Delta(\mathcal{K}(S))$, and hence $(\zeta(\theta))^\infty \in \Delta[(\mathcal{K}(S))^\infty] \subset \Delta[(\mathcal{K}(S^\infty))]$.[12] By the Choquet Theorem, there exists a (unique) belief function on S^∞ corresponding to $(\zeta(\theta))^\infty$. Denote it by θ^∞, so that

$$\zeta(\theta^\infty) = (\zeta(\theta))^\infty .$$

Since Hendon et al [14] propose this rule in the case of finite Cartesian products, we refer to θ^∞ as the *Hendon i.i.d. product.*[13] When θ is additive, and thus a probability measure, then θ^∞ is the usual i.i.d. product.

The belief-function utility V on $\mathcal{F}$ is called an *i.i.d. (belief-function) utility* if there exists $\theta \in Bel(S)$ such that

$$V(f) = V_{\theta^\infty}(f) \equiv \int f d(\theta^\infty), \text{ for all } f \in \mathcal{F}.$$

It is easily verified (using (2.4)) that[14]

$$\theta^\infty(A_I \times A_J \times S^\infty) = \theta^\infty(A_I \times S^\infty)\ \theta^\infty(A_J \times S^\infty), \tag{4.1}$$

for $A_I \in \Sigma_I$ and $A_J \in \Sigma_J$, where $I, J \subset \mathbb{N}$ are finite and disjoint, which suggests one sense in which "stochastic independence" is embodied in θ^∞.

Our main result is that Symmetry and WOI characterize utility functions that are "mixtures" of i.i.d. utilities.

---

[11] Talagrand [28] contains the study of symmetric belief functions, where WOI for the corresponding utility function is not assumed.

[12] Recall that ζ denotes the homeomorphism defined by Theorem 2.1. We denote by ν a generic belief function on S^∞ and by θ a generic belief function on S.

[13] Recall the discussion in the introduction and the connection to Ghirardato [12].

[14] A more general "product relation" involving (nonindicator) acts is aso satisfied; see (B.1) in the proof of Corollary 4.2 below.

Theorem 4.1. *Let U be a belief function utility. Then the following statements are equivalent:*

(i) U satisfies Symmetry and Weak Orthogonal Independence.

(ii) There exists a (necessarily unique) Borel probability measure μ on $Bel(S)$ such that

$$U(f) = \int_{Bel(S)} V_{\theta^\infty}(f)\, d\mu(\theta), \quad \text{for every } f \text{ in } \mathcal{F}. \tag{4.2}$$

(iii) There exists a (necessarily unique) Borel probability measure μ on $Bel(S)$ such that ν, the belief-function corresponding to U, can be expressed in the form

$$\nu(A) = \int_{Bel(S)} \theta^\infty(A)\, d\mu(\theta), \quad \text{for every } A \text{ in } \Sigma. \tag{4.3}$$

As emphasized earlier, we interpret the de Finetti Theorem as a result regarding preference that assumes subjective expected utility. With this interpretation, we generalize his result to the framework of belief function preference.

The more general representation (4.3) also admits a "conditionally i.i.d." interpretation. The fact that each θ is a belief function rather than a probability measure leads to a difference in interpretation. Consider the above coin-tossing setting for concreteness, so that $S = \{H, T\}$. Let the bias be unknown and let tossing technique be thought to be important. In the Bayesian model, each experiment is characterized by a single number in the unit interval - the probability of Heads. Here, instead an experiment is characterized by an interval of probabilities for Heads, which is nondegenerate because even given the physical bias of the coin, the influence of tossing technique is poorly understood. For any $\theta \in Bel(S)$ appearing in (1.2), the interval is $[\theta(H), 1 - \theta(T)]$.

Nonadditivity of belief functions leaves scope for ambiguity about correlation. At the functional form level, the latter can be seen by examining more closely the form of each i.i.d. utility function V_{θ^∞}, and thus expressions of the form

$$V_{\theta^\infty}(f) \equiv \int f d(\theta^\infty).$$

Let $m_\theta \in \Delta(\mathcal{K}(S))$ denote the measure on subsets implied by Choquet Theorem 2.1, in which case the i.i.d. product $(m_\theta)^\infty \in \Delta[(\mathcal{K}(S))^\infty]$ is the corresponding measure for θ^∞. Then, by (2.5),

$$V_{\theta^\infty}(f) = \int_{(\mathcal{K}(S))^\infty} \left(\inf_{P \in \Delta(K_1 \times K_2 \times \ldots)} P \cdot f \right) d(m_\theta)^\infty(K_1, K_2, \ldots). \tag{4.4}$$

Aversion to correlation between experiments, such as expressed by (3.2) or (3.3), derives from the minimization over *all joint* measures on each $K_1 \times K_2 \times ....$[15]

Paralleling de Finetti's contribution for a Bayesian framework, the representation (1.2) can aid in forming a (probabilistic) prior even where experiments are ambiguous. For example, it is arguably easier to decide on which intervals might describe every coin and on a probability distribution over these intervals than to arrive at beliefs, in the form of a belief function, directly over all possible sample realizations.

Finally, for those readers who are concerned about practicality, we offer some reassurance. One obstacle to evaluating the utility functions appearing in the theorem, is evaluation of i.i.d. utility functions V_{θ^∞}, and thus expressions of the form

$$V_{\theta^\infty}(f) \equiv \int f d(\theta^\infty).$$

But as just noted in (4.4), $V_{\theta^\infty}(f)$ equals a standard integral with respect to an additive product measure. One need only derive m_θ for each relevant θ. For a binary state space, the one-to-one map between θ and m_θ was described at the end of Section 2. Also more generally, if S is finite, then m_θ can be constructed explicitly from θ by the so-called Mobius inversion formula[16]

$$m_\theta(A) = \Sigma_{B\subset A}(-1)^{\#(A\setminus B)}\theta(B), \text{ if } A \subset S.$$

See Hendon *et al* [14, p. 100], for example.

