

# SYMMETRY OF EVIDENCE WITHOUT EVIDENCE OF SYMMETRY\*

Larry G. Epstein      Kyoungwon Seo

December 4, 2008

## Abstract

The de Finetti Theorem is a cornerstone of the Bayesian approach. Bernardo [4, p. 5] writes that its “message is very clear: if a sequence of observations is judged to be exchangeable, then any subset of them must be regarded as a random sample from some model, and there exists a prior distribution on the parameter of such model, hence *requiring* a Bayesian approach.” We argue that while exchangeability, interpreted as symmetry of evidence, is a weak assumption, when combined with subjective expected utility theory, it implies also complete confidence that experiments are identical. When evidence is sparse, and there is little evidence of symmetry, this implication of de Finetti’s hypotheses is not intuitive. We adopt multiple-priors utility as the benchmark model of preference and generalize the de Finetti Theorem to this framework. The resulting model also features a “conditionally IID” representation, but it differs from de Finetti in permitting the degree of confidence in the evidence of symmetry to be subjective.

## 1. INTRODUCTION

### 1.1. Motivation and Objectives

An individual is confronted with a sequence of possibly biased coins to be tossed in turn. Information about the coins is symmetric. How would she rank bets on the outcomes? The usual way to model this situation is to assume subjective expected utility (SEU) theory with a suitable prior over the possible sequences of outcomes. Further, it is assumed that the individual does not attach any weight to the fact that many coins are involved - she ranks bets just as though a single coin were to be tossed repeatedly. Consequently, she would not view outcomes across stages as independent. De Finetti identified *exchangeability* as the natural property of beliefs in such settings - a probability measure on the set

---

\*Boston University, lepstein@bu.edu and Department of Managerial Economics and Decision Sciences, Northwestern University, k-seo@kellogg.northwestern.edu. We are grateful to Asen Kochov, Fabio Maccheroni, Massimo Marinacci, Uzi Segal, participants at RUD 08 at Oxford, and especially to Martin Schneider for helpful comments and discussions, and to Yi-Chun Chen for providing us with his unpublished work which helped us to better understand “regularity” in the multiple-priors model. Epstein acknowledges also the financial support of NSF (award SES-0611456).

of sequences of outcomes is exchangeable if the probability of any finite set of outcomes depends only on the numbers of Heads and Tails and not on their order. Moreover, de Finetti shows that this property of beliefs is equivalent to the view that outcomes are “conditionally i.i.d.” - this is his celebrated theorem (see [16, 21]).

An exchangeable prior is often justified on the grounds of *symmetry of evidence* - no information is given that would *imply* a distinction between the first coin and any other. However, if little information is provided about any of the coins, in which case there is little *evidence of symmetry*, then a thoughtful and cautious individual might very well *admit the possibility* that the coins may differ in some way, and this may influence her ranking of bets. The distinction between the two forms of symmetry is due to Walley [33], who also argued that this distinction is behaviorally meaningful and that it cannot be accommodated within the Bayesian framework.

Naturally, coin tossing is but one example of what we have in mind. More generally imagine a countable number of experiments. Information about the experiments is symmetric. *We aim to capture the following perception: there exist some factors affecting all experiments (like the physical make-up of a fixed coin in a sequence of tosses), but there are also idiosyncratic and poorly understood factors that potentially can vary across experiments in an unrelated or independent way.*<sup>1</sup>

There are different levels at which one can see that the Bayesian model *cannot* capture intuitively a situation where evidence is symmetric but where the individual cares that there is little or no evidence of symmetry. First, symmetry of evidence suggests an exchangeable prior, which leaves no room to capture the perception of possible differences between coins. Second, de Finetti shows that an exchangeable prior admits the following interpretation: though the individual is uncertain about the bias of each coin, she is certain that the coins are independent and have the *identical* bias. Third, and most significantly, behavior that is intuitive given a concern with the lack of evidence of symmetry is excluded by the Bayesian (SEU) model. We elaborate on the behavioral critique in the next subsection.

However, the above distinction *can* be accommodated within the multiple-priors model (Gilboa and Schmeidler [18]) as we show in this paper. Our model is axiomatic - we add axioms to those put forth by Gilboa and Schmeidler and show that they capture alternative hypotheses about how the relationship between experiments is perceived (see Corollary 3.4 and Theorem 4.2). In the SEU special case, where the set of priors is a singleton, each result is equivalent to the classic de Finetti theorem. However, if SEU is not assumed, then they are not equivalent and they provide alternative generalizations of his theorem to the multiple-priors framework.

There are many examples, other than coin tossing, where the distinction we are discussing would seem to be relevant. Walley describes another concrete example (p. 462): experiment  $j$  refers to the age at which person  $j$  dies, where individuals are ordered alphabetically by name. If you are given no information about the individuals or how they were chosen, then evidence is symmetric and you would presumably be indifferent be-

---

<sup>1</sup>Since “independence” has a different meaning in an axiomatic context, below we often refer to the “unrelatedness” or “stochastic independence” of experiments, though the latter should not be understood in the usual sense of probability theory.

tween bets on  $j$  versus  $i$  living past 60. However, you may *not* be *certain* that individuals are identical in health, and so on, and feel that these other relevant factors are unrelated across individuals.

More generally, think of a statistician or empiricist as the decision-maker in our model and of an experiment as part of a statistical model of how data are generated. For example, the context could be that of the literature attempting to explain cross-country differences in growth rates, in which case an ‘experiment’ corresponds to a country. Invariably symmetry is assumed at some level - perhaps after correcting for perceived asymmetries, such as heteroscedasticity of errors in a regression model. Standard statistical methods presume that, after such corrections, the *identical* statistical model applies to all observations in the cross-country data set. Brock and Durlauf [5, p. 231] criticize this presumption and argue that “a major source of skepticism about the empirical growth literature, and one that incorporates many of the usual criticisms, is the failure of ... exchangeability to hold in conventional empirical growth exercises.” We view this paper as *potentially* providing decision-theoretic foundations for statistical procedures that would permit the analyst to express a judgement of “similarity” but also a concern that the relevant experiments are not identical.<sup>2</sup>

## 1.2. A Role for Randomization

Modify the coin-tossing example slightly and suppose that it is a single coin that is tossed repeatedly. Then outcomes are related through the coin’s inherent (or physical) bias, determined by its material composition and distribution. The individual believes that other time-varying factors may also play a role - for example, the way in which the coin is tossed. (The latter may vary because, for example, each toss is conducted by a different person.) She has a poor understanding of this latter factor - she does not know the relevant physics and she cannot even describe what is involved beyond the statement that the “manner of tossing can matter.” However, she is certain that, whatever is the coin’s true physical bias, the other factors influence outcomes of the various tosses “independently” - for example, if the manner of the first toss inadvertently biases the outcome towards Heads, nothing is implied about the bias that might be imparted inadvertently in the second toss. Moreover, in light of the symmetry of information, she believes that the same set of possible influences is operative for each coin toss.

To illustrate why the Independence Axiom is not intuitive, it is enough to consider the special case where the physical bias is known with certainty. Then the coin tosses are stochastically independent, as in the special case of our model considered in [13]. The natural state space is  $S^\infty$ , where  $S = \{H, T\}$ . Symmetry suggests indifference between any bet and its image under a permutation of coins; in particular,

$$H_1T_2 \sim T_1H_2.$$

Here  $H_1T_2$  is the bet that pays 1 if the first toss yields heads and the second Tails; the bet  $T_1H_2$  is interpreted similarly. Payoffs are denominated in utils, as in the Anscombe-Aumann model, so that the agent is indifferent to risk in outcomes.

---

<sup>2</sup>“Potentially” is emphasized because we have nothing to offer here in the way of new statistical methods.

Consider now the choice between either of the above bets and  $\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2$ , the bet paying  $\frac{1}{2}$  if  $\{H_1T_2, T_1H_2\}$  and 0 otherwise. The Independence Axiom would imply that

$$\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2 \sim H_1T_2.$$

This is intuitive given certainty that the coins are identical, since then there is nothing to be gained by mixing; neither is there a cost to randomizing because the individual is risk neutral. On the other hand, if she admits the *possibility* that the coins are not identical, and her meagre information suggests that she may, then she may strictly prefer the mixture because the bets  $H_1T_2$  and  $T_1H_2$  hedge one another: the former pays well if coin toss 1 is biased towards Heads and coin toss 2 is biased towards Tails, pays poorly if the opposite bias pattern is valid, and these “good” and “bad” states are reversed for act  $T_1H_2$ . Thus  $\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2$  hedges uncertainty about the bias pattern perfectly, and as such, suggests the ranking

$$\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2 \succ H_1T_2 \sim T_1H_2. \tag{1.1}$$

It is important to realize that concern with the coins not being identical is *not a (probabilistic) risk* - if it were, then the individual’s risk neutrality would imply that there is no value to hedging the risk and hence to randomization. It is not possible to model the noted concern by using a single probability measure, since symmetry of information suggests immediately that the associated probability measure is exchangeable, leaving no room for possible differences between coins. This is the heart of Walley’s criticism of the exchangeable Bayesian model.

However, randomization *is* a matter of indifference for *some* bets, and precisely when such indifference prevails can be interpreted in terms of the individual’s perception of how experiments are related to one another. The extreme where the Independence Axiom is satisfied and where there is indifference to randomization for any pair of indifferent bets, we interpret as reflecting certainty that the experiments are identical. A weaker axiom formulated below, called Orthogonal Independence, is interpretable as reflecting a more modest perception. To illustrate, consider bets on a Head resulting from the first, or alternatively, from the second coin toss. Clearly,  $H_1 \sim H_2$  given the symmetry of evidence, but what about  $\frac{1}{2}H_1 + \frac{1}{2}H_2$ ? We claim that the indifference

$$\frac{1}{2}H_1 + \frac{1}{2}H_2 \sim H_1 \sim H_2$$

is intuitive. The reason is that, even though the payoff to  $H_1$  is uncertain, with the above perception  $H_2$  does not hedge this uncertainty. An unfavorable tossing technique for the first coin cannot be compensated by that for the second coin because of the perceived unrelatedness of technique across coins. Neither does it hedge uncertainty about the physical bias, if it exists - since the bias is identical for both tosses, mixing  $H_1$  and  $H_2$  does not moderate payoff uncertainty. Thus the mixture  $\frac{1}{2}H_1 + \frac{1}{2}H_2$  is no better than either component bet, and neither is it worse given risk neutrality.

Our central axioms build on the intuition just described. Their interpretations as capturing the perception of poorly understood factors that are unrelated are supported by the representations for utility and beliefs that they deliver.

### 1.3. Related Literature

Kreps [25, 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 - this is warranted if and only if samples are assessed *ex ante* as being exchangeable. As a result, and also because similarity judgements naturally play a central role in statistical analysis, the notion of exchangeability underlies much of common empirical practice. See Kingman [24] for an overview. For its role in: (i) empirical practice, and regression analysis in particular, see Arnold [2]; (ii) data analysis, see Draper *et al* [10]; (iii) predictive modeling and inference, see Draper [9] and Draper *et al* [10].

Bayesians often refer to exchangeability as a weak assumption. Schervish [30, p. 8] writes: “The motivation for the definition of exchangeability is to express symmetry of beliefs ... in the weakest possible way. The definition ... does not require any judgement of independence or that any limit of relative frequencies will exist. It merely says that the labeling of random quantities is immaterial.” We agree that “symmetry of beliefs”, in the sense of “symmetry of evidence”, is a weak assumption. Our objection is to the (implicit) companion hypothesis of SEU preferences. To improve upon exchangeability, Bayesians have proposed weaker notions (“partial exchangeability” is one example) that build in less symmetry, while maintaining SEU; see Schervish’s Ch. 8, for example. *Such extensions within the Bayesian framework do not permit the separate modeling of a concern with evidence of symmetry in an environment where evidence is symmetric.*

We have already acknowledged our debt to Walley [33] for the critique that motivates this paper and for the distinction that we have adopted as a title. His contribution to modeling the distinction is described briefly in Section 3.1.

Finally, Epstein and Schneider [13] model the distinction between symmetry of evidence and evidence of symmetry in the special case where experiments are viewed as being completely unrelated (in the context of the above example of repeated tosses of a single coin, they assume that the physical bias is known with certainty). They speak in terms of experiments that are “indistinguishable” as opposed to “identical”, a distinction that we adopt also below. In [14], those authors study the more general case dealt with here (unknown physical bias). There are two important differences from this paper. First, they describe functional forms and provide informal justification, partly through applications, while here the focus is on axiomatic foundations. Secondly, in both of the just cited papers the models are dynamic - experiments are conducted in a fixed temporal order, the individual updates beliefs and preferences in response to previously observed outcomes, and dynamic consistency is an important consideration. We ignore these issues here and deal with a formally atemporal setup - we model *ex ante* beliefs and preferences only. An advantage is that the analysis applies to cross-sectional experiments, such as cross-country growth patterns [5]. However, the question of how inference could or should be conducted is not addressed here.

## 2. PRELIMINARIES

### 2.1. The Bayesian Model

There exists a countable infinity of experiments - they are ordered and indexed by the set  $\mathbb{N} = \{1, 2, \dots\}$ . Each experiment yields an outcome in the finite set  $S$ . The set of possible outcomes for the  $i^{\text{th}}$  experiment is sometimes denoted  $S_i$ , though  $S_i = S$  for all  $i$ .