Another possible concern is whether a decision-maker can plausibly arrive at a prior μ over belief functions. However, belief functions are often not such complicated objects. For example, with repeated coin-tossing, when S is binary, each belief function θ corresponds to a unique interval, and the latter corresponds to an ordered pair of real numbers - in other words, one need only formulate a prior over an unknown two-dimensional parameter. More generally, since each θ corresponds to a unique Mobius inverse m_θ, the task is to form a prior over $\Delta(\mathcal{K}(S))$. This is perhaps more difficult than forming a prior over $\Delta(S)$, as required by de Finetti, but is qualitatively comparable to the latter. Some guidance for our

[15] The former is immediate. For the latter, let μ attach positive probability to θ, where $\theta(H),\ \theta(T) > 0$, and $\theta(H) + \theta(T) < 1$. Then, by straightforward but tedious calculations, $V_{\theta^\infty}(f) = (\theta(H))^2 + (\theta(T))^2 < V_{\theta^\infty}\left(\frac{1}{2}f + \frac{1}{2}\pi f\right) = (\theta(H))^2 + (\theta(T))^2 + \theta(H)\theta(T)(1-\theta(H)-\theta(T))$. Therefore, (3.3) follows.

[16] Since m_θ is additive, it is uniquely defined as a measure on $\mathcal{K}(S)$ by its values $m_\theta(A)$ for every $A \subset S$.

agents in calibrating beliefs is provided in the next section, when the connection to empirical frequencies is considered.

A special case of the model is where there is certainty about the true belief function, and hence where the utility function U itself is IID, that is,

$$U(\cdot) = V_{\theta^\infty}(\cdot) \text{ for some } \theta \text{ in } Bel(S). \tag{4.5}$$

This case is of particular interest, especially from the perspective of the question of how to form independent products of belief functions (Hendon [14] and Ghirardato [12]). The next result gives a behavioral justification for using the Hendon product.

Denote by $\varkappa$ the *shift* operator, so that, for any act,

$$(\varkappa f)(s_1, s_2, s_3, ...) = f(s_2, s_3, ..).$$

It is straightforward to show that Symmetry implies also indifference to shifts,[17]

$$\varkappa f \sim f \quad \text{for all } f \in \mathcal{F}.$$

We define also the pointwise product of two acts. For any two acts f^* and f, $f^* \cdot f$ denotes the act satisfying

$$(f^* \cdot f)(\omega) = f^*(\omega) f(\omega), \text{ for all } \omega.$$

Let $\mathcal{F}_1 \subset \mathcal{F}$ denote the set of acts depending on the first experiment only ($f : S_1 \longrightarrow [0,1]$). Then, for any f in $\mathcal{F}_1$,

$$(f \cdot \varkappa f)(s_1, s_2, ...) = f(s_1) f(s_2).$$

For example, if f is the bet H_1 on the outcome Heads, then $f \cdot \varkappa f$ is the bet $H_1 H_2$ that the first two tosses both yield Heads.

Corollary 4.2. *Let U be a belief function utility. Then the following statements are equivalent:*

(i) The utility function U has the IID form in (4.5).

(ii) U satisfies Symmetry and Weak Orthogonal Independence, and for all $f \in \mathcal{F}_1$,

$$f \sim p \implies f \cdot \varkappa f \sim p^2. \tag{4.6}$$

[17] See our earlier paper [8].

To interpret the behavioral condition in (ii), consider for simplicity the special case of bets (binary acts). Suppose that a bet on A is indifferent to the bet on a coin with known objective probability p.[18] How would an individual rank two-fold repetitions of each? In the case of the coin, the two tosses would be independent and thus have probability p^2 of success. For the subjective bet, the repetitions are not plausibly viewed as independent in general, where there is a common unknown element connecting experiments - like the unknown bias of a coin that is tossed repeatedly. In that case, (4.6) is replaced by $f \sim p \implies f \cdot \varkappa f \succeq p^2$, (Epstein and Seo [8, Section 5]), generalizing the well-known fact that for any random sequence (X_t) having an exchangeable probability law, the covariance of any X_i and X_j is non-negative. Intuitively, indifference holds as in (4.6) if and only if there is no unknown common element, which the Corollary translates into the precise statement that utility is IID.

5. FREQUENCIES

In this section, we relate subjective uncertainty about the true i.i.d. belief function θ^∞, represented by μ, to beliefs about empirical frequencies. Formally, our result is a corollary of our de Finetti-style representation and a law of large numbers (LLN) for belief functions due to Maccheroni and Marinacci [18].

The coin-tossing setting conveys the point most clearly. Let $S = \{H, T\}$ and denote by $\Psi_n(\omega)$ the proportion of Heads realized in the first n experiments along the sequence ω. Then, for any $\theta \in Bel(S)$, the noted LLN for belief functions implies that

$$\theta^\infty\left(\{\theta(H) \leq \liminf \Psi_n(\omega) \leq \limsup \Psi_n(\omega) \leq 1 - \theta(T)\}\right) = 1. \tag{5.1}$$

Further, these bounds on empirical frequencies are tight in the sense that[19]

$$\begin{aligned} [a &> \theta(H) \text{ or } b < 1 - \theta(T)] \implies 0 = \\ &\theta^\infty\left(\{a \leq \liminf \Psi_n(\omega) \leq \limsup \Psi_n(\omega) \leq b\}\right). \end{aligned} \tag{5.2}$$

Therefore, the representation (4.3) implies that, for every $0 \leq a \leq b \leq 1$,

$$\begin{aligned} &\mu\left(\{\theta : a \leq \theta(H) \leq 1 - \theta(T) \leq b\}\right) \\ = &\nu\left(\{a \leq \liminf \Psi_n(\omega) \leq \limsup \Psi_n(\omega) \leq b\}\right). \end{aligned} \tag{5.3}$$

[18] Recall (2.2).

[19] This is proven in the context of proving the Corollary below.

This equality admits an appealing interpretation. In the Bayesian setting, each coin toss is described by a common unknown probability of Heads, and the LLN justifies interpreting uncertainty about this "parameter" in terms of uncertainty about the limiting empirical frequency of Heads. In our setting, the individual is not certain that empirical frequencies converge to a fixed point, and she thinks in terms of intervals that will contain all limit points. Supposing for simplicity that μ has finite support, then (5.3) is equivalent to:

$$\mu(\theta) = \nu(\{\omega : \theta(H) \leq \liminf \Psi_n(\omega) \leq \limsup \Psi_n(\omega) \leq 1 - \theta(T)\}) . \tag{5.4}$$

Thus the prior subjective probability of the unknown (but nonrandom) parameter θ equals the prior likelihood, according to ν, that the interval $[\theta(H), 1-\theta(T)]$ will contain the random interval of empirical frequencies in the long run. This provides a frequentist perspective for the probability measure μ over belief functions.