- $\Omega = S^\infty = S_1 \times S_2 \times \dots$
- $\mathcal{S}_i$ :  $\sigma$ -algebra on  $S_i$  (the power set given that  $S$  is finite), identified with a  $\sigma$ -algebra on  $\Omega$
- Consider  $(\Omega, \Sigma)$ , where  $\Sigma$  is the product  $\sigma$ -algebra,  $\Sigma = \sigma(\bigvee_{j=1}^\infty \mathcal{S}_j)$
- Probability measures on  $(\Omega, \Sigma)$  are understood to be countably additive unless specified otherwise

An *act* is a  $\Sigma$ -measurable function from  $\Omega$  into  $[0, 1]$ . For example, when  $S = \{H, T\}$ , then the act  $f$ ,

$$f(s_1, \dots, s_i, \dots) = \begin{cases} 1 & (s_1, s_2) = (H, T) \\ 0 & \text{otherwise,} \end{cases}$$

is the bet on Heads followed by Tails, which was denoted  $H_1T_2$  above (we use similar abbreviations below when referring to coin-tossing). Preference, denoted  $\succeq$ , is defined on the set  $\mathcal{F}$  of all acts.

Denote by  $\Pi$  the set of finite permutations of  $\mathbb{N}$ ; all permutations appearing in the paper should be understood to be finite. For any  $\pi$  in  $\Pi$  and probability measure  $P$  on  $(S^\infty, \Sigma)$ , define  $\pi P$  to be the unique probability measure on  $S^\infty$  satisfying (for all rectangles)

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

Given an act  $f$ , define the permuted act  $\pi f$  by  $(\pi f)(s_1, \dots, s_t, \dots) = f(s_{\pi(1)}, \dots, s_{\pi(t)}, \dots)$ . Abbreviate  $\int f dP$  by  $Pf$ , or  $P(f)$ . Then, for all  $P, f$  and  $\pi$ ,

$$(\pi P)f = P(\pi f).$$

The probability measure  $P$  is *exchangeable* if  $\pi P = P$  for all  $\pi$ . In behavioral terms, assuming subjective expected utility preference with prior  $P$ , exchangeability of  $P$  is equivalent to the universal indifference between an act and any permuted variant, that is,

$$f \sim \pi f \text{ for all acts } f \text{ and permutations } \pi.$$

For any probability measure  $\ell$  on  $S$  (write  $\ell \in \Delta(S)$ ),  $\ell^\infty$  denotes the corresponding i.i.d. product measure on  $(\Omega, \Sigma)$ .

**Theorem 2.1 (de Finetti).** *The probability measure  $P$  on  $(\Omega, \Sigma)$  is exchangeable if and only if there exists a Borel probability measure  $\mu$  on  $\Delta(S)$  such that*

$$P(\cdot) = \int_{\Delta(S)} \ell^\infty(\cdot) d\mu(\ell). \quad (2.1)$$

Moreover,  $\mu$  is unique.

A noteworthy and problematic feature of the framework, which we adopt also below, is that payoffs to acts depend on the outcomes of infinitely many experiments, which is problematic for a positive model. In particular, the domain of preference includes acts whose payoffs depend on the truth/falsity of tail events, which are not observed in finite time, and thus are, in fact, unobservable.<sup>3</sup> This concern was emphasized also by de Finetti; see [29] for extensive discussion of de Finetti's view, and also [11]. However, a decision-maker might be able to conceive of payoffs that depend on tail events (receive  $x^*$  if the limiting empirical frequency of Heads in an infinite sequence of tosses is greater than  $\frac{1}{2}$  and  $x$  otherwise). Thus the de Finetti theorem and its generalizations below seem useful in a normative context.<sup>4</sup>

Another objection to the de Finetti-Savage model is the one raised by Walley and described in the introduction - that symmetry of evidence in their model implies also that experiments are necessarily viewed as being identical. This is reflected in de Finetti's representation (2.1) which admits the following interpretation: though the individual is uncertain ex ante which likelihood function applies, she is certain that the same likelihood function applies to all experiments. Accommodating this critique is the objective of this paper, and is the reason that we move from SEU to the multiple-priors model.

## 2.2. Multiple-Priors Preference

Gilboa and Schmeidler [18] propose a specification for utility that is commonly written in the following form:

$$U(f) = \min_{P \in C} \int_{\Omega} u(f) dP. \quad (2.2)$$

Here: acts map states into  $\Delta(Z)$ , the set of lotteries over an underlying set of outcomes  $Z$ ;  $u : \Delta(Z) \rightarrow \mathbb{R}$  is affine and nonconstant; and  $C \subset ba_+^1(\Omega)$  is a convex and weak\*-compact set of finitely additive probability measures on  $(\Omega, \Sigma)$ .<sup>5</sup>

<sup>3</sup>The tail  $\sigma$ -algebra is defined by  $\Sigma^{tail} = \cap_{t=1}^{\infty} \sigma(\vee_{j=t}^{\infty} \mathcal{S}_j)$ .

<sup>4</sup>In fact, we have overstated the problem somewhat in as much as our central axioms concern only the ranking of acts that depend on finitely many experiments.

<sup>5</sup>For any compact metric space  $X$ ,  $ba(X)$  and  $ca(X)$  denote the spaces of finite variation set functions on the Borel  $\sigma$ -algebra that are finitely additive (charges) and countably additive respectively;  $ba_+^1(X)$  and  $ca_+^1(X)$  are the corresponding subsets of positive and normalized measures. The notation  $ca_+^1(X)$  is useful when it is important to draw a distinction between finitely and countably additive probability measures; otherwise, the simpler notation  $\Delta(X)$  is used. Unless otherwise specified,  $\Delta(X)$  is endowed with the weak-convergence topology induced by continuous functions, which renders it also compact metric. For any  $\mathcal{P} \subset \Delta(X)$ ,  $cl(\mathcal{P})$  denotes its weak-convergence closure. By the weak\* topology on  $ba(\Omega)$ , we mean the topology induced by bounded measurable functions. Finally,  $\mathcal{K}(X)$  denotes the space of compact subsets of  $X$ , endowed with the Hausdorff metric topology, which renders it compact metric. When  $X$  is a lts,  $\mathcal{K}^c(X)$  denotes the space of compact and convex subsets of  $X$ .

They prove that this model is characterized by a simple set of axioms. Most basic are the axioms of completeness, transitivity, a form of continuity and monotonicity, which we state here for convenience:

*MONOTONICITY*: If  $f(\omega) \geq g(\omega)$  for all  $\omega$ , abbreviated  $f \geq g$ , then  $f \succeq g$ .

Two core axioms reflect the way in which the model addresses aversion to ambiguity. Mixtures are defined in the familiar Anscombe-Aumann fashion:

$$(\alpha f + (1 - \alpha) g)(\omega) \equiv \alpha f(\omega) + (1 - \alpha) g(\omega).$$

*AMBIGUITY AVERSION*: If  $f \succeq g$ , then  $\alpha f + (1 - \alpha) g \succeq g$  for all  $\alpha$  in  $[0, 1]$ .

*CERTAINTY INDEPENDENCE*: For all  $\alpha$  in  $(0, 1)$ , acts  $f$  and  $g$  and constant acts  $x$ ,  $f \succeq g \iff f + (1 - \alpha)x \succeq \alpha g + (1 - \alpha)x$ .

Jointly, the two axioms weaken Independence. The intuition for them is well-known. Randomizing between acts is not harmful because, by its definition, randomization is linear in utils. It can be beneficial, however, contrary to Independence, if  $f$  and  $g$  hedge one another, in which case the mixture smooths outcomes across states and thus makes ambiguity about likelihoods less of a concern.

For completeness, we state:

*INDEPENDENCE*: For all  $\alpha$  in  $(0, 1)$ ,

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

For our benchmark model of preference, we specialize the Gilboa-Schmeidler model in two ways. First, because it greatly simplifies the sequel, we suppress  $Z$  and  $\Delta(Z)$  and measure outcomes instead in utils, which lie in the unit interval (as specified in the definition of an act given in the outline of the Bayesian model). This procedure is easily justified if it is assumed that there exist best and worst outcomes, so that (identifying outcomes with constant acts),

$$\underline{z} \preceq z \preceq \bar{z} \text{ for all } z.$$

Without loss of generality, let  $u(\underline{z}) = 0$  and  $u(\bar{z}) = 1$ . Then  $u(f(\omega))$  is the unique probability so that

$$f(\omega) \sim (\bar{z}, u(f(\omega)); \underline{z}, 1 - u(f(\omega))). \quad (2.3)$$

Such calibration renders the util-outcomes of any act observable, and these are the  $[0, 1]$ -valued outcomes we assume henceforth, and that justify rewriting utility in the form

$$U(f) = \min_{P \in C} \int_{\Omega} f dP = \min_{P \in C} Pf, \quad f \in \mathcal{F}. \quad (2.4)$$

The fact that *outcomes are “equivalent” probabilities* will be important below - multiplying outcomes, which may seem unnatural, will amount to the very natural operation

of multiplying probabilities. Observe finally that with the preceding interpretation,  $U(f)$  is also in probability units - it equals the unique probability such that

$$f \sim (\bar{z}, U(f); \underline{z}, 1 - U(f)). \quad (2.5)$$

This is to be contrasted with the more familiar certainty equivalent outcome of  $f$ , which, given (2.2), equals  $u^{-1}(\int_{\Omega} u(f)dP)$ .

The second specialization is more substantive, but has both a clear rationale and simple behavioral foundations. The rationale is that just as countable additivity is assumed almost universally in the central theorems of probability theory, including the de Finetti Theorem that concerns us here, we specialize multiple-priors utility to provide a counterpart of countable additivity.<sup>6</sup> The form this takes is to assume that the Gilboa-Schmeidler set of priors  $C$  is the weak\*-closure of a suitable set of countably additive priors, or equivalently, that there exists a convex set  $\mathcal{P}$ ,

$$\mathcal{P} \subset ca_+^1(\Omega),$$

such that utility can be written as

$$U(f) = \inf_{\mathcal{P}} Pf, \quad f \in \mathcal{F}. \quad (2.6)$$

Further, we assume that  $\mathcal{P}$  is compact in the weak-convergence topology on  $ca_+^1(\Omega)$ .

An alternative way to express “countable additivity” for a set of priors, that is put forth by Chateauneuf *et al* [7], is to assume that  $\mathcal{P}$  itself is weak\*-compact, and hence that  $C$  (equals  $\mathcal{P}$  and) consists exclusively of countably additive measures. Our assumption is weaker (weak\*-compactness implies weak-convergence compactness for any set of priors), but generality is not the reason for our choice of compactness assumption. We show in Section 4.5 that if one requires that sets of priors be weak\*-compact subsets of  $ca_+^1(\Omega)$ , then one cannot capture even the extreme case (that we refer to as IID below) where experiments are perceived to be stochastically independent in a natural sense, without preference collapsing to expected utility.

Turn to the foundations for the specification (2.6) with (weak-convergence) compact set  $\mathcal{P}$ . We need some additional notation. The set of all  $[0, 1]$ -valued acts on  $\Omega$  is  $\mathcal{F}$ . Denote by  $\mathcal{F}^u$  the set of all upper semicontinuous (usc) and simple (finite-ranged) acts, and by  $\mathcal{F}^l$  the set of lower semicontinuous (lsc) and simple acts. As shown above, under suitable conditions there is a unique probability-equivalent utility function  $U$ , defined in (2.5), that represents preference. Thus we can state the sought-after condition in terms of that utility function.<sup>7</sup>

---

<sup>6</sup>See [29] and [11, 12] for approaches assuming only finite additivity.

<sup>7</sup>We state Regularity for any utility function  $U$ . As shown in [15], the axiom is readily expressed explicitly in terms of preference for a large class of preferences.

*REGULARITY*: A utility function  $U : \mathcal{F} \rightarrow [0, 1]$ , and the corresponding preference order, are regular if both of the following conditions are satisfied:

**Inner Regularity**  $U(h) = \sup\{U(g) : g \leq h, g \in \mathcal{F}^u\}, \forall h \in \mathcal{F}^\ell$ ; and

**Outer Regularity**  $U(f) = \inf\{U(h) : h \geq f, h \in \mathcal{F}^\ell\}, \forall f \in \mathcal{F}$ .

There is an obvious parallel with the notion of regularity for a measure - think of the special case of acts that are indicator functions, and note that the indicator  $\mathbf{1}_A$  is simple and usc (lsc) if  $A$  is closed (open). This parallel inspired the closely related definition of regularity of preference due to Epstein and Wang [15].<sup>8</sup> The relation is that  $U$  is regular in the above sense if and only if its conjugate  $U^*$ ,

$$U^*(f) = 1 - U(1 - f), \quad f \in \mathcal{F},$$

is regular in the sense of [15]. For another perspective on the difference between the two definitions of regularity, observe that the Epstein-Wang notion requires that

$$U(f) = \sup\{U(g) : g \leq f, g \in \mathcal{F}^u\}, \forall f \in \mathcal{F},$$

that is, the utility of arbitrary acts can be approximated from below (by simple usc acts). In contrast, Outer Regularity above postulates that the utility of arbitrary acts can be approximated from above (by simple lsc acts). Approximation from above seems more intuitive given the conservatism inherent in aversion to ambiguity or to limited evidence. Since any probability measure coincides with its conjugate ( $P(A) = 1 - P(\Omega \setminus A)$ ), the two notions of regularity coincide in the SEU case, where  $U(f) = Pf$  for a fixed  $P$ , with the usual notion of regularity of the measure  $P$ .<sup>9</sup>

For any subset  $I$  of  $\mathbb{N}$ , let  $\Sigma_I = \sigma(\bigvee_{i \in I} \mathcal{S}_i)$  and denote by  $\mathcal{F}_I$  the set of all acts that are  $\Sigma_I$ -measurable. (When  $I = \{i\}$ , we write  $\Sigma_i$  and  $f \in \mathcal{F}_i$ .) Such acts will be said to depend only on experiments in  $I$ . Particularly important, for empirical relevance, are acts that depend on finitely many experiments, that is, acts in

$$\mathcal{F}_{fin} = \bigcup_{I \text{ finite}} \mathcal{F}_I.$$

Refer to such acts as *finitely-based*.