Consistent with the normative slant of our model, it is also worthwhile noting that (5.4) can also help a decision-maker, a statistician for example, to calibrate her uncertainty about the true θ. That is because $\mu(\theta)$ equals that prize which, if received with certainty, would be indifferent for her to betting (with prizes 1 and 0) on the event that

$$[\liminf \Psi_n(\omega), \limsup \Psi_n(\omega)] \subset [\theta(H), 1 - \theta(T)].$$

We elaborate briefly on the formal meaning of the preceding. Any $\theta \in Bel(S)$ is completely determined by the two numbers $\theta(H)$ and $\theta(T)$, or equivalently by the interval

$$I_\theta = [\theta(H), 1 - \theta(T)].$$

Moreover, $\theta \longmapsto I_\theta$ is one-to-one. Thus the representing measure μ can be thought of as a measure over intervals I_θ. Formally, let $\mathcal{I}$ be the collection of all compact subintervals of $[0, 1]$. As a subset of $\mathcal{K}([0, 1])$, $\mathcal{I}$ inherits the Hausdorff metric and the associated Borel σ-algebra. Moreover, $Bel(S)$ is homeomorphic to $\mathcal{I}$, and thus there is a one-to-one correspondence, denoted e, between probability measures on $Bel(S)$, and probability measures on intervals, that is, measures in $\Delta(\mathcal{I})$. In particular, μ can be identified with a unique $\widehat{\mu} = e(\mu)$ in $\Delta(\mathcal{I})$. Thus (5.3) can be written in the form

$$\begin{aligned} & \widehat{\mu}(\{I : I \subset [a, b]\}) \\ = \;& \nu(\{\omega : [\liminf \Psi_n(\omega), \limsup \Psi_n(\omega)] \subset [a, b]\}) . \end{aligned} \tag{5.5}$$

The general (nonbinary) case is similar. Denote by $\Psi_n(\cdot)(\omega)$ the empirical frequency measure given the sample ω; $\Psi_n(B)(\omega)$ is the empirical frequency of $B \in \Sigma_S$ in the first n experiments. The above reasoning can be extended to prove:

Corollary 5.1. *Let $U = U_\nu$ be a belief function utility satisfying Symmetry and Weak Orthogonal Independence. Then the equivalent statements in Theorem 4.1 are equivalent also to the following: There exists a probability measure μ on $Bel(S)$ satisfying both (i) μ represents U in the sense of (4.2); and (ii) for every finite collection $\{A_1, ..., A_I\} \subset \Sigma_S$, and for all $a_i \leq b_i$, $i = 1, ..., I$,*

$$\begin{aligned} & \mu\left(\bigcap_{i=1}^{I}\{\theta : [\theta(A_i), 1 - \theta(S \backslash A_i)] \subset [a_i, b_i]\}\right) \qquad (5.6) \\ = \; & \nu\left(\bigcap_{i=1}^{I}\{\omega : [\liminf \Psi_n(A_i)(\omega), \limsup \Psi_n(A_i)(\omega)] \subset [a_i, b_i]\}\right). \end{aligned}$$

Equation (5.6) relates the prior μ over belief functions to ex ante beliefs about empirical frequencies for the events $A_1, ..., A_I$. More precisely, the μ-measures of only the sets shown are so related. However, as our final result shows, μ is completely determined by its values on these sets.

Proposition 5.2. *If $\mu, \mu' \in \Delta(Bel(S))$ coincide on all sets of the form*

$$\{\theta \in Bel(S) : \theta(A_1) \geq a_1, ..., \theta(A_I) \geq a_I\},$$

where A_i, a_i and I vary over Σ_S, $[0,1]$ and the positive integers respectively, then $\mu = \mu'$.

A. Appendix: Proof of Theorem 4.1

First we prove the measurability required to show that the integrals in (4.2) and (4.3) are well-defined. (Recall that the Borel probability measure μ has a unique extension to the class of all universally measurable subsets.)

Lemma A.1. *Both $\theta \longmapsto V_{\theta^\infty}(f)$ and $\theta \longmapsto \theta^\infty(A)$ are universally measurable for any $f \in \mathcal{F}$ and $A \in \Sigma$.*

Proof. Since $Bel(S)$ and $\Delta(\mathcal{K}(S))$ are homeomorphic, and in light of (2.5), it is enough to prove analytical (and hence universal) measurability of the mapping from $\Delta(\mathcal{K}(S))$ to $\mathbb{R}$ given by

$$\ell \longmapsto \int_{[\mathcal{K}(S)]^\infty} \inf_{\omega \in K} f(\omega)\, d\ell^\infty(K).$$

Step 1. $\Delta(\mathcal{K}(S))$ and $\{\ell^\infty : \ell \in \Delta(\mathcal{K}(S))\}$ are homeomorphic when the latter set is endowed with the relative topology inherited from $\Delta([\mathcal{K}(S)]^\infty)$.

Step 2. $P \longmapsto \int \hat{f} dP$ from $\Delta([\mathcal{K}(S)]^\infty)$ to $\mathbb{R}$ is analytically measurable for any bounded analytically measurable function $\hat{f}$ on $[\mathcal{K}(S)]^\infty$: If $\hat{f}$ is simple (has a finite number of values), then $P \longmapsto \int \hat{f} dP$ is analytically measurable by [3, p. 169]. More generally, $\int \hat{f} dP$ equals the pointwise limit of $\lim \int \hat{f}_n dP$ for some simple and analytically measurable $\hat{f}_n$, which implies the desired measurability.

Step 3. Note that

$$\left\{K \in \mathcal{K} : \inf_{\omega \in K} f(\omega) \geq t\right\} = \{K \in \mathcal{K} : K \subset \{\omega : f(\omega) \geq t\}\} \tag{A.1}$$

is co-analytic by [23, p. 772], and hence analytically measurable.

Steps 1, 2 and 3 complete the proof. ■

For Theorem 4.1, we show (iii)⇒(ii)⇒(i)⇒(iii). If $\nu \in Bel$, let $m = \zeta(\nu)$. We use (2.5) repeatedly without reference.