An implication of Regularity is that utility is completely determined by its values on finitely-based acts. Much as the Kolmogorov Extension Theorem tells us that a probability measure on  $\Omega = S^\infty$ , (which is necessarily regular given that  $\Omega$  is metric), is completely determined by its values on finite cylinders, a generalized extension theorem proven in [15, Theorem D.2] implies that a regular utility is uniquely determined by its values on  $\mathcal{F}_{fin}$ .<sup>10</sup>

---

<sup>8</sup>The reader is referred to [15] for detailed discussion of regularity of preferences and the formal relationship to regular probability measures and also regular capacities.

<sup>9</sup>More precisely, it follows from [15, Theorem 4.1] that: an SEU preference with prior  $P$  is regular in the sense of Epstein-Wang if and only if it is regular in the sense of this paper if and only if  $P$  is a regular measure.

<sup>10</sup>The different meaning of “regularity,” explained above, does not affect the validity of the Kolmogorov-style theorem.

Similarly, given Regularity, it can be shown that it is enough to restrict the core axioms characterizing multiple-priors utility, Monotonicity, Ambiguity Aversion and Certainty Independence, to finitely-based acts.

We can now state the main result of this section, which serves as the starting point for our modeling of “exchangeability”.<sup>11</sup>

**Theorem 2.2.** *Let  $U$  be a multiple-priors utility as in (2.4). Then  $U$  satisfies Regularity if and only if it can be expressed in the form (2.6) for some  $\mathcal{P} \subset ca_+^1(\Omega)$  that is convex and (weak-convergence) compact. Moreover, the set  $\mathcal{P}$  is unique.*

Henceforth, when referring to *multiple-priors preference or utility*, regularity is assumed, even if not mentioned explicitly. Similarly, by a *set of priors* we mean a set  $\mathcal{P} \subset ca_+^1(\Omega) = \Delta(\Omega)$  that is convex and compact.

We conclude this section with an elementary lemma that we use repeatedly. Say that  $P \in \mathcal{C}$  is a *minimizing*, or *supporting*, *measure for  $f$*  if  $P$  is a solution in (2.4). If  $f$  is continuous (or even lsc), such as if  $f$  is finitely-based, then there is a minimizer in  $\mathcal{P}$ , but not so in general.

**Lemma 2.3.** *Let  $f_i \in \mathcal{F}_{fin}$  and  $\alpha_i > 0$ ,  $i = 1, \dots, n$ , with  $\sum_{i=1}^n \alpha_i = 1$ . Then*

$$U(\sum_{i=1}^n \alpha_i f_i) = \sum_{i=1}^n \alpha_i U(f_i) \tag{2.7}$$

*if and only if every measure supporting  $\sum_{i=1}^n \alpha_i f_i$  also supports every  $f_i$ . In particular, (2.7) implies that, for any  $m \leq n$ ,*

$$U(\sum_{i=1}^m \beta_i f_i) = \sum_{i=1}^m \beta_i U(f_i),$$

*for any  $\beta_i > 0$ ,  $\sum_{i=1}^m \beta_i = 1$ .*

**Proof.** Let  $P^*$  support the mixed act. Then

$$U(\sum_i \alpha_i f_i) = \sum_i \alpha_i P^* f_i > \sum_i \alpha_i U(f_i),$$

if  $P^*$  is not minimizing for some  $f_i$ . The rest of the proof is obvious. ■

### 3. EVIDENCE OF SYMMETRY

Turn finally to the core question - how to model the distinction described in the title.

The first part is obvious - if evidence is symmetric, then it is intuitive that an individual would satisfy:

---

<sup>11</sup>See Appendix A for a proof; it relies, for one direction, on a result by Chen [8]. We remind the reader that since the probability-equivalent utility  $U$  corresponds uniquely to preference, the theorem could be restated in terms of the latter. Finally, see [28, Proposition 1] for a related result dealing with lower envelopes of sets of priors rather than with preferences over acts.

**Axiom 1 (SYMMETRY).** For all finitely-based acts  $f$  and permutations  $\pi$ ,  $f \sim \pi f$ .

Assuming subjective expected utility, Symmetry is equivalent to exchangeability of the prior, as noted above. But Symmetry in itself is a relatively weak assumption following, for example, from symmetry of information about all the experiments. The force of the assumption of Symmetry, as reflected in de Finetti's theorem, stems largely from the added assumption of expected utility theory, or a single prior, as will be evident in the sequel. In a multiple-priors framework, relatively little structure is implied for the set of priors.

**Theorem 3.1.** Let  $\succeq$  be represented by multiple-priors utility as in (2.6), with set of priors  $\mathcal{P}$ . Then  $\succeq$  satisfies Symmetry iff for every finite permutation  $\pi$ ,

$$P \in \mathcal{P} \implies \pi P \in \mathcal{P}. \quad (3.1)$$

Say that  $\mathcal{P}$  is *symmetric* if it satisfies (3.1).

**Proof.** The axiom implies that preference over  $\mathcal{F}_{fin}$  is represented both by  $\mathcal{P}$  and by  $\{\pi P : \pi \in \Pi, P \in \mathcal{P}\}$ . It is clear that the latter set is convex and compact. Hence they represent the same preference over  $\mathcal{F}$  by [15, Theorem D.2]. Therefore, they must be identical by the uniqueness of the representing set of priors (Theorem 2.2). The other direction is obvious. ■

### 3.1. An Extreme Case

The heart of the paper concerns modeling the perception of “evidence of symmetry.” One way to do so is described next. The axiom is redundant in the Bayesian case because Symmetry and subjective expected utility jointly imply also the following axiom.

**Axiom 2 (STRONG EXCHANGEABILITY).** For all finitely-based acts  $f$  and all  $\alpha$  in  $[0, 1]$ ,

$$\alpha f + (1 - \alpha) \pi f \sim f.$$

The axiom implies Symmetry - set  $\alpha = 0$ . But it is much stronger - recall the example and discussion in the introduction. Based on the intuition given there, we interpret the indifference in the axiom as expressing the individual's view that all experiments are identical, just as in de Finetti's model.

**Lemma 3.2.**  $\mathcal{P}^{exch} \equiv \{P \in \mathcal{P} : \pi P = P \text{ for every } \pi\}$  is convex and compact.

**Proof.** Convexity is obvious. A measure  $P$  is exchangeable if *and only if*, for every  $\pi$  and for every  $f \in \mathcal{F}_{fin}$ ,

$$Pf = P(\pi f).$$

But since  $f \in \mathcal{F}_{fin}$  is continuous, this equality is preserved in the (weak-convergence) limit. ■

Strong Exchangeability is satisfied if and only if every prior is exchangeable.

**Theorem 3.3.** Let  $\succeq$  be represented by regular multiple-priors utility as in (2.6), with set of priors  $\mathcal{P}$ . Then  $\succeq$  satisfies Strong Exchangeability if and only if

$$\mathcal{P} = \mathcal{P}^{exch} \equiv \{P \in \mathcal{P} : \pi P = P \text{ for all } \pi\}. \quad (3.2)$$

**Proof.** If  $\mathcal{P} = \mathcal{P}^{exch}$ , then every  $P$  in  $\mathcal{P}$  satisfies: for every act  $f$  and permutation  $\pi$ ,

$$P(\pi f) = (\pi P)f = Pf.$$

For any finitely-based act  $f$ , if

$$U(f) = \min_{P \in \mathcal{P}} Pf = P^* f$$

for some  $P^*$  in  $\mathcal{P}$ , then

$$U(\pi f) = \min_{P \in \mathcal{P}} P(\pi f) = \min_{P \in \mathcal{P}} (\pi P)f = \min_{P \in \mathcal{P}} Pf = P^* f,$$

that is,  $P^*$  is also minimizing for  $\pi f$ . Therefore, Lemma 2.3 gives the result.

Conversely, assume Strong Exchangeability. We claim that the indifference asserted in the axiom extends, because of Regularity, to all acts. To see this, let  $\widehat{U}(f) = U(\alpha f + (1 - \alpha)\pi f)$ . By Strong Exchangeability,  $\widehat{U} = U$  on  $\mathcal{F}_{fin}$ . Lemma A.1 shows that  $\widehat{U}$  satisfies Regularity. Therefore,  $\widehat{U}(f) = U(f)$  for all  $f \in \mathcal{F}$  by [15, Theorem D.2].

Refer to  $P^*$  in  $\mathcal{P}$  as an *exposed point* if there exists a continuous act  $f$  such that

$$\{P^*\} = \arg \min_{P \in \mathcal{P}} Pf.$$

Then,  $\alpha f + (1 - \alpha)\pi f \sim f \implies$  there is a common minimizing measure for  $f$  and  $\pi f \implies P^* = \pi P^*$ , and this is true for every  $\pi$ . That is,  $P^* \in \mathcal{P}^{exch}$ .

Argue next that  $\mathcal{P}$  equals the closed convex hull of its exposed points: Since  $\Omega$  is separable, so is  $ca(\Omega, \Sigma)$  with the weak convergence topology. Therefore,  $C(\Omega)$ , the Banach space of continuous real-valued functions with the sup norm, is an Asplund space [27, Theorem 2.12]. The assertion now follows from [27, Theorem 5.12].

Finally, (3.2) is implied by the fact that  $\mathcal{P}^{exch}$  is closed and convex. ■

The de Finetti Theorem gives another way to express the implication for  $\mathcal{P}$  of Strong Exchangeability. Since each prior  $P$  in  $\mathcal{P}$  is exchangeable, it has a representation as in (2.1). Denote by  $\mu_P$  the measure on  $\Delta(S)$  corresponding to  $P$  in (2.1) and let  $\mathcal{M} = \{\mu_P : P \in \mathcal{P}\}$ . Then

$$\mathcal{P} = \left\{ \int \ell^\infty(\cdot) d\mu(\ell) : \mu \in \mathcal{M} \right\}. \quad (3.3)$$

Thus we have proven:

**Corollary 3.4.** Let  $\succeq$  be represented by regular multiple-priors utility as in (2.6), with set of priors  $\mathcal{P}$ . Then  $\succeq$  satisfies Strong Exchangeability if and only if it can be expressed in the form (3.3) for some set  $\mathcal{M} \subset \Delta(\Delta(S))$ .

The Corollary clarifies how a model with Strong Exchangeability differs from the de Finetti model. The representation (3.3) suggests the interpretation whereby the individual is uncertain *ex ante* which likelihood function applies, but she is certain that the same likelihood function applies to all experiments. This is just as for the Bayesian case - experiments are perceived as identical. The difference here is that the *ex ante* uncertainty is in general not representable by a single probability measure - there is ambiguity rather than risk regarding the true likelihood function. For the introductory example of coin tossing, there is ambiguity about the given coin's bias, but the way in which the coin is tossed is viewed as fixed across tosses, and hence these are viewed as identical.

**Remark 1.** *Walley [33, Ch. 9] defines and discusses exchangeability for “previsions”  $\nu$ , where  $\nu(f)$  is interpreted as the maximum price (in utils) the individual would be willing to pay for the act  $f$ , that is, so that the act  $f - \nu(f)$  is just desirable. Symmetry of evidence is expressed through indifference between an act  $f$  and any permutation  $\pi f$ , in the sense that  $\nu(f) = \nu(\pi f)$ . Walley suggests an additional axiom, which he calls exchangeability, which states that*

$$\nu(\pi f - f) = 0, \text{ for all } f \text{ and } \pi.$$

*The axiom, and his representation result, bear some similarity to Strong Exchangeability and Theorem 3.3. His formulation leads to results that follow almost by definition - for example, the heavy machinery invoked in the proof of our theorem is not needed. In addition, outcomes in his model are utils and, unlike the case here, his analysis is not readily translated into decision-theoretic terms using only “actual” outcomes.*

### 3.2. A New Axiom

We have argued that Strong Exchangeability is “too strong” in that it implies overwhelming evidence of symmetry - all experiments are viewed as identical. We propose an alternative axiom here (called Orthogonal Independence) that weakens the classic Independence axiom so as to permit modeling the sought-after perception - there are poorly understood factors affecting all experiments in a suitably unrelated fashion.

Nevertheless, the axiom is not comparable with Strong Exchangeability. The reason is that it rules out the sort of prior ambiguity represented by the nonsingleton set  $\mathcal{M}$  in (3.3). In terms of the implied representation to be described below (Theorem 4.2), it will have in common with de Finetti's (2.1) a single prior, but it will differ from his in featuring (in a suitable sense) multiple likelihoods. We have not yet achieved a general model that admits also prior ambiguity, though Section 5 describes an intuitive axiom and a conjectured representation.

To motivate Orthogonal Independence, recall again the coin-tossing example, and suppose that  $H_2 \succeq T_2$ . The Independence axiom requires that mixing these bets with any common third bet should leave the ranking invariant. We have already seen that this is not intuitive, in general, given a concern that the experiments may not be identical. However, consider the special case where the third bet is  $H_3$ . We claim that *if*, as we do, *one excludes ambiguity about the coin's bias*, then there is no reason for an individual to

reverse his ranking in this case, that is, we would expect the ranking<sup>12</sup>

$$\frac{1}{2}H_2 + \frac{1}{2}H_3 \succeq \frac{1}{2}T_2 + \frac{1}{2}H_3. \quad (3.4)$$

This is because the bet  $H_3$  on the third toss cannot hedge the uncertainty regarding a bet on the second toss if the individual perceives the poorly understood factors (tossing style) to vary “independently” across tosses. Moreover, there is similar intuition for the implication that: if  $T_1H_2 \succeq T_1T_2$ , then also

$$\frac{1}{2}T_1H_2 + \frac{1}{2}T_1H_3 \succeq \frac{1}{2}T_1T_2 + \frac{1}{2}T_1H_3. \quad (3.5)$$

In this case, the three bets  $T_1H_2, T_1T_2$  and  $T_1H_3$  depend on overlapping tosses, though in a specific way. Our central axiom describes the general restriction suggested by the preceding intuition.