(iii)⇒(ii): Let Σ' be the σ-algebra generated by the class

$$\{K \in \mathcal{K} : K \subset A\}_{A \in \Sigma}.$$

We claim that $m(\cdot) = \int_{Bel(S)} \zeta(\theta^\infty)(\cdot)\, d\mu(\theta)$ on Σ'. Since the latter is a probability measure on $\mathcal{K}$, it is enough to show that

$$m(\{K \in \mathcal{K} : K \subset A\}) = \int_{Bel(S)} \zeta(\theta^\infty)(\{K \in \mathcal{K} : K \subset A\})\, d\mu(\theta)$$

for each $A \in \Sigma$. This is equivalent to

$$\nu(A) = \int_{Bel(S)} \theta^\infty(A)\, d\mu(\theta),$$

which is true given (iii).

By a standard argument using the Lebesgue Dominated Convergence Theorem,

$$\int_{\mathcal{K}} \hat{f} dm = \int_{Bel(S)} \left(\int_{\mathcal{K}} \hat{f} d\zeta(\theta^\infty) \right) d\mu(\theta),$$

for all Σ'-measurable $\hat{f} : \mathcal{K} \to [0,1]$. Since $K \longmapsto \inf_{\omega \in K} f(\omega)$ is Σ'-measurable by (A.1),

$$\begin{aligned} U_\nu(f) &= \int_{\mathcal{K}} \inf_{\omega \in K} f(\omega)\, dm(K) = \int_{Bel(S)} \left(\int_{\mathcal{K}} \inf_{\omega \in K} f(\omega)\, d\zeta(\theta^\infty) \right) d\mu(\theta) \\ &= \int_{Bel(S)} V_{\theta^\infty}(f)\, d\mu(\theta). \end{aligned}$$

(ii)⇒(i): It is enough to show that V_{θ^∞} satisfies Symmetry and WOI. Let $m = \zeta(\theta^\infty) = (\zeta(\theta))^\infty$. Then, m is an i.i.d. measure on $[\mathcal{K}(S)]^\infty$. Since m is symmetric,

$$\begin{aligned} V_{\theta^\infty}(\pi f) &= \int_{\mathcal{K}} \inf_{\omega \in K} \pi f(\omega)\, dm(K) = \int_{\mathcal{K}} \inf_{\omega \in K} f(\pi\omega)\, dm(K) \\ &= \int_{\mathcal{K}} \inf_{\pi\omega \in \pi K} f(\pi\omega)\, dm(K) = \int_{\mathcal{K}} \inf_{\omega \in K} f(\omega)\, d(\pi m)(K) \\ &= \int_{\mathcal{K}} \inf_{\omega \in K} f(\omega)\, dm(K) = V_{\theta^\infty}(f). \end{aligned}$$

Show (3.1) to prove WOI. For simplicity, let $f \in \mathcal{F}_1$ and $g \in \mathcal{F}_2$. The general case is similar. For $0 < \alpha \leq 1$,

$$
\begin{aligned}
& V_{\theta^\infty}\left(\alpha f + (1-\alpha) g\right) \\
= & \int_{\mathcal{K}} \inf_{\omega \in K} \left[\alpha f(\omega) + (1-\alpha) g(\omega)\right] dm(K) \\
= & \int_{[\mathcal{K}(S)]^\infty} \inf_{s_1 \in K_1, s_2 \in K_2} \left[\alpha f(s_1) + (1-\alpha) g(s_2)\right] dm(K_1, K_2, ...)
\end{aligned}
$$

$$
\begin{aligned}
= & \int_{[\mathcal{K}(S)]^\infty} \alpha \left[\inf_{s_1 \in K_1} f(s_1)\right] + (1-\alpha)\left[\inf_{s_2 \in K_2} (1-\alpha) g(s_2)\right] dm(K_1, K_2, ...) \\
= & \ \alpha \int_{[\mathcal{K}(S)]^\infty} \left[\inf_{s_1 \in K_1} f(s_1)\right] dm(K_1, K_2, ...) \\
& + (1-\alpha) \int_{[\mathcal{K}(S)]^\infty} \left[\inf_{s_2 \in K_2} g(s_2)\right] dm(K_1, K_2, ...) \\
= & \ \alpha V_{\theta^\infty}(f) + (1-\alpha) V_{\theta^\infty}(g).
\end{aligned}
$$

The second equality follows because $K \in [\mathcal{K}(S)]^\infty$, *a.s.*-$m[K]$.

(i)⇒(iii): For $C \subset \mathcal{K}$, let $\pi C = \{\pi K \in \mathcal{K} : K \in C\}$, and for $m \in \Delta(\mathcal{K})$, define $\pi m \in \Delta(\mathcal{K})$ by $\pi m(C) = m(\pi C)$ for each $C \in \Sigma_{\mathcal{K}}$.

Lemma A.2. *For any $m \in \Delta(\mathcal{K})$, $m = \pi m$ for all π if and only if $m = \zeta(\nu)$ for some symmetric belief function ν on Ω.*

Proof. If $m = \zeta(\nu)$, then $\nu(K) = m(\{K' \in \mathcal{K} : K' \subset K\})$, and

$$
\begin{aligned}
\nu(\pi K) &= m(\{K' \in \mathcal{K} : K' \subset \pi K\}) = m(\{\pi K' \in \mathcal{K} : \pi K' \subset \pi K\}) \\
&= m(\{\pi K' \in \mathcal{K} : K' \subset K\}) = m(\pi(\{K' \in \mathcal{K} : K' \subset K\})).
\end{aligned}
$$

The asserted equivalence follows, because the class $\{K' \in \mathcal{K} : K' \subset K\}_{K \in \mathcal{K}}$ generates the Borel σ-algebra on $\mathcal{K}$. ■

Lemma A.3. *Let ν be a belief function on S^∞ and $m = \zeta(\nu)$ the corresponding measure on $\mathcal{K}(S^\infty)$. If U_ν satisfies WOI, then $m[(\mathcal{K}(S))^\infty] = 1$.*

Proof. For any $\omega \in S^{\infty}$ and disjoint sets $I, J \subset \mathbb{N}$, ω_I denotes the projection of ω onto S^I, and we write $\omega = (\omega_I, \omega_J, \omega_{-I-J})$. When $I = \{i\}$, we write ω_i, rather than $\omega_{\{i\}}$, to denote the i-th component of ω.

Let $\mathcal{A}$ be the collection of compact subsets K of S^{∞} satisfying: For any $n > 0$, and $\omega^1, \omega^2 \in K$, and for every partition $\{1, ..., n\} = I \cup J$,

$$\exists \omega^* \in K, \text{ such that } \omega^*_I = \omega^1_I \text{ and } \omega^*_J = \omega^2_J. \tag{A.2}$$

In other words, for every n, the projection of K onto S^n is a Cartesian product.