Given any two acts  $f^*$  and  $f$ , then  $f^* \cdot f$  denotes the pointwise product, that is, the act given by

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

Recall from Section 2.2 that the outcome produced by  $f$  in state  $\omega$  can be viewed as a coin toss which gives the best ‘true’ underlying outcome  $\bar{z}$ , or utility 1, with objective probability  $f(\omega)$ , and the worst outcome  $\underline{z}$ , or utility 0, with the complementary probability. Similarly, in state  $\omega$  the product act  $f^* \cdot f$  gives a lottery where 1 util is received with objective probability  $f^*(\omega) f(\omega)$ , corresponding to the independent tosses of the two coins associated with  $f^*$  and  $f$ .

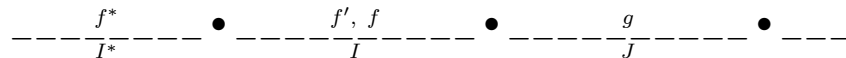
Our central axiom can now be stated.

**Axiom 3 (ORTHOGONAL INDEPENDENCE (OI)).** For all  $0 < \alpha \leq 1$ , acts  $f^* \in \mathcal{F}_{I^*}, f', f \in \mathcal{F}_I$  and  $g \in \mathcal{F}_J$ , with  $I^*, I$  and  $J$  finite and disjoint,

$$f^* \cdot f' \succeq f^* \cdot f \text{ iff } \alpha f^* \cdot f' + (1 - \alpha) f^* \cdot g \succeq \alpha f^* \cdot f + (1 - \alpha) f^* \cdot g.$$

The reason for the name is that one might refer to acts  $f$  and  $g$  as in the statement as being *orthogonal* because they depend on different experiments. Formally, say that  $f$  and  $g$  are (mutually) *orthogonal*, written  $f \perp g$ , if  $f \in \mathcal{F}_I$  and  $g \in \mathcal{F}_J$  for some disjoint  $I$  and  $J$ . Note that all the acts in the axiom statement are finitely-based.

The diagram below illustrates the orthogonality assumed in the axiom. The positioning of acts above the line indicates that  $g$  depends only on experiments in  $J$ ,  $f$  only on those in  $I$ , and so on.



<sup>12</sup>On the other hand, the intuition *against* the ranking in (3.4) is clear if there is ambiguity about the physical bias, even in the simplest scenario where the individual believes that *only* the bias influences outcomes. Then  $H_3$  hedges  $T_2$  but not  $H_2$ , and thus the individual may strictly prefer  $\frac{1}{2}T_2 + \frac{1}{2}H_3$ .

The special case  $f^*$  constant (or  $I^*$  empty) is easiest to understand. Then, because  $g$  depends on different experiments than do  $f'$  and  $f$ , it does not hedge either and mixing with  $g$  does not reverse the ranking of  $f'$  versus  $f$  - this is illustrated in the coin-tossing example by (3.4). The more general case is illustrated by the expanded coin-tossing example surrounding (3.5).

Note finally that Orthogonal Independence is weaker than the Independence Axiom, and stronger than Certainty-Independence, one of the axioms characterizing multiple-priors utility (see Section 2.2), which is obtained by restricting  $I^*$  to be empty and  $g$  to be a constant act.

Orthogonal Independence implies a form of additivity for utility that is used repeatedly (and without reference).

**Lemma 3.5.** *If  $f^* \in \mathcal{F}_{I^*}$ ,  $f \in \mathcal{F}_I$  and  $g \in \mathcal{F}_J$ , with  $I^*$ ,  $I$  and  $J$  finite and disjoint, then*

$$U(\alpha f^* \cdot f + (1 - \alpha)f^* \cdot g) = \alpha U(f^* \cdot f) + (1 - \alpha)U(f^* \cdot g).$$

**Proof.** Suppose to the contrary that

$$U(\alpha f^* \cdot f + (1 - \alpha)f^* \cdot g) > \alpha U(f^* \cdot f) + (1 - \alpha)U(f^* \cdot g).$$

By continuity, it is enough to treat the case  $U(f^* \cdot f) < 1$ . Then there exists a constant  $p$  such that

$$U(f^* \cdot p) > U(f^* \cdot f) \text{ and } U(\alpha f^* \cdot f + (1 - \alpha)f^* \cdot g) > \alpha U(f^* \cdot p) + (1 - \alpha)U(f^* \cdot g),$$

that is,

$$f^* \cdot p \succ f^* \cdot f \text{ and } \alpha f^* \cdot f + (1 - \alpha)f^* \cdot g \succ \alpha f^* \cdot p + (1 - \alpha)f^* \cdot g,$$

which contradicts Orthogonal Independence. ■

We provide two examples to illustrate what is excluded by Orthogonal Independence.

**Example 3.6.** *Let  $P^0$  be any countably additive (not necessarily exchangeable) measure and define  $\mathcal{P}$  to be the closed convex hull of  $\{\pi P^0 : \pi \in \Pi\}$ . By construction,  $\mathcal{P}$  is symmetric. However, it violates OI.*

*There is a simple interpretation:  $P^0$  reflects some asymmetries across experiments, for example, it might be believed that coin 1 is biased towards Heads and that the others are unbiased. If beliefs are instead that there exists exactly one biased coin, though its identity is completely unknown, one is led to  $\{\pi P^0 : \pi \in \Pi\}$ .<sup>13</sup> Then the agent would be indifferent between betting on Tails for any two coins, but, contrary to OI, she would strictly prefer to randomize, that is,*

$$\frac{1}{2}T_1 + \frac{1}{2}T_2 \succ T_1 \sim T_2.$$

---

<sup>13</sup>Taking the closed convex hull has no consequence for decisions.

Since the worst case scenario for  $T_1$  ( $T_2$ ) is that the first (second) coin is the biased one, the mixture smooths out these uncertainties and guarantees at least one coin that is not biased against Tails. Hence it is strictly preferable. OI is violated because the poorly understood factor - which coin is the biased one - relates the outcomes of the different coins since there is certainty that only one is biased.

**Example 3.7.** Fix a probability measure  $\ell^*$  in  $\Delta(S)$  and let

$$\mathcal{P} = \{P \in \Delta(\Omega) : \text{mrg}_{S_i} P = \ell^* \text{ for all } i \}.$$

Thus  $\mathcal{P}$  consists of all measures that agree with  $\ell^*$  on each  $S_i$ , with joint distributions across different experiments being unrestricted. The interpretation is that there is no ambiguity about the nature of any single experiment, but there is complete ignorance about how experiments are correlated. This perception of the experiments is not covered by our model. The individual in our model is uncertain that experiments are identical because she views each experiment as being affected also by poorly understood factors that vary across experiments, but she is certain that these are unrelated across experiments. Here, in contrast, she is concerned with the possible correlation of these factors across experiments.<sup>14</sup>

Though  $\mathcal{P}$  is obviously symmetric, compact and convex, it lies outside the scope of our model because it violates OI as we now show.

For concreteness, let  $S = \{H, T\}$  and let  $\ell^*$  describe an unbiased coin. OI would imply that

$$\begin{aligned} & U\left(\left(\frac{1}{2}H_1 + \frac{1}{2}\right)\left(\frac{1}{2}T_2 + \frac{1}{2}\right)\left(\frac{1}{2}H_3 + \frac{1}{2}\right)\right) \\ &= \frac{1}{2}U\left(\left(\frac{1}{2}H_1 + \frac{1}{2}\right)\left(\frac{1}{2}T_2 + \frac{1}{2}\right)H_3\right) + \frac{1}{2}U\left(\left(\frac{1}{2}H_1 + \frac{1}{2}\right)\left(\frac{1}{2}T_2 + \frac{1}{2}\right)\right) \\ & \dots \\ &= \frac{1}{8} \left[ \begin{array}{c} U(1) + U(H_1) + U(T_2) + U(H_3) \\ +U(H_1 \cdot T_2) + U(T_2 \cdot H_3) + U(H_1 \cdot H_3) + U(H_1 \cdot T_2 \cdot H_3) \end{array} \right]. \end{aligned}$$

Thus, there is a common minimizing measure, say  $P$ , for the acts  $H_1, T_2, H_3, H_1 \cdot T_2, T_2 \cdot H_3, H_1 \cdot H_3$  and  $H_1 \cdot T_2 \cdot H_3$ . Compute that

$$\begin{aligned} U(H_1) &= U(T_2) = U(H_3) = \frac{1}{2}, \text{ and} \\ U(H_1 \cdot T_2) &= U(T_2 \cdot H_3) = U(H_1 \cdot H_3) = U(H_1 \cdot T_2 \cdot H_3) = 0, \end{aligned}$$

where, for example,  $U(H_1 \cdot T_2) = 0$  because the worst-case scenario for this act is that coins 1 and 2 are perfectly positively correlated. Since  $P$  is a common minimizer, deduce that

$$\begin{aligned} P(H_1) &= P(T_2) = P(H_3) = \frac{1}{2}, \text{ and} \\ P(H_1 T_2) &= P(T_2 H_3) = P(H_1 H_3) = P(H_1 T_2 H_3) = 0. \end{aligned}$$

---

<sup>14</sup>In fact, the difference is more subtle, since, as shown in the sequel, OI does permit the perception of some degree of dependence between experiments.

But there does not exist a probability measure satisfying these conditions. (Since  $P(H_1T_2) = 0$ ,  $P(H_1H_3) = 0$  and  $P(H_1) = \frac{1}{2}$ , it follows that

$$\begin{aligned} P(H_1H_2H_3) &= 0, & P(H_1H_2T_3) &= \frac{1}{2}, \\ P(H_1T_2H_3) &= 0 & \text{and } P(H_1T_2T_3) &= 0. \end{aligned}$$

Combine these with  $P(T_2) = \frac{1}{2}$  and  $P(T_2H_3) = 0$  to deduce that

$$P(T_1T_2H_3) = 0 \text{ and } P(T_1T_2T_3) = \frac{1}{2}.$$

Finally, use  $P(H_3) = \frac{1}{2}$  to conclude that  $P(T_1H_2H_3) = \frac{1}{2}$ . But then  $P(H_1H_2T_3) + P(T_1T_2T_3) + P(T_1H_2H_3) > 1$ .

Finally, note that a weaker form of *OI*, where  $f^* \equiv 1$  or  $I^*$  is empty, is satisfied. Let  $f \in \mathcal{F}_I$  and  $g \in \mathcal{F}_J$ , with  $I$  and  $J$  finite and disjoint. It suffices to show that

$$U(\alpha f + (1 - \alpha)g) = \alpha U(f) + (1 - \alpha)U(g).$$

Let  $P$  and  $P'$  be minimizing measures for  $f$  and  $g$  respectively. Then  $\text{mrg}_{\mathbb{N} \setminus J} P \otimes \text{mrg}_J P'$ , the product measure formed from the  $\mathbb{N} \setminus J$ -marginal of  $P$  and the  $J$ -marginal of  $P'$ , lies in  $\mathcal{P}$  and is minimizing for both  $f$  and  $g$ . The noted additivity follows from the existence of a common minimizer.

### 3.3. A Final Axiom: Dominance

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

$$(\theta f)(s_1, s_2, s_3, \dots) = f(s_2, s_3, \dots);$$

$\theta^n$  denotes the  $n$ -fold replication of  $\theta$ . It is straightforward to show that Symmetry implies also indifference to shifts.

**Lemma 3.8.**  $\theta f \sim f$  for all  $f \in \mathcal{F}$ .

**Proof.** Let  $\widehat{U}(f) = U(\theta f)$ . By Symmetry,  $\widehat{U} = U$  on  $\mathcal{F}_{fin}$ . By [15, Theorem D.2] and Lemma B.11, the two functions coincide everywhere. ■

Product acts continue to play a role below. For any act  $f^* \in \mathcal{F}_{\{1, \dots, n\}}$ ,  $f^* \cdot \theta^n f$  is the act satisfying

$$(f^* \cdot \theta^n f)(\omega) = f^*(s_1, \dots, s_n) f(s_{n+1}, s_{n+2}, \dots).$$

Note that the acts  $f^*$  and  $\theta^n f$  are orthogonal (depend on different experiments).

Say that  $f$  and  $g$  *do not hedge* one another if they (are finitely-based and) satisfy:<sup>15</sup> for all  $f^* \in \mathcal{F}_{\{1, \dots, n\}}$  and  $p \in [0, 1]$ ,

$$f^* \cdot \theta^n f \sim p \implies [\frac{1}{2}f^* \cdot \theta^n f + \frac{1}{2}f^* \cdot \theta^n g \sim \frac{1}{2}p + \frac{1}{2}f^* \cdot \theta^n g]. \quad (3.6)$$

<sup>15</sup>By Lemma 2.3, one would obtain an equivalent definition if one required indifference for all  $\alpha$ -mixtures, and not just  $\alpha = \frac{1}{2}$ .

If  $f^* \cdot \theta^n f \sim p$ , then nonindifference in the second comparison would arise only if the hedging properties of the two mixtures differ. However, there is no hedging in the mixture involving the constant act  $p$ . Therefore, the asserted indifference in (3.6) indicates that, in an informal sense,  $f^* \cdot \theta^n f$  and  $f^* \cdot \theta^n g$  do not hedge one another. Take  $f^* = 1$  to conclude that the same can be said for  $f$  and  $g$ . Conversely, (3.6) can be “derived”, because, as in the discussion of Orthogonal Independence, the common orthogonal factor  $f^*$  does not affect the incentive to randomize.