Step 1. For any *continuous* acts $f \in \mathcal{F}_I$ and $g \in \mathcal{F}_J$ with finite disjoint I and J,

$$\min_{\omega \in K} \left[\tfrac{1}{2} f(\omega) + \tfrac{1}{2} g(\omega)\right] = \tfrac{1}{2} \min_{\omega \in K} f(\omega) + \tfrac{1}{2} \min_{\omega \in K} g(\omega), \tag{A.3}$$

a.s.-$m[K]$: This is where WOI enters - by (3.1) it implies that

$$U_\nu\left(\tfrac{1}{2} f + \tfrac{1}{2} g\right) = \tfrac{1}{2} U_\nu(f) + \tfrac{1}{2} U_\nu(g).$$

Since $U_\nu(f) = \int_{\mathcal{K}} \inf_{\omega \in K} f(\omega)\, dm(K)$,

$$\int_{\mathcal{K}} \min_{\omega \in K} \left[\tfrac{1}{2} f(\omega) + \tfrac{1}{2} g(\omega)\right] dm(K) = \tfrac{1}{2} \int_{\mathcal{K}} \min_{\omega \in K} f(\omega)\, dm(K) + \tfrac{1}{2} \int_{\mathcal{K}} \min_{\omega \in K} g(\omega)\, dm(K).$$

The assertion follows from

$$\min_{\omega \in K} \left[\tfrac{1}{2} f(\omega) + \tfrac{1}{2} g(\omega)\right] \geq \tfrac{1}{2} \min_{\omega \in K} f(\omega) + \tfrac{1}{2} \min_{\omega \in K} g(\omega).$$

Let $\mathcal{G}$ be the set of all pairs (f, g) such that f and g are continuous and $f \in \mathcal{F}_I$, $g \in \mathcal{F}_J$ for some finite disjoint I and J. Let $\mathcal{B}_{f,g}$ be the collection of $K \in \mathcal{K}$ satisfying (A.3), given f and g. Step 1 implies $m(\mathcal{B}_{f,g}) = 1$ for each $(f, g) \in \mathcal{G}$.

Step 2. $m\left(\bigcap_{(f,g) \in \mathcal{G}} \mathcal{B}_{f,g}\right) = 1$: Since the set of continuous finitely-based acts is separable under the sup-norm topology (see [2, Lemma 3.99]), it is easy to see that $\mathcal{G}$ is also separable. Let $\{(f_n, g_n)\}$ be a countable dense subset of $\mathcal{G}$. By Step 1,

$$m\left(\mathcal{K} \backslash \left(\bigcap_{i=1}^{\infty} \mathcal{B}_{f_i, g_i}\right)\right) = m\left(\bigcup_{i=1}^{\infty} (\mathcal{K} \backslash \mathcal{B}_{f_i, g_i})\right) \leq \sum m(\mathcal{K} \backslash \mathcal{B}_{f_i, g_i}) = 0.$$

Thus it is enough to show that $\bigcap_{i=1}^{\infty} \mathcal{B}_{f_i,g_i} = \bigcap_{(f,g)\in\mathcal{G}} \mathcal{B}_{f,g}$.

Only $\subset$ requires proof. Let $K \in \bigcap_{i=1}^{\infty} \mathcal{B}_{f_i,g_i}$, $(f,g) \in \mathcal{G}$ and assume wlog that $(f_i, g_i) \to (f,g)$. Then, by the Maximum Theorem [2, Theorem 17.31],

$$\begin{aligned}
\min_{\omega\in K} \left[\tfrac{1}{2} f(\omega) + \tfrac{1}{2} g(\omega)\right] &= \lim_i \min_{\omega\in K} \left[\tfrac{1}{2} f_i(\omega) + \tfrac{1}{2} g_i(\omega)\right] \\
&= \lim_i \left[\tfrac{1}{2} \min_{\omega\in K} f_i(\omega) + \tfrac{1}{2} \min_{\omega\in K} g_i(\omega)\right] \\
&= \tfrac{1}{2} \min_{\omega\in K} f(\omega) + \tfrac{1}{2} \min_{\omega\in K} g(\omega).
\end{aligned}$$

Thus $K \in \bigcap_{(f,g)\in\mathcal{G}} \mathcal{B}_{f,g}$.

Step 3. If $K \in \bigcap_{(f,g)\in\mathcal{G}} \mathcal{B}_{f,g}$, then $K \in \mathcal{A}$: Let $n \geq 0$, $\omega^1, \omega^2 \in K$ and $\{1, ..., n\} = I \cup J$, with I and J disjoint. For each i, take closed sets

$$\begin{aligned}
A_i &= \left\{\omega : \sum_{t\in I} 2^{-t} d\left(\omega_t, \omega_t^1\right) \geq \frac{1}{i}\right\} \text{ and} \\
B_i &= \left\{\omega : \sum_{t\in J} 2^{-t} d\left(\omega_t, \omega_t^2\right) \geq \frac{1}{i}\right\},
\end{aligned}$$

where $d(\cdot,\cdot)$ is the metric on S. By Urysohn's Lemma, there are continuous functions f_i and g_i such that, for each i,

$$\begin{aligned}
f_i(\omega) &= 1 \text{ if } \omega \in A_i \text{ and } 0 \text{ if } \omega_I = \omega_I^1, \text{ and} \\
g_i(\omega) &= 1 \text{ if } \omega \in B_i \text{ and } 0 \text{ if } \omega_J = \omega_J^2.
\end{aligned}$$

Since $A_i \in \Sigma_I$ and $B_i \in \Sigma_J$, we can take $f_i \in \mathcal{F}_I$, and $g_i \in \mathcal{F}_J$. Then, $\min_{\omega\in K} f_i(\omega) = \min_{\omega\in K} g_i(\omega) = 0$ and, since $K \in \mathcal{B}_{f_i,g_i}$,

$$\min_{\omega\in K} \left[f_i(\omega) + g_i(\omega)\right] = 0.$$

Hence, there exists $\hat{\omega}^i \in K$ such that $f_i\left(\hat{\omega}^i\right) = g_i\left(\hat{\omega}^i\right) = 0$. By the construction of f_i and g_i, we have $\hat{\omega}^i \notin A_i, B_i$, which implies

$$\sum_{t\in I} 2^{-t} d\left(\hat{\omega}_t^i, \omega_t^1\right) + \sum_{t\in J} 2^{-t} d\left(\hat{\omega}_t^i, \omega_t^2\right) < \frac{2}{i}.$$

Since $\{\hat{\omega}^i\} \subset K$ and K is compact, there is a limit point $\omega^* \in K$ satisfying (A.2).

Step 4. $m(\mathcal{A}) = 1$: By Steps 2-3, $1 \geq m(\mathcal{A}) \geq m\left(\bigcap_{(f,g)\in\mathcal{G}} \mathcal{B}_{f,g}\right) = 1$.

Step 5. $\mathcal{A} = (\mathcal{K}(S))^\infty$: Clearly $\mathcal{A} \supset (\mathcal{K}(S))^\infty$. For the other direction, take $K \in \mathcal{A}$ and assume $\omega^1, \omega^2, ... \in K$. It suffices to show that

$$\omega^* = \left(\omega_1^1, \omega_2^2, ..., \omega_n^n, ...\right) \in K. \tag{A.4}$$

Since $K \in \mathcal{A}$ and $\omega^1, \omega^2 \in K$, there exists $\hat{\omega}^2 \in K$ such that $\left(\hat{\omega}_1^2, \hat{\omega}_2^2\right) = (\omega_1^1, \omega_2^2)$. Similarly, since $\hat{\omega}^2, \omega^3 \in K$, there exists $\hat{\omega}^3 \in K$ such that $\left(\hat{\omega}_1^3, \hat{\omega}_2^3, \hat{\omega}_3^3\right) = \left(\hat{\omega}_1^2, \hat{\omega}_2^2, \omega_3^3\right) = (\omega_1^1, \omega_2^2, \omega_3^3)$, and so on, giving a sequence $\{\hat{\omega}^n\}$ in K. Any limit point ω^* satisfies (A.4). ■

Finally, we prove (i)⇒(ii). Let ν be a belief function on S^∞ and suppose that U_ν satisfies Symmetry and WOI. By Lemma A.3, $m \equiv \zeta(\nu)$ can be viewed as a measure on $[\mathcal{K}(S)]^\infty$, and by Lemma A.2, m is symmetric. Thus we can apply de Finetti's Theorem [15] to m, viewing $\mathcal{K}(S)$ as the one-period state space, to obtain: There exists $\hat{\mu} \in \Delta(\Delta(\mathcal{K}(S)))$ such that

$$m(C) = \int_{\Delta(\mathcal{K}(S))} \ell^\infty(C)\, d\hat{\mu}(\ell) \text{ for all } C \in \Sigma_{[\mathcal{K}(S)]^\infty}.$$

Here each ℓ lies in $\Delta(\mathcal{K}(S))$ and ℓ^∞ is the i.i.d. product measure on $[\mathcal{K}(S)]^\infty$. Extend each measure ℓ^∞ to $\Sigma_\mathcal{K}$ and write

$$m(C) = \int_{\Delta(\mathcal{K}(S))} \ell^\infty(C)\, d\hat{\mu}(\ell) \text{ for all } C \in \Sigma_{\mathcal{K}(S^\infty)}.$$

We claim that the equation extends also to $C \in \Sigma'$, where Σ' is the σ-algebra generated by the class

$$\{K \in \mathcal{K} : K \subset A\}_{A\in\Sigma}.$$

First, note that $\ell \longmapsto \ell^\infty(C)$ is universally measurable by Lemma A.1, and hence the integral is well-defined. By a standard argument using the Lebesgue Dominated Convergence Theorem, $C \longmapsto \int_{\Delta(\mathcal{K}(S))} \ell^\infty(C)\, d\hat{\mu}(\ell)$ is countably additive on Σ'. This completes the argument because m has a unique extension to the σ-algebra of universally measurable sets, and the latter contains Σ'.

Let $\mu \equiv \hat{\mu} \circ \zeta \in \Delta(Bel(S))$ and apply the Change of Variables Theorem to derive, for any $A \in \Sigma$,

$$\begin{aligned}
\nu(A) &= m(\{K \in \mathcal{K} : K \subset A\}) \\
&= \int_{\Delta(\mathcal{K}(S))} \ell^{\infty}(\{K \in \mathcal{K} : K \subset A\})\, d\hat{\mu}(\ell) \\
&= \int_{\Delta(\mathcal{K}(S))} \ell^{\infty}(\{K \in \mathcal{K} : K \subset A\})\, d\mu \circ \zeta^{-1}(\ell)
\end{aligned}$$

$$\begin{aligned}
&= \int_{Bel(S)} [\zeta(\theta)]^{\infty}(\{K \in \mathcal{K} : K \subset A\})\, d\mu(\theta) \\
&= \int_{Bel(S)} \zeta(\theta^{\infty})(\{K \in \mathcal{K} : K \subset A\})\, d\mu(\theta) \\
&= \int_{Bel(S)} \theta^{\infty}(A)\, d\mu(\theta).
\end{aligned}$$

Uniqueness of μ follows from the uniqueness of $\hat{\mu}$ provided by de Finetti's Theorem. ■

B. Appendix: Proof of Corollary 4.2

(i) $\Longrightarrow$ (ii): In the proof of Theorem 4.1, we showed that each $V_{\theta^{\infty}}$ satisfies Symmetry and WOI. The argument used there is readily adapted to prove: for all finite and disjoint I and J, subsets of $\mathbb{N}$, and for all $f \in \mathcal{F}_I$ and $g \in \mathcal{F}_J$,

$$V_{\theta^{\infty}}(f \cdot g) = V_{\theta^{\infty}}(f)\, V_{\theta^{\infty}}(g). \tag{B.1}$$

Alternatively, the latter follows from (4.4).