This notion of (non)hedging is subjective - it is tied to the given preference. In terms of utility, for any multiple-priors utility, (Certainty Independence is enough), condition (3.6) can be written in the form: for all  $f^* \in \mathcal{F}_{\{1, \dots, n\}}$ ,

$$U\left(\frac{1}{2}f^* \cdot \theta^n f + \frac{1}{2}f^* \cdot \theta^n g\right) = \frac{1}{2}U(f^* \cdot \theta^n f) + \frac{1}{2}U(f^* \cdot \theta^n g).$$

Given Orthogonal Independence, orthogonal acts do not hedge one another; but nonorthogonal acts can also be nonhedging. This is clear, since it is immediate from the definition that  $f^* \cdot \theta^n f$  and  $f^* \cdot \theta^n g$  are nonhedging whenever  $f$  and  $g$  are. In fact, we can say more.

**Lemma 3.9.** *If  $f$  and  $g$  do not hedge one another, then, assuming the preceding axioms, neither do  $f^{**} \cdot \theta^m f$  and  $f^* \cdot \theta^m g$ , for all  $f^{**}, f^* \in \mathcal{F}_{\{1, \dots, m\}}$ , with  $f^{**} \perp f^*$ .*

**Proof.** Compute that

$$\begin{aligned} & U\left(\frac{1}{4}f^{**} \cdot \theta^m f + \frac{1}{4}f^* \cdot \theta^m f + \frac{1}{4}f^{**} \cdot \theta^m g + \frac{1}{4}f^* \cdot \theta^m g\right) \\ &= U\left(\left(\frac{1}{2}f^{**} + \frac{1}{2}f^*\right) \cdot \left(\frac{1}{2}\theta^m f + \frac{1}{2}\theta^m g\right)\right) \\ &= \frac{1}{2}U\left(\left(\frac{1}{2}f^{**} + \frac{1}{2}f^*\right) \cdot \theta^m f\right) + \frac{1}{2}U\left(\left(\frac{1}{2}f^{**} + \frac{1}{2}f^*\right) \cdot \theta^m g\right) \quad (\text{since } f \text{ and } g \text{ do not hedge}) \\ &= \frac{1}{4}U(f^{**} \cdot \theta^m f) + \frac{1}{4}U(f^* \cdot \theta^m f) + \frac{1}{4}U(f^{**} \cdot \theta^m g) + \frac{1}{4}U(f^* \cdot \theta^m g) \quad (\text{by OI}). \end{aligned}$$

It follows from Lemma 2.3, (existence of a common minimizer), that

$$U\left(\frac{1}{2}f^{**} \cdot \theta^m f + \frac{1}{2}f^* \cdot \theta^m g\right) = \frac{1}{2}U(f^{**} \cdot \theta^m f) + \frac{1}{2}U(f^* \cdot \theta^m g).$$

In the same way, one can prove the additivity

$$\begin{aligned} & U\left(\frac{1}{2}f^{***} \cdot \theta^n(f^{**} \cdot \theta^m f) + \frac{1}{2}f^{***} \cdot \theta^n(f^* \cdot \theta^m g)\right) \\ &= \frac{1}{2}U(f^{***} \cdot \theta^n(f^{**} \cdot \theta^m f)) + \frac{1}{2}U(f^{***} \cdot \theta^n(f^* \cdot \theta^m g)), \end{aligned}$$

for any  $f^{***} \in \mathcal{F}_{\{1, \dots, n\}}$ . ■

Our final axiom is:

**Axiom 4 (DOMINANCE).** *If  $g^* \geq h^* \in \mathcal{F}_{\{1, \dots, m\}}$  are nonhedging, then*

$$\frac{1}{2}g^* \cdot \theta^{n+m} f' + \frac{1}{2}h^* \cdot \theta^m f \succeq \frac{1}{2}g^* \cdot \theta^m f + \frac{1}{2}h^* \cdot \theta^{n+m} f', \quad (3.7)$$

if: either (i)  $f' \geq f$ , where  $f', f \in \mathcal{F}_{\{1, \dots, n\}}$ ; or  
(ii)  $f' = \theta^n(\alpha F' + (1 - \alpha) F)$ ,  $f = (\alpha \theta^n F' + (1 - \alpha) F)$ , where  $F', F \in \mathcal{F}_{\{1, \dots, n\}}$ .

Two implications of the axiom provide some perspective. By taking  $g^* = 1$  and  $h^* = 0$ , the axiom implies (given Symmetry and OI) Monotonicity and Ambiguity Aversion. For the former, if  $f' \succeq f$ , then (3.7) implies

$$f' \sim \theta^{n+m} f' \succeq \theta^m f \sim f.$$

For the latter, let  $F \succeq F'$ , and define  $f'$  and  $f$  as in (ii). Then (3.7) implies

$$\begin{aligned} \alpha F' + (1 - \alpha) F &\sim \theta^{2n+m} (\alpha F' + (1 - \alpha) F) \succeq \theta^m (\alpha \theta^n F' + (1 - \alpha) F) \\ &\sim (\alpha \theta^n F' + (1 - \alpha) F) \succeq \theta^n F' \sim F', \end{aligned}$$

where the second weak preference is due to OI (since  $\theta^n F'$  and  $F$  are orthogonal). Thus, given Symmetry and OI, one can view Dominance as strengthening Monotonicity and Ambiguity Aversion, properties that are meaningful in an abstract state space, for our setting of many symmetric and suitably independent experiments.

For further interpretation, notice first that each of the acts being compared in (3.7) is a mixture of nonhedging acts: since  $g^*$  and  $h^*$  are nonhedging, and since  $\theta^{n+m} f'$  and  $\theta^m f$  are orthogonal to one another and to both  $g^*$  and  $h^*$ , Lemma 3.9 implies that  $g^* \cdot \theta^{n+m} f'$  and  $h^* \cdot \theta^m f$  are nonhedging; similarly for  $g^* \cdot \theta^n f'$  and  $h^* \cdot \theta^{n+m} f'$ . Thus hedging considerations do not play a role in the expressed ranking.

Next, each of the conditions (i) and (ii) is interpretable as a form of dominance, paralleling first and second-order stochastic dominance for lotteries. Consider (i), where  $f' \succeq f$ . Though  $\theta^n f' \not\succeq f$ , *any* preference satisfying Symmetry and Monotonicity would rank  $\theta^n f'$  above  $f$ , ( $\theta^n f' \sim f' \succeq f$ ). In that sense,  $\theta^n f'$  first-order dominates  $f$ . The interpretation for case (ii) is only slightly more complicated. Compare  $\theta^n (\alpha F' + (1 - \alpha) F)$  with  $(\alpha \theta^n F' + (1 - \alpha) F)$ . Under Symmetry, this is equivalent to the comparison between  $(\alpha F' + (1 - \alpha) F)$  and  $(\alpha \theta^n F' + (1 - \alpha) F)$ . The essential difference between them is that the former mixture permits hedging the ambiguity associated with  $F'$  and  $F$ , while the latter does not - there are no hedging gains from randomizing between  $\theta^n F'$  and  $F$  because they are orthogonal. Thus *any* multiple-priors preference satisfying Symmetry and OI would rank  $f'$  above  $f$ . In this sense,  $\theta^n f'$  second-order dominates  $f$ , where “second-order” refers to aversion to ambiguity rather than aversion to risk.

Finally, we are in a position to interpret the axiom. Think of the randomization in (3.7) as coming from the toss of an objectively unbiased coin. Denote the two mixed acts by  $LH$  (left side) and  $RH$  (right side). Then  $LH$  yields a “dominating act” ( $g^* \cdot \theta^{n+m} f'$  versus  $g^* \cdot \theta^m f$ ) if Heads occurs, while  $RH$  yields a “dominating act” ( $h^* \cdot \theta^{n+m} f'$  versus  $h^* \cdot \theta^m f$ ) if Tails occurs. Because there is no hedging (or interaction) involved in either mixture, to rank them we can think in terms of the above gains and losses and how they compare. The differences between  $LH$  and  $RH$  are “almost” mirror images, (they arise by replacing  $\theta^m f$  by  $\theta^{n+m} f'$ ), but they are compounded differently - the gain in  $RH$  is compounded by  $h^*$ , while the gain in  $LH$  is compounded by the larger act  $g^*$ . Thus the latter difference dominates and the ranking  $LH \succeq RH$  results.

Though we believed, at an earlier stage in this research, that Dominance was implied by the other axioms, that remains an open question. On the other hand, Dominance does not imply Orthogonal Independence, even given the other axioms, as illustrated by the utility function (5.1) discussed in the concluding section.

## 4. REPRESENTATIONS: “CONDITIONALLY IID”

### 4.1. A Definition

Our next objective is to describe the representation implied by Symmetry, Orthogonal Independence and Dominance. It is our counterpart, or generalization, of the “conditionally i.i.d.” representation in de Finetti’s theorem. Thus we begin with a definition of “stochastic independence” of experiments for our framework. (Here we mean that experiments are unrelated, not even by a common bias in the case of coin tossing - think of the case where the bias is known with certainty.) In the Bayesian setting, it amounts to beliefs being represented by a product measure. However, the situation is more complicated in a multiple-priors framework - there are different ways of defining a “product” set of priors consistent with given sets of marginals (for example, see Hendon et al [20] and Ghirardato [17]).

We define “product” in terms of utility functions rather than directly in terms of sets of priors. Say that the multiple-priors utility function  $U$ , as in (2.6), is a *product utility function* if

$$U(f \cdot g) = U(f)U(g) \text{ for all orthogonal } f, g \in \mathcal{F}_{fin}. \quad (4.1)$$

If also preference represented by  $U$  satisfies Symmetry, then refer to an *IID utility function*, and to the corresponding set of priors  $\mathcal{P}$  as an *IID set of priors*. Following [13], the acronym IID stands for “independently and *indistinguishably* (as opposed to identically) distributed.”<sup>16</sup>

The rationale for (4.1) may seem obvious, but its behavioral meaning should be made clear. Recall the probability-equivalence nature of outcomes and utility (see (2.3) and (2.5)). The acts  $f$  and  $g$  are assumed to depend on different experiments - for concreteness, let  $f \in \mathcal{F}_1$  and  $g \in \mathcal{F}_2$ . Then, in state  $(s_1, s_2)$ ,  $f \cdot g$  yields (the equivalent of) successive and independent tosses of an objective  $f(s_1)$ -coin and an objective  $g(s_2)$ -coin.<sup>17</sup> If experiments 1 and 2 are “stochastically independent,” it is intuitive to perceive this prospect as though the “order” of coin tossing were: toss all  $f(s_1)$ -coins as  $s_1$  varies over  $S_1$ , and separately and *independently* toss all  $g(s_2)$ -coins as  $s_2$  varies over  $S_2$ . But the prospect consisting of the first set of coin tosses is equivalent to  $f$ , and the second to  $g$ . Further,  $f$  is indifferent to a  $U(f)$ -coin and  $g$  is indifferent to a  $U(g)$ -coin. Conclude that  $f \cdot g$  is indifferent to winning 1 util if both the  $U(f)$ - and the  $U(g)$ -coins, tossed independently, produce favorable outcomes, which is equivalent to a  $U(f)U(g)$ -coin. This “proves” that (4.1) is implied if experiments are seen to be independent. The converse is just as intuitive.

There is another way, that was invoked above when motivating Orthogonal Independence, in which the perceived independence of experiments is reflected intuitively - through the additivity

$$U(\alpha f + (1 - \alpha)g) = \alpha U(f) + (1 - \alpha)U(g), \quad (4.2)$$

<sup>16</sup>We continue to use the lower case acronym iid when referring to single measures, with the usual meaning of “independently and identically distributed.”

<sup>17</sup>A  $p$ -coin is one that yields 1 util (or the best outcome  $\bar{z}$ ) with objective probability  $p$  and 0 utils (or the worst outcome  $\underline{z}$ ) with probability  $1 - p$ .

for all  $\alpha$  and all finitely-based orthogonal acts  $f$  and  $g$ . The latter is implied by our definition of IID utility, as is easily shown. In fact, we can prove much more.

**Lemma 4.1.** *If  $U$  is an IID utility function, then  $U$  satisfies both Orthogonal Independence and Dominance.*

**Proof.** If  $f$  and  $g$  are orthogonal, then

$$\begin{aligned} U\left(\frac{1}{4}f \cdot g + \frac{1}{4}f + \frac{1}{4}g + \frac{1}{4}\right) &= U\left(\left(\frac{1}{2}f + \frac{1}{2}\right) \cdot \left(\frac{1}{2}g + \frac{1}{2}\right)\right) \\ &= U\left(\frac{1}{2}f + \frac{1}{2}\right) U\left(\frac{1}{2}g + \frac{1}{2}\right) \\ &= \frac{1}{4}[U(f \cdot g) + U(f) + U(g) + 1], \end{aligned}$$

which implies (4.2) by Lemma 2.3. Further, if  $f^*$  is orthogonal to both  $f$  and  $g$ , then

$$\begin{aligned} U(f^* \cdot (\alpha f + (1 - \alpha)g)) &= U(f^*)U((\alpha f + (1 - \alpha)g)) \\ &= \alpha U(f^*)U(f) + (1 - \alpha)U(f^*)U(g) \\ &= \alpha U(f^* \cdot f) + (1 - \alpha)U(f^* \cdot g). \end{aligned}$$

For Dominance, OI and Lemma 3.9 imply that

$$\begin{aligned} U\left(\frac{1}{2}g^* \cdot \theta^{n+m}f' + \frac{1}{2}h^* \cdot \theta^m f\right) &= \frac{1}{2}U(g^* \cdot \theta^{n+m}f') + \frac{1}{2}U(h^* \cdot \theta^m f), \\ U\left(\frac{1}{2}h^* \cdot \theta^{n+m}f' + \frac{1}{2}g^* \cdot \theta^m f\right) &= \frac{1}{2}U(h^* \cdot \theta^{n+m}f') + \frac{1}{2}U(g^* \cdot \theta^m f). \end{aligned}$$

Therefore,

$$\begin{aligned} U\left(\frac{1}{2}g^* \cdot \theta^{n+m}f' + \frac{1}{2}h^* \cdot \theta^m f\right) &\geq U\left(\frac{1}{2}h^* \cdot \theta^{n+m}f' + \frac{1}{2}g^* \cdot \theta^m f\right) \iff \\ U(g^* \cdot \theta^{n+m}f') - U(g^* \cdot \theta^m f) &\geq U(h^* \cdot \theta^{n+m}f') - U(h^* \cdot \theta^m f) \iff \\ U(g^*)(U(\theta^{n+m}f') - U(\theta^m f)) &\geq U(g^*)(U(\theta^{n+m}f') - U(\theta^m f)) \iff \\ (U(g^*) - U(h^*))(U(f') - U(f)) &\geq 0. \end{aligned}$$

Dominance follows. ■

To help fix ideas, we describe one example of an IID utility. Fix a (closed) set  $\mathcal{L}$  of probability measures on  $S$ , thought of as the set of priors applying to any single experiment. Let<sup>18</sup>

$$\mathcal{P}_{WF} = clh(\mathcal{L}^\infty), \text{ where } \mathcal{L}^\infty \equiv \{\otimes_{i \in \mathbb{N}} \ell_i : \ell_i \in \mathcal{L} \text{ for every } i\}. \quad (4.3)$$

Since the utility of any finitely-based act is a minimum over  $\mathcal{L}^\infty$ , which consists exclusively of product measures, (4.1) is obvious; so is symmetry. Therefore,  $U_{WF}$  defined by

$$U_{WF}(f) = \inf_{P \in \mathcal{L}^\infty} Pf, \quad f \in \mathcal{F}, \quad (4.4)$$

---

<sup>18</sup> $\otimes_{t \in I} \ell_t$  denotes the unique countably additive product measure with marginals  $\ell_t$ . Since  $\mathcal{L}^\infty$  is not convex, we take its closed convex hull, denoted by  $clh(\mathcal{L}^\infty)$ , in order to conform to the normalization that sets of priors be closed and convex. The sets  $clh(\mathcal{L}^\infty)$  and  $\mathcal{L}^\infty$  generate the identical preference.

is an IID utility function. This product is adapted from Walley and Fine [34], and has been studied also by Gilboa and Schmeidler [18].

We emphasize that  $U_{WF}$  is just one example of an IID utility function. It is well-known in the decision theory literature (see Hendon *et al* [20] and Ghirardato [17]) that stochastic independence is multi-faceted in the multiple-priors (or nonadditive probability) framework, and hence that there is more than one way to form an independent product from a given set  $\mathcal{L}$  of priors over  $S$ . In other words, in general, and in contrast to the Bayesian setting, there are many utility functions satisfying (4.1), and hence the “stochastic independence” embodied in it, that also agree on the ranking of acts over any single experiment.

## 4.2. The Main Result

Some preliminaries are needed in order to state the representation. Any set of priors  $\mathcal{P}$  lies in  $\mathcal{K}^c(\Delta(\Omega))$ , the space of compact and convex subsets of  $\Delta(\Omega)$ ; the Hausdorff metric topology renders it compact metric.

Each  $\mathcal{P} \in \mathcal{K}^c(\Delta(\Omega))$  corresponds to a unique multiple-priors preference, or equivalently, to a unique multiple-priors utility function  $U_{\mathcal{P}} : \mathcal{F} \rightarrow \mathbb{R}$ , given by

$$U_{\mathcal{P}}(f) = \inf_{P \in \mathcal{P}} Pf.$$

This correspondence induces a compact metric topology on

$$\mathcal{U} = \{U_{\mathcal{P}} : \mathcal{P} \in \mathcal{K}^c(\Delta(\Omega))\}.$$

The subset of IID utility functions,

$$\mathcal{V} = \{U \in \mathcal{U} : U \text{ is IID}\},$$

inherits the induced topology.

**Theorem 4.2.** *The preference  $\succeq$  on  $\mathcal{F}$  is a multiple-priors preference and satisfies Symmetry, Orthogonal Independence and Dominance if and only if it admits representation by a utility function  $U$  of the form in (2.6) satisfying*

$$U(f) = \int_{\mathcal{V}} V(f) d\mu(V), \text{ for all } f \text{ in } \mathcal{F}, \tag{4.5}$$

for some Borel probability measure  $\mu$  on  $\mathcal{V}$ . Moreover,  $\mu$  is unique.

The proof of sufficiency is relegated to Appendix B. Here consider briefly necessity (see the proof in the appendix for further details). The first step is to verify that the integrand on the right is well-defined for every  $f$ . This is done by showing that the function  $V \mapsto V(f)$  is universally measurable, and by making use of the fact that any measure  $\mu$  admits a unique extension, also denoted  $\mu$ , to the universal completion of  $\Sigma$ . (A similar procedure is used throughout, without explicit mention, to make sense of integrals where measurability issues arise.)

Turn to axioms. Since  $U$  is a mixture of symmetric utility functions, it is also symmetric. We showed above (Lemma 4.1) that Orthogonal Independence and Dominance are satisfied by any IID utility function - the argument is readily extended to any mixture of IID utility functions as in the representation.

The theorem generalizes de Finetti's, wherein each IID utility function in the support of  $\mu$  is an expected utility function with i.i.d. probabilistic beliefs. The more general representation (4.5) suggests an interpretation similar to that familiar for a mixture of i.i.d. beliefs. Any IID utility function reflects the view that experiments are indistinguishable (because of Symmetry) and unrelated or independent. Thus experiments would be IID (indistinguishable and independent) if the individual knew which IID utility function were appropriate or correct. However, she is uncertain of that, as reflected by the measure  $\mu$ . Overall, therefore, she views experiments as being IID conditionally on the correct  $V$ . Because the possible functions  $V$  correspond to multiple-priors rather than to expected utility, the individual may value randomization, as illustrated in the introduction, and accordingly not view experiments as being identical.

To illustrate, suppose that each IID function in the support of  $\mu$  has the form in (4.4). Then, in a slight abuse of notation, where the uncertainty modeled by  $\mu$  is translated into uncertainty about the true set  $\mathcal{L}$ ,

$$U(f) = \int \left( \inf_{\mathcal{L}^\infty} P f \right) d\mu(\mathcal{L}). \quad (4.6)$$

Given resolution of that uncertainty and thus a specific  $\mathcal{L}$ , the same set is assumed to describe each experiment (because of the minimum is over  $\mathcal{L}^\infty$ ). This implies that experiments are indistinguishable (or viewed symmetrically). However, experiments are not viewed as identical because  $\mathcal{L}^\infty$  admits that different likelihoods from  $\mathcal{L}$  apply to different experiments.

In the concrete setting of coin-tossing, any (convex) set  $\mathcal{L}$  of likelihoods can be identified with an interval  $\Theta = [\theta_m, \theta_M] \subset [0, 1]$ , interpreted as a set of possible probabilities for Heads. There is ex ante uncertainty about which interval is the correct one, but conditional on knowing  $\Theta$ , coin tosses are viewed as indistinguishable ambiguous experiments, independent from one another in the specific sense of (4.3). It is easily seen how the model can accommodate the value of randomization in our motivating example (1.1), and this is so even when there is certainty about  $\Theta$ . Suppose  $\theta_m < \frac{1}{2} < \theta_M$ . Then

$$\begin{aligned} U\left(\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2\right) &= \min_{P \in \mathcal{L}^2} P\left(\frac{1}{2}H_1T_2 + \frac{1}{2}T_1H_2\right) \\ &= \min_{\ell_1, \ell_2 \in \mathcal{L}} \frac{1}{2} \left( \ell_1(H_1)\ell_2(T_2) + \frac{1}{2}\ell_1(T_1)\ell_2(H_2) \right) \\ &= \min_{\ell_1, \ell_2 \in \mathcal{L}} \frac{1}{2} \left[ \ell_1(H_1)(1 - \ell_2(H_2)) + (1 - \ell_1(H_1))\ell_2(H_2) \right] \\ &= \min \{ \theta_m(1 - \theta_m), \theta_M(1 - \theta_M) \} \\ &\geq \theta_m(1 - \theta_M) = U(H_1T_2) = U(T_1H_2). \end{aligned}$$

The representation result leads to an interesting implication about the perceived value of repetition, which, at a mathematical level, extends the fact that for any random sequence  $(X_t)$  having an exchangeable probability law,  $X_i$  and  $X_j$  are positively correlated if  $i \neq j$ .<sup>19</sup>

**Theorem 4.3.** *If the multiple-priors preference  $\succeq$  satisfies Symmetry, Orthogonal Independence and Dominance, then, for any act  $f \in \mathcal{F}_{\{1, \dots, n\}}$ ,*

$$f \sim p \implies f \cdot \theta^n f \succeq p^2. \quad (4.7)$$

**Proof.** By the representation,

$$\begin{aligned} U(f \cdot \theta^n f) &= \int V(f \cdot \theta^n f) d\mu(V) = \int V(f) V(\theta^n f) d\mu(V) \\ &= \int (V(f))^2 d\mu(V) \geq \left( \int V(f) d\mu(V) \right)^2 = (U(f))^2 = p^2, \end{aligned}$$

where we use the fact that every  $V$  is symmetric (and hence also invariant to shifts) and a product utility function, and also the familiar property that the geometric average is at least as large as the arithmetic average.  $\blacksquare$

For simplicity, consider 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  $p$ . 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, as de Finetti pointed out in the Bayesian setting. Where there is a common element connecting experiments - like the uncertain bias of a coin that is tossed repeatedly - experiments are presumably viewed as “positively correlated”, which makes bets such as  $A \times A$  more attractive than two-fold independent replicas of the bet on  $A$ . This intuition relies only on the individual having a “conditionally i.i.d. (or IID)” view of experiments and not on the experiments conforming to a Bayesian (probabilistic) model.<sup>20</sup>

### 4.3. The Representation of Sets of Priors

The representation given in Theorem 4.2 is for utility functions, while de Finetti’s theorem is about beliefs. The former seems more appropriate for a decision-theoretic model, but it is interesting to consider also a formulation that is closer to de Finetti’s. In his theorem, every exchangeable measure is represented as a mixture of i.i.d. measures. Here, every set of priors consistent with our axioms is a (suitably defined) mixture of IID sets of priors.

To state this formally, define

$$\Gamma = \{ \mathcal{Q} \in \mathcal{K}^c(\Delta(\Omega)) : \mathcal{Q} \text{ is an IID set} \}.$$

<sup>19</sup>In Hewitt and Savage [21], see their Theorem 5.1.

<sup>20</sup>Even in the Bayesian case, we have not found intuition for (4.7) that relies solely on the axioms, without recourse to the representation.

Since  $\Gamma$  is homeomorphic to  $\mathcal{V}$ , each measure on  $\mathcal{V}$  corresponds to a unique measure on  $\Gamma$ , and we use the same symbol to denote both. Given  $\mu \in \Delta(\Gamma)$ , use Aumann's integral for a correspondence to define the set of priors  $\int_{\Gamma} \mathcal{Q} d\mu(\mathcal{Q})$ . (Technical details are provided in Appendix B, which also contains, in Section B.1, all the ingredients of a proof of the following Corollary.)

**Corollary 4.4.** *Let  $\succeq$  be represented by multiple-priors utility  $U$  as in (2.6) with set of priors  $\mathcal{P}$ . Then each hypothesis in Theorem 4.2 is equivalent to  $\mathcal{P}$  being expressible in the form<sup>21</sup>*

$$\mathcal{P} = cl \left( \int_{\Gamma} \mathcal{Q} d\mu(\mathcal{Q}) \right),$$

for the Borel probability measure  $\mu$  on  $\Gamma$  corresponding to the measure on  $\mathcal{V}$  appearing in (4.5).

When all IID sets have the form in (4.3), one obtains a representation even closer to de Finetti's. Then, with the obvious abuse of notation,

$$\mathcal{P} = cl \left( \int clh(\mathcal{L}^{\infty}) d\mu(\mathcal{L}) \right).$$

De Finetti's representation (2.1) is the special case where there is certainty that each set of likelihoods is a singleton, and hence that each experiment is described by the same likelihood. Here, by contrast, *multiple likelihoods* are associated with each experiment.<sup>22</sup>

#### 4.4. Ambiguity and Dissimilarity

The next theorem describes axiomatically the gap between de Finetti's model and ours.

**Theorem 4.5.** *Let the multiple priors preference  $\succeq$  satisfy Symmetry, Orthogonal Independence and Dominance. Then the following statements are equivalent:*

- (i)  $\succeq$  is an expected utility preference.
- (ii)  $\succeq$  satisfies Strong Exchangeability on the subdomain of acts over  $S_1 \times S_2$ , that is,

$$\alpha f + (1 - \alpha) \pi f \sim f \text{ for all } f \in \mathcal{F}_{\{1,2\}}.$$

- (iii)  $\succeq$  satisfies the Independence Axiom on the subdomain  $\mathcal{F}_1$  of acts over  $S_1$ .

(i) is the de Finetti model. The other conditions describe alternative characterizations of how it differs from ours. According to intuition given earlier, (ii) says that the first two experiments, and hence also any other pair, are perceived as identical. Following Gilboa and Schmeidler, we think of violations of Independence as reflecting (aversion to) ambiguity. Therefore, (iii) says that the first experiment is unambiguous. Conclude that our model permits any *two* experiments to be nonidentical by allowing ambiguity about

<sup>21</sup> $cl(\cdot)$  denotes closure. Below  $clh(\cdot)$  denotes closed convex hull.

<sup>22</sup>Contrast also with the representation (3.3), corresponding to Strong Exchangeability, where every experiment is described by the same likelihood but where there is ambiguity about which likelihood is the correct one.

any *single* experiment. This connection seems to us to be intuitive (however, see Example 3.7, for a specification where it is violated).