(ii) $\Longrightarrow$(i): By the representation,

$$U(f \cdot \varkappa f) = \int V_{\theta^{\infty}}(f \cdot \varkappa f)\, d\mu(\theta) = \int V_{\theta^{\infty}}(f)\, V_{\theta^{\infty}}(\varkappa f)\, d\mu(V)$$

$$= \int (V_{\theta^{\infty}}(f))^2\, d\mu(\theta) \geq \left(\int V_{\theta^{\infty}}(f)\, d\mu(\theta)\right)^2 = (U(f))^2 = p^2,$$

where we use the fact that every $V_{\theta^{\infty}}$ is symmetric (and hence also invariant to shifts) and satisfies (B.1), and also the familiar property that the geometric average is at least as large as the arithmetic average. Evidently, the indifference in (4.6) implies that

$$V_{\theta^\infty}(f) \text{ is constant } \mu\text{-}a.s.[\theta]. \tag{B.2}$$

The exceptional set depends on f. Since the set of continuous acts in $\mathcal{F}_1$ is separable under the sup-norm topology [2, Lemma 3.99], there exists a μ-null set of θ's that works for all continuous acts in $\mathcal{F}_1$ and also for all upper semicontinuous acts in $\mathcal{F}_1$. Since any belief function utility is regular in the sense of [8], the same μ-null set of θ's works for all acts. ■

C. Appendix: Proofs for Section 5

Proof of Corollary 5.1: Since $\theta^\infty(A) = \theta(A)$ for $A \in \Sigma_S$, $\theta \longmapsto \theta(A)$ is universally measurable by Lemma A.1. Hence, every set of the form

$$\{\theta \in Bel(S) : [\theta(A), 1 - \theta(S \backslash A)] \subset [a, b]\}$$

is universally measurable and the statement of the Corollary is well-defined.

We need two lemmas. Recall that $\Psi_n(A)(\omega) = \frac{1}{n}\sum_{i=1}^n I(\omega_i \in A)$ where ω_i is the i-th component of ω. Similarly define $\widehat{\Psi}_n(A)(K) = \frac{1}{n}\sum_{i=1}^n I(K_i \subset A)$ for $K \in [\mathcal{K}(S)]^\infty$, where K_i is the i-th component of K.

Lemma C.1. *Let $K \in [\mathcal{K}(S)]^\infty$, $K = K_1 \times K_2 \times \ldots$, and $\alpha \in \mathbb{R}$. Then the following are equivalent:*
(i) $\liminf_n \Psi_n(A)(\omega) > \alpha$ *for every* $\omega_i \in K_i$, $i = 1, \ldots$
(ii) $\liminf_n \widehat{\Psi}_n(A)(K) > \alpha$.

Proof. (i)⇒(ii): If $K_i \subset A$, let ω_i be any element in K_i, and otherwise, let ω_i be any element in $K_i \backslash A$. Then, $I(K_i \subset A) = I(\omega_i \in A)$ and thus (ii) is implied.

(ii)⇒(i): If $\omega_i \in K_i$, $I(K_i \subset A) \leq I(\omega_i \in A)$. Thus, if $\omega_i \in K_i$ for $i = 1, \ldots$, then,

$$\liminf_n \Psi_n(A)(\omega) \geq \liminf_n \widehat{\Psi}_n(A)(K) > \alpha.$$

■

Lemma C.2. *(i)* $\theta^{\infty}\left(\{\omega : \theta(A) < \liminf_n \Psi_n(A)(\omega)\}\right) = 0$ *for each* $A \in \Sigma_S$*; and (ii)* $\theta^{\infty}\left(\{\omega : \limsup_n \Psi_n(A)(\omega) < 1 - \theta(S \backslash A)\}\right) = 0$ *for each* $A \in \Sigma_S$.

Proof. Fix $A \in \Sigma_S$. Then,

$$
\begin{aligned}
& \theta^{\infty}\left(\left\{\omega : \theta(A) < \liminf_n \Psi_n(A)(\omega)\right\}\right) \\
= \; & [\zeta(\theta)]^{\infty}\left(\left\{K \in [\mathcal{K}(S)]^{\infty} : K \subset \left\{\omega : \theta(A) < \liminf_n \Psi_n(A)(\omega)\right\}\right\}\right) \\
= \; & [\zeta(\theta)]^{\infty}\left(\left\{K \in [\mathcal{K}(S)]^{\infty} : \liminf_n \widehat{\Psi}_n(A)(K) > \theta(A)\right\}\right) \quad \text{(by Lemma C.1).}
\end{aligned}
$$

By the Law of Large Numbers, $\widehat{\Psi}_n(A)(K)$ converges to $\zeta(\theta)(\{K_1 \in \mathcal{K}(S) : K_1 \subset A\}) = \theta(A)$ almost surely-$[\zeta(\theta)]^{\infty}$, which implies (i). The proof of (ii) is similar. ■

Return to the Corollary. By the LLN in [18], Lemma C.2 and the monotonicity of belief functions,

$$
\begin{aligned}
& \theta^{\infty}\left(\{\omega : [\liminf \Psi_n(A)(\omega), \limsup \Psi_n(A)(\omega)] \subset [a,b]\}\right) = 1 \\
\Leftrightarrow \; & [\theta(A), 1 - \theta(S \backslash A)] \subset [a,b]
\end{aligned}
$$

and

$$
\begin{aligned}
& \theta^{\infty}\left(\{\omega : [\liminf \Psi_n(A)(\omega), \limsup \Psi_n(A)(\omega)] \subset [a,b]\}\right) = 0 \\
\Leftrightarrow \; & [\theta(A), 1 - \theta(S \backslash A)] \text{ is not a subset of } [a,b].
\end{aligned}
$$

Moreover, for any belief function γ on Ω, if $\gamma(A) = \gamma(B) = 1$, then $\gamma(A \cap B) = 1$ by the Choquet theorem. Therefore,

$$
\begin{aligned}
& \nu\left(\bigcap_{i=1}^{I} \{\omega : [\liminf \Psi_n(A_i)(\omega), \limsup \Psi_n(A_i)(\omega)] \subset [a_i, b_i]\}\right) \\
= \; & \int_{Bel(S)} \theta^{\infty}\left(\bigcap_{i=1}^{I} \{\omega : [\liminf \Psi_n(A_i)(\omega), \limsup \Psi_n(A_i)(\omega)] \subset [a_i, b_i]\}\right) d\mu(\theta) \\
= \; & \mu\left(\bigcap_{i=1}^{I} \{\theta : [\theta(A_i), 1 - \theta(S \backslash A_i)] \subset [a_i, b_i]\}\right). \qquad ■
\end{aligned}
$$

Proof of Proposition 5.2: By exploiting the homeomorphism defined in the Choquet Theorem, we can identify μ' and μ with measures on $\Delta(\mathcal{K}(S))$. Modulo this identification, we are given that μ' and μ agree on the collection of all subsets of $\Delta(\mathcal{K}(S))$ of the form

$$\bigcap_{i=1}^{I} \{\ell \in \Delta(\mathcal{K}(S)) : \ell(\{K \in \mathcal{K}(S) : K \subset A_i\}) \geq a_i\},$$

for all $I > 0$, $A_i \in \Sigma_S$ and $a_i \in [0,1]$. They necessarily agree also on the generated σ-algebra, denoted Σ^*. Therefore, it suffices to show that

$$\Sigma_{\Delta(\mathcal{K}(S))} \subset \Sigma^*,$$

where $\Sigma_{\Delta(\mathcal{K}(S))}$ is the Borel σ-algebra on $\Delta(\mathcal{K}(S))$.

Step 1. $\ell \longmapsto \ell(C)$ is Σ^*-measurable for measurable $C \in \Sigma_{\mathcal{K}(S)}$: Let $\mathcal{C}$ be the collection of measurable subsets C of $\mathcal{K}(S)$ such that $\ell \longmapsto \ell(C)$ is Σ^*-measurable. Every set of the form $\{K' \in \mathcal{K}(S) : K' \subset K\}$ for $K \in \mathcal{K}(S)$ lies in $\mathcal{C}$. Since the collection $\{K' \in \mathcal{K}(S) : K' \subset K\}_{K \in \mathcal{K}(S)}$ generates $\Sigma_{\mathcal{K}(S)}$, it is enough to show that $\mathcal{C}$ is a σ-algebra: (i) $C \in \mathcal{C}$ implies $\mathcal{K}(S) \backslash C \in \mathcal{C}$; (ii) if each $C_i \in \mathcal{C}$, then $\ell \longmapsto \ell(\cup_{i=1}^{\infty} C_i)$ is Σ^*-measurable because it equals the pointwise limit of $\ell \longmapsto \ell(\cup_{i=1}^{n} C_i)$ - hence $\cup_{i=1}^{\infty} C_i \in \mathcal{C}$.

Step 2. $\ell \longmapsto \int \hat{f} d\ell$ is Σ^*-measurable for all Borel-measurable $\hat{f}$ on $\mathcal{K}(S)$: Identical to Step 2 in Lemma A.1.

Step 3. $\Sigma_{\Delta(\mathcal{K}(S))} \subset \Sigma^*$: By Step 2, $\left\{\ell : \int \hat{f} d\ell \geq a\right\} \in \Sigma^*$ for all Borel-measurable $\hat{f}$ on $\mathcal{K}(S)$. But $\Sigma_{\Delta(\mathcal{K}(S))}$ is the smallest σ-algebra containing the sets $\left\{\ell : \int \hat{f} d\theta \geq a\right\}$ for all continuous $\hat{f}$ and $a \in \mathbb{R}$.

References

[1] N. Al-Najjar and L.De Castro, A subjective foundation of objective probability, 2009.

[2] C.D. Aliprantis and K.C. Border, *Infinite Dimensional Analysis*, 3rd edition, Springer, 2006.

[3] D.P. Bertsekas and S.E. Shreve, *Stochastic Optimal Control: The Discrete Time Case*, Academic Press, New York, 1978.

[4] A. Castaldo, F. Maccheroni and M. Marinacci, Random correspondences as bundles of random variables, *Sankhya* (Series A) 80 (2004), 409-427.

[5] G. Choquet, Theory of capacities, *Ann. Inst. Fourier* 5 (1954), 131-295.

[6] A.P. Dempster, Upper and lower probabilities induced by a multi-valued mapping, *Ann. Math. Statist.* 38 (1967), 325-339.

[7] L.G. Epstein, M. Marinacci and K. Seo, Coarse contingencies and ambiguity, *Theor. Econ.* 2 (2007), 355-394.

[8] L.G. Epstein and K. Seo, Symmetry of evidence without evidence of symmetry, *Theor. Econ.*, forthcoming.

[9] L.G. Epstein and T. Wang, Intertemporal asset pricing under Knightian uncertainty, *Econometrica* 62 (1994), 283-322.

[10] B. de Finetti, La prevision: ses lois logiques, ses sources subjectives. *Ann. Inst. H. Poincare* 7 (1937), 1-68. English translation in *Studies in Subjective Probability*, 2nd edition, H.E. Kyburg and H.E. Smokler eds., Krieger Publishing, Huntington NY, 1980, pp. 53-118.

[11] P. Ghirardato, Coping with ignorance, unforeseen contingencies and nonadditive uncertainty, *Econ. Theory* 17 (2001), 247-276.

[12] P. Ghirardato, On independence for non-additive measure, with a Fubini theorem, *J. Econ. Theory* 73 (1997), 261-291.

[13] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.

[14] E. Hendon, H.J. Jacobson. B. Sloth and T. Tranaes, The product of capacities and belief functions, *Math. Soc. Sc.* 32 (1996), 95-108.

[15] E. Hewitt and L.J. Savage, Symmetric measures on Cartesian products, *Trans. Amer. Math. Soc.* 80 (1955), 470-501.

[16] P.J. Huber, *Robust Statistics*, Wiley, New York, 1981.

[17] D.M. Kreps, *Notes on the Theory of Choice*, Westview, 1988.

[18] F. Maccheroni and M. Marinacci, A strong law of large numbers for capacities, *Ann. Prob.* 33 (2005), 1171-1178.

[19] I. Molchanov, *Theory of Random Sets*, Springer, 2005.

[20] S. Mukerji, Understanding the nonadditive probability decision model, *Econ. Theory* 9 (1997), 23-46.

[21] H.T. Nguyen, On random sets and belief functions, *J. Math. Anal. and Appl.* 65 ((1978), 531-542.

[22] R. Phelps, *Lectures on Choquet's Theorem*, Springer, 2001.

[23] F. Philippe, G. Debs and J.Y. Jaffray, Decision making with monotone lower probabilities of infinite order, *Math. Oper. Res.* 24 (1999), 767-784.

[24] L. Savage, *The Foundations of Statistics*, Dover, 1972.

[25] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* 57 (1989), 571-587.

[26] G. Shafer, *A Mathematical Theory of Evidence*, Princeton U. Press, Princeton, 1976.

[27] G. Shafer, Belief functions and parametric models (with commentary), *J. Royal Stat. Soc. Series B* 44 (1982), 322-352.

[28] M. Talagrand, Capacités invariantes extrémales, *Ann. Inst. Fourier* (Grenoble) 28 (1978), 79-146.