**Proof.** (i) $\implies$ (ii) : clear.

(ii) $\implies$ (iii) : Let  $U$  represent  $\succeq$  and assume (ii). For  $f, g \in \mathcal{F}_1$ ,

$$\begin{aligned}
& U\left(\frac{1}{8}[f \cdot \theta g + \theta g + f + 1 + g \cdot \theta f + \theta f + g + 1]\right) \\
&= U\left(\frac{1}{2}\left(\frac{1}{2}f + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta g + \frac{1}{2}\right) + \frac{1}{2}\left(\frac{1}{2}g + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta f + \frac{1}{2}\right)\right) \\
&= U\left(\left(\frac{1}{2}f + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta g + \frac{1}{2}\right)\right) \quad (\text{by (ii)}) \\
&= \frac{1}{2}U\left(\left(\frac{1}{2}f + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta g + \frac{1}{2}\right)\right) + \frac{1}{2}U\left(\left(\frac{1}{2}g + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta f + \frac{1}{2}\right)\right) \quad (\text{by Symmetry}) \\
&= \frac{1}{8}[U(f \cdot \theta g) + U(\theta g) + U(f) + U(1) + U(g \cdot \theta f) + U(\theta f) + U(g) + U(1)]. \\
&\quad (\text{by Orthogonal Independence})
\end{aligned}$$

Thus, by Lemma 2.3, there is a common minimizing measure for  $f$  and  $g$ , and (iii) follows.

(iii) $\implies$ (i) : By Theorem 4.2, there exists  $\mu \in \Delta(\mathcal{V})$  such that  $U(f) = \int V(f) d\mu(V)$ , for all  $f$  in  $\mathcal{F}$ . By (iii),

$$\int [V(\alpha f + (1 - \alpha)g) - \alpha V(f) - (1 - \alpha)V(g)] d\mu(V) = 0 \text{ for all } f, g \in \mathcal{F}_1.$$

Since the integrand is nonnegative for all  $V \in \mathcal{V}$ , conclude that: for all  $f, g \in \mathcal{F}_1$ , *a.s.*- $\mu[V]$ ,

$$V(\alpha f + (1 - \alpha)g) = \alpha V(f) + (1 - \alpha)V(g). \quad (4.8)$$

Thus it suffices to show that if  $V \in \mathcal{V}$  satisfies (4.8), then  $V$  is an expected utility function.

Assume  $V$  satisfies (4.8) and let  $\mathcal{P}$  be the corresponding set of measures for  $V$ . Write  $V(B)$  instead of  $V(1_B)$ . By the assumption, there exists  $\ell \in \Delta(S)$  such that, for all  $A \in \Sigma_1$ ,

$$V(A) = \ell(A).$$

Claim: If  $P \in \mathcal{P}$ , then  $P(A_1 \times A_2 \times \cdots \times A_n) = V(A_1 \times A_2 \times \cdots \times A_n)$  for all  $A_i \in \Sigma_i$ ,  $i \leq n$ .

Let  $A = A_1 \times A_2 \times \cdots \times A_n$ . Since  $S$  is finite,

$$\begin{aligned}
P(A) &= 1 - P(\Omega \setminus A) \leq 1 - V(\Omega \setminus A) \\
&\leq 1 - \sum_{(s_1, \dots, s_n) \notin A} V(\{(s_1, \dots, s_n)\}) \\
&= 1 - \sum_{(s_1, \dots, s_n) \notin A} \prod_{i=1}^n V(\{s_i\}) = 1 - \sum_{(s_1, \dots, s_n) \notin A} \prod_{i=1}^n \ell(\{s_i\}) \\
&= \sum_{(s_1, \dots, s_n) \in A} \prod_{i=1}^n \ell(\{s_i\}) = \prod_{i=1}^n \ell(\{A_i\}) = \prod_{i=1}^n V(\{A_i\}) \\
&= V(A).
\end{aligned}$$

But,  $P(A) \geq \min_{P' \in \mathcal{P}} P'(A) = V(A)$ . Thus,  $P(A) = V(A)$ .

Conclude that all  $P \in \mathcal{P}$  agree with  $V$ , and therefore, with one another, on finite rectangles. Since finite rectangles generate the Borel  $\sigma$ -algebra  $\Sigma$ ,  $\mathcal{P}$  is a singleton.  $\blacksquare$

#### 4.5. A “Technical” Point: Topology Matters

In Section 2.2, when introducing the basic multiple-priors model for an abstract state space, we restricted sets of priors to be compact in the weak convergence topology. We noted, however, that the stronger restriction of weak\*-compactness had been proposed (by Chateauneuf *et al* [7]). On the surface, the latter would seem to be a natural way to capture the counterpart of countable additivity of single measures, since weak\*-compactness of  $\mathcal{P} \subset ca(\Omega)$  corresponds exactly to the set of priors that enters into the Gilboa-Schmeidler functional form (2.2) consisting exclusively of countably additive measures. It may seem plausible also at the more meaningful behavioral level. Chateauneuf *et al* show that weak\*-compactness of  $\mathcal{P}$  is characterized behaviorally by Monotone Continuity:

*MONOTONE CONTINUITY:* Given  $f \succ g$ , outcome  $x$ , and a sequence  $\{A_n\}$  in  $\Sigma$ , with  $A_n \searrow \emptyset$ , then, there exists  $N$  such that  $(x, A_N; f(\cdot), \Omega \setminus A_N) \succ g$  and  $f \succ (x, A_N; g(\cdot), \Omega \setminus A_N)$ .

As the cited authors point out, this axiom is used by Arrow [3] to characterize countable additivity of the Savage prior. Moreover, though stronger than Regularity, Monotone Continuity is obviously much simpler.

However, in the setting of infinitely many experiments that are viewed symmetrically, Monotone Continuity is *too strong* - that is our interpretation of the following theorem.

**Theorem 4.6.** *If  $V$  is an IID utility function that satisfies Monotone Continuity, then  $V$  is an expected utility function (with an i.i.d. prior).*

The theorem shows that Monotone Continuity conflicts (in the sense of implying expected utility) with two assumptions - Symmetry, and (the perception of) stochastic independence of experiments in the sense of the product rule (4.1).<sup>23</sup> (Our other axioms - OI and Dominance - do not enter.) Thus it is too strong to admit even the important special case where experiments are viewed as “suitably” unrelated, but not identical.

One qualification is that one might argue that only a weaker form of stochastic independence is appropriate - instead of requiring

$$V(f \cdot g) = V(f)V(g),$$

for all (finitely-based) orthogonal acts, as in (4.1), one might require that it obtain only for acts that depend on (different) *single* experiments ( $f \in \mathcal{F}_i, g \in \mathcal{F}_j$  and  $i \neq j$ ), corresponding to experiments being *pairwise independent*. Maccheroni and Marinacci [26] assume such pairwise independence (albeit only for acts that are indicator functions). We feel, however, that in many cases, where pairwise independence would be intuitive, then so also would be the stronger notion (4.1). Therefore, at the very least, our theorem reflects on the limited scope of results about sequences of experiments (or random variables) that rely on Monotone Continuity.

---

<sup>23</sup>Recall that (4.1) has a simple behavioral meaning that was provided above.

Theorem 4.6 was surprising to us. However, there is reason to suspect that Monotone Continuity may be overly strong where ambiguity is important. For example, it implies that, for all measurable sets,<sup>24</sup>

$$\text{if } B_n \nearrow B, \text{ then } V(B_n) \nearrow V(B). \quad (4.9)$$

This is true in particular for all tail events  $B$ , which case is the central to the proof of the theorem. The worrisome feature is the assumption that tail events can be approximated in preference from *below*, and that this is so for *any* sequence of measurable sets that converges set-wise ( $\cup B_n = B$ ). As suggested when discussing Regularity, approximation from below may be unintuitive when modeling ambiguity averse individuals.

Contrast (4.9) with (Inner) Regularity, which imposes (4.9) only when  $B$  is open, (thereby excluding tail events), and every  $B_n$  is compact. On the other hand, universal approximation from *above*, in the sense that, for all measurable  $B_n$  and  $B$ ,

$$\text{if } B_n \searrow B, \text{ then } V(B_n) \searrow V(B),$$

is satisfied by every multiple-priors utility function as in (2.6), even those violating Regularity.<sup>25</sup> Though these differences may seem minor, the theorem shows that they are significant. For example, Regularity is consistent with the Walley-Fine IID utility function (4.4), but that model is excluded if Monotone Continuity is assumed.<sup>26</sup>

The key to the proof of the theorem is to show that Monotone Continuity, plus Symmetry and the stochastic independence condition (4.1), imply that  $V$  is 0-1 valued *and* additive on  $\Sigma^{\text{tail}}$ . That implies that all measures in  $\mathcal{P}$  are 0-1 valued and that they agree on  $\Sigma^{\text{tail}}$ . The rest is straightforward.

## 5. A CONCLUDING CONJECTURE

Our central model assumes that uncertainty about the coin's bias (or about the true IID utility  $V$  in (4.5), or about the true set of likelihoods  $\mathcal{L}$  in (4.6)) is probabilistic. The model based on Strong Exchangeability, on the other hand, permits prior ambiguity, but imposes that experiments are viewed as identical. It remains to formulate a general model that features both prior ambiguity, as in (3.3), and indistinguishable, but not necessarily identical, experiments.

The following functional form for utility seems to provide both elements:

$$U(f) = \inf_{\mu \in \mathcal{M}} \int_{\mathcal{Y}} V(f) d\mu, \quad (5.1)$$

<sup>24</sup>Let  $B_n \nearrow B$ . Define  $A_n = B \setminus B_n \searrow \emptyset$ ,  $f = 1_B$  and  $f_n = 1_{B_n} = (0, A_n; f, \Omega \setminus A_n)$ . Then Monotone Continuity implies that, for every  $\epsilon > 0$ , there exists  $N$  such that  $V(B_N) = V(f_N) > (1 - \epsilon)V(B)$ .

<sup>25</sup>If  $\mathcal{P} \subset ca_+^1(\Omega)$ , then  $\lim V(B_n) = \inf_n V(B_n) = \inf_n \inf_{\mathcal{P}} P(B_n) = \inf_{\mathcal{P}} \inf_n P(B_n) = \inf_{\mathcal{P}} P(B)$ . The last equality follows from the Monotone Convergence Theorem applied to every  $P$ .

<sup>26</sup>One can show directly that, for example,  $\mathcal{L}^\infty$  is not weak\*-compact unless  $\mathcal{L}$  is a singleton. For simplicity, consider  $S = \{H, T\}$ . Let  $\ell_0, \ell_1 \in \mathcal{L}$  and  $\ell_0 \neq \ell_1$ . For any  $r \in [0, 1]$ , we can find  $\{i_t\}_{t=1}^\infty \in \{0, 1\}^\infty$  such that  $\frac{1}{N} \sum i_t$  converges to  $r$ . Then, by [19, Theorem 2.19], the measure  $\otimes_t \ell_{i_t} \in \mathcal{L}^\infty$  assigns 1 to the event  $A_r$ , where  $A_r$  is the set that the limiting empirical frequency of Head is  $(1 - r)\ell_0(H) + r\ell_1(H)$ . If  $\mathcal{L}^\infty$  were weak\*-compact, there would be  $Q \in \Delta(S^\infty)$  such that  $Q(A) = 0$  implies  $P(A) = 0$  for all  $P \in \mathcal{L}^\infty$ , by [7, Lemma 3]. Thus,  $Q(A_r) > 0$  for all  $r \in [0, 1]$ , which cannot be true.

where  $\mathcal{M} \subset \Delta(\mathcal{V})$  is a set of probability measures over the set  $\mathcal{V}$  of IID utility functions. Moreover, in addition to Symmetry and Dominance, it satisfies the following intuitive axiom.<sup>27</sup>

**Axiom 5 (ORTHOGONAL INDIFFERENCE).** For all  $0 < \alpha \leq 1$ ,  $f^* \in \mathcal{F}_{I^*}$  and  $f \in \mathcal{F}_I$ , with  $I^*$  and  $I$  finite and disjoint, if  $\pi(I) \cap (I^* \cup I) = \emptyset$ , then

$$\alpha f^* \cdot f + (1 - \alpha) f^* \cdot (\pi f) \sim f^* \cdot f.$$

An example is the indifference

$$\frac{1}{2}T_1H_2 + \frac{1}{2}T_1H_3 \sim T_1H_2,$$

which can be understood in the now familiar way.

The axiom is weaker than both Strong Exchangeability and Orthogonal Independence (in the latter case, if we assume Symmetry). Strong Exchangeability implies that

$$\alpha f^* \cdot f + (1 - \alpha) \pi(f^* \cdot f) \sim f^* \cdot f,$$

for all acts and permutations. In particular, if  $f^*$  and  $f$  are as in the statement of Orthogonal Indifference, and if  $\pi$  is the identity on  $I^*$ , then  $\pi(f^* \cdot f) = f^* \cdot (\pi f)$ , which implies

$$\alpha f^* \cdot f + (1 - \alpha) f^* \cdot (\pi f) \sim f^* \cdot f,$$

for all admissible permutations. Second, the hypothesis in the new axiom implies that  $\pi f$  is orthogonal to  $f$ , and Symmetry implies that  $f \sim \pi f$ . Therefore, Orthogonal Independence implies  $\alpha f^* \cdot f + (1 - \alpha) f^* \cdot (\pi f) \sim f^* \cdot f$ .

We conjecture that the representation (5.1) is characterized by Symmetry, Orthogonal Indifference and Dominance.

---

<sup>27</sup>To verify Dominance, note that  $U_\mu$ ,  $U_\mu(f) = \int_{\mathcal{V}} V(f) d\mu$ , satisfies it for every  $\mu$ . Thus

$$U_\mu \left( \frac{1}{2}g^* \cdot \theta^{n+m} f' + \frac{1}{2}h^* \cdot \theta^m f \right) \geq U_\mu \left( \frac{1}{2}g^* \cdot \theta^m f + \frac{1}{2}h^* \cdot \theta^{n+m} f' \right).$$

Obviously the inequality is preserved by taking the infimum over  $\mu$ .

## A. Appendix: Regularity

*Proof of Theorem 2.2:*

$\Leftarrow$ : Chen [8, Proposition 1] proves that  $V$  is regular in the sense of Epstein-Wang, where

$$V(f) = \sup_{P \in \mathcal{P}} Pf = \max_{P \in cl(\mathcal{P})} Pf, f \in \mathcal{F}.$$

( $cl^*(\mathcal{P})$  denotes the weak\*-closure of  $\mathcal{P}$  in  $ba_+^1(\Omega)$ .) Therefore,  $U = V^*$ , the conjugate of  $V$ , satisfies Regularity.

$\Rightarrow$ : The multiple-priors utility function  $U$  can be extended in the obvious way to  $C(\Omega)$ , the set of all continuous real-valued functions on  $\Omega$ , and the extension is norm-continuous, superadditive, monotone, and  $U(1) = 1$ . Therefore, it is a support function for a unique compact and convex set  $\mathcal{P} \subset ca_+^1(\Omega)$ . In particular,

$$U(f) = \min_{\mathcal{P}} Pf, \text{ for every continuous act } f.$$

Let  $g \in \mathcal{F}^u$ . By Outer Regularity, there exist  $h_i \in \mathcal{F}^\ell$  such that

$$h_i \geq g \text{ and } U(h_i) < U(g) + 2^{-i}.$$

Further, there exist continuous acts  $f_i$  such that

$$h_i \geq f_i \geq g.$$

(When  $h_i = \mathbf{1}_{G_i}$  and  $g = \mathbf{1}_K$  are indicator acts, this follows from Urysohn's Lemma. More generally, the assertion follows from a straightforward extension of Urysohn's Lemma for simple acts - see [15, Lemma A.1].) Finally, it is wlog to assume that  $f_i \searrow g$  (see [1, Theorem 3.13]). It follows that

$$U(g) = \inf_i U(f_i) = \inf_i \inf_{\mathcal{P}} Pf_i = \inf_{\mathcal{P}} \inf_i Pf_i = \inf_{\mathcal{P}} Pg;$$

the last equality follows because  $\inf_i Pf_i = Pg$  for every  $P$  by the Monotone Convergence Theorem.

Define

$$\bar{U}(f) = \inf_{\mathcal{P}} Pf = \min_{cl(\mathcal{P})} Pf, f \in \mathcal{F}.$$

By the first part of the proof,  $\bar{U}$  is regular, while  $U$  is regular by assumption. As just shown, the two utility functions agree on  $\mathcal{F}^u$ . It follows immediately from Regularity that they must agree on all of  $\mathcal{F}$ . By the uniqueness of the (weak\*-compact and convex) set of priors, proven by Gilboa and Schmeidler,  $C$  is the weak\* closure of  $\mathcal{P}$ .  $\blacksquare$

The following lemma was used in the proof of Theorem 3.3.

**Lemma A.1.** *If  $U$  is regular, then so is  $\hat{U}$ , where  $\hat{U}(f) = U(\alpha f + (1 - \alpha)\pi f)$ ,  $f \in \mathcal{F}$ , for any fixed  $\alpha$  and  $\pi$ .*

**Proof.** Inner Regularity: Take  $h \in \mathcal{F}^\ell$ . It is clear that  $U(h) \geq \sup\{U(g) : g \leq h, g \in \mathcal{F}^u\}$ . Next show equality. By [1, Theorem 3.13], we can take continuous  $f_n$  such that  $f_n(\omega) \nearrow h(\omega)$  for each  $\omega \in \Omega$ . By [15, Lemma D.3], there exist finitely-based  $h'_n$  such that  $f_n \leq h'_n \leq h$ . Thus,

$$\begin{aligned} \lim_n U(\alpha h'_n + (1 - \alpha) \pi h'_n) &= \lim_n \min_{P \in \mathcal{P}} \int (\alpha h'_n + (1 - \alpha) \pi h'_n) dP \\ &= \min_{P \in \mathcal{P}} \lim_n \int (\alpha h'_n + (1 - \alpha) \pi h'_n) dP \\ &= \min_{P \in \mathcal{P}} \int (\alpha h + (1 - \alpha) \pi h) dP \\ &= U(\alpha h + (1 - \alpha) \pi h). \end{aligned}$$

The second equality follows from Terkelsen's minimax theorem [32, Corollary, p.407] and the third by the Monotone Convergence Theorem for each  $P$ .

Outer Regularity: Note that

$$\begin{aligned} U(\alpha f + (1 - \alpha) \pi f) &= \inf_{P \in \mathcal{P}} \int (\alpha f + (1 - \alpha) \pi f) dP \\ &= \inf_{P \in \mathcal{P}} \inf_{f \leq h \in \mathcal{F}^\ell} \int (\alpha h + (1 - \alpha) \pi h) dP \\ &= \inf_{f \leq h \in \mathcal{F}^\ell} \inf_{P \in \mathcal{P}} \int (\alpha h + (1 - \alpha) \pi h) dP \\ &= \inf_{f \leq h \in \mathcal{F}^\ell} U(\alpha h + (1 - \alpha) \pi h). \end{aligned}$$

The second equality follows because  $f \mapsto \int (\alpha f + (1 - \alpha) \pi f) dP$  satisfies Regularity. ■

## B. Appendix: Proof of Theorem 4.2

After proving necessity, the bulk of the proof concerns sufficiency of the axioms. Here we adapt the Hewitt-Savage [21] proof strategy for the de Finetti theorem to our setting. In broad terms, it amounts to showing that the set  $\mathcal{U}^*$  of multiple-priors utility functions satisfying Symmetry, OI and (a suitably weakened form of) Dominance is compact and convex, and then using the Choquet Theorem [27, p.14] to express any such utility function as an integral over extreme points of  $\mathcal{U}^*$ . The proof of uniqueness concludes.

### B.1. Necessity

Show first that the integral  $\int_{\mathcal{V}} V(f) d\mu(V)$  is well-defined for all  $f$  in  $\mathcal{F}$ . Denote by

$$\mathcal{Q}^{IID} = \{\mathcal{P} \in \mathcal{K}^c(\Delta(\Omega)) : U_{\mathcal{P}} \in \mathcal{V}\}$$

the set of all IID sets of priors. We show below that  $\mathcal{V}$ , and hence also  $\mathcal{Q}^{IID}$ , are compact, hence Borel measurable. Since  $\mu$  is well-defined on  $\bar{\Sigma}$ , the universal completion of  $\Sigma$ , it

suffices to show that the function  $V \mapsto V(f)$  is universally measurable. This is true if every set of the form

$$\{\mathcal{P} \in \mathcal{Q}^{IID} : \exists P \in \mathcal{P}, Pf < c\} = \text{proj}_{\Delta(\Omega)} \left( \{(\mathcal{P}, P) \in \mathcal{K}^c(\Delta(\Omega)) \times \Delta(\Omega) : \mathcal{P} \in \mathcal{Q}^{IID}, P \in \mathcal{P}, Pf < c\} \right)$$

lies in  $\bar{\Sigma}$ . But the set being projected is Borel-measurable (in the product  $\sigma$ -algebra). Therefore, the projection is universally measurable by the Lusin-Choquet-Meyer Theorem [22, Theorem A.1.8, p. 457].

Define  $U$  by (4.5), that is,

$$U(f) = \int_{\mathcal{V}} V(f) d\mu(V), \text{ for all } f \text{ in } \mathcal{F}.$$

We need to show that  $U$  is a (regular) multiple-priors utility function. To do so, we establish a suitable set of priors for  $U$ .

Since  $\mathcal{U}$  is homeomorphic to  $\mathcal{K}^c(\Delta(\Omega))$ ,  $\mu \in \Delta(\mathcal{V})$  can be viewed as a measure on  $\mathcal{K}^c(\Delta(\Omega))$ . Thus, we can write

$$U(f) = \int U_{\mathcal{Q}}(f) d\mu(\mathcal{Q}). \tag{B.1}$$

We define the Aumann integral  $\int \mathcal{Q} d\mu(\mathcal{Q})$  as follows: For a measurable  $\phi : \mathcal{K}^c(\Delta(\Omega)) \rightarrow \Delta(\Omega)$ , define  $\int \phi(\mathcal{Q}) d\mu(\mathcal{Q}) = \int \phi d\mu \in \Delta(\Omega)$ , by<sup>28</sup>

$$\left( \int \phi(\mathcal{Q}) d\mu(\mathcal{Q}) \right) (A) = \int \phi(\mathcal{Q})(A) d\mu(\mathcal{Q}) \text{ for all } A \in \Sigma.$$

Let  $\Phi$  be the identity function from  $\mathcal{K}^c(\Delta(\Omega))$  to  $\mathcal{K}^c(\Delta(\Omega))$  and  $\text{Sel } \Phi$  the set of all measurable selections from  $\Phi$ , that is,  $\phi \in \text{Sel } \Phi$  iff  $\phi$  is a measurable function from  $\mathcal{K}^c(\Delta(\Omega))$  to  $\Delta(\Omega)$  satisfying  $\phi(\mathcal{Q}) \in \mathcal{Q}$ . Then

$$\int \mathcal{Q} d\mu(\mathcal{Q}) \equiv \left\{ \int \phi d\mu : \phi \in \text{Sel } \Phi \right\}.$$

Below we use the next lemma, which can be proven by a standard argument using the Lebesgue Dominated Convergence Theorem.

**Lemma B.1.** *Let  $\phi : \mathcal{K}^c(\Delta(\Omega)) \rightarrow \Delta(\Omega)$  be measurable and  $P = \int \phi d\mu$ . Then, for any  $f \in \mathcal{F}$ ,*

$$\int_{\mathcal{K}^c(\Delta(\Omega))} \left[ \int_{\Omega} f d\phi(\mathcal{Q}) \right] d\mu(\mathcal{Q}) = \int f dP.$$

---

<sup>28</sup>The right side is well-defined because  $\mathcal{Q} \mapsto \mathcal{Q}(A)$  is measurable by [1, Lemma 15.16]. It is easy to see that one obtains a countably additive measure.

**Lemma B.2.** Let  $\mu \in \Delta(\mathcal{K}^c(\Delta(\Omega)))$ . Then, for all  $f \in \mathcal{F}$ ,

$$\begin{aligned} \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP &= \int_{\mathcal{K}^c(\Delta(\Omega))} \left( \inf_{P \in \mathcal{Q}} \int f dP \right) d\mu(\mathcal{Q}). \\ &= \int_{\mathcal{K}^c(\Delta(\Omega))} U_{\mathcal{Q}}(f) d\mu(\mathcal{Q}) \equiv U(f), \end{aligned}$$

where  $U$  is defined by (B.1).

**Proof.** We use a result of Castaldo et al [6, Theorem 3.2], which translated into our setup, states

$$\inf_{\phi \in \text{Sel } \Phi} \int_{\Delta(\Omega)} \hat{f}(P) d(\mu \circ \phi^{-1})(P) = \int \inf_{P \in \mathcal{Q}} \hat{f}(P) d\mu(\mathcal{Q})$$

for any measurable  $\hat{f} : \Delta(\Omega) \rightarrow \mathbb{R}$ . Given  $f \in \mathcal{F}$ , define  $\hat{f}(P) = \int f dP$ . Then

$$\begin{aligned} \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP &= \inf_{\phi \in \text{Sel } \Phi} \int_{\mathcal{K}^c(\Delta(\Omega))} \left[ \int_{\Omega} f d\phi(\mathcal{Q}) \right] d\mu(\mathcal{Q}) \\ &= \inf_{\phi \in \text{Sel } \Phi} \int_{\mathcal{K}^c(\Delta(\Omega))} \hat{f}(\phi(\mathcal{Q})) d\mu(\mathcal{Q}) \\ &= \inf_{\phi \in \text{Sel } \Phi} \int_{\Delta(\Omega)} \hat{f}(P) d(\mu \circ \phi^{-1})(P) \\ &= \int \inf_{P \in \mathcal{Q}} \hat{f}(P) d\mu(\mathcal{Q}) = \int_{\mathcal{K}^c(\Delta(\Omega))} \left[ \inf_{P \in \mathcal{Q}} \int f dP \right] d\mu(\mathcal{Q}), \end{aligned}$$

where the third equality follows by the Change of Variable Theorem [1, Theorem 13.46], and the fourth by the result cited above.  $\blacksquare$

**Lemma B.3.** Let  $\bar{\mu}$  denote the weak-convergence closure of  $\int \mathcal{Q} d\mu(\mathcal{Q})$ . Then:

- (i)  $\bar{\mu} \in \mathcal{K}^c(\Delta(\Omega))$ ; and
- (ii)  $\int U_{\mathcal{Q}}(f) d\mu(\mathcal{Q}) = \inf_{P \in \bar{\mu}} \int f dP$  for all  $f \in \mathcal{F}$ .

In other words, the utility function  $U$  defined as in (4.5), or equivalently, as in (B.1), is a regular multiple-priors utility function with set of priors  $\bar{\mu}$ .

**Proof.** (i) Convexity is clear because if  $\phi, \phi'$  are measurable selections, then so is any convex combination. The set  $\bar{\mu}$  is closed by construction.

- (ii) Define

$$\bar{U}(f) = \inf_{P \in \bar{\mu}} \int f dP, f \in \mathcal{F}.$$

By (i) and Theorem 2.2,  $\bar{U}$  is regular.

Step 1: By the preceding lemma,

$$U(f) \equiv \int U_{\mathcal{Q}}(f) d\mu(\mathcal{Q}) = \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP, \text{ for every } f \in \mathcal{F}.$$

Step 2:

$$\bar{U}(f) = \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP, \text{ for every lsc } f.$$

It is enough to show that  $\min_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP$  exists for every lsc  $f$ . By the preceding lemma,

$$\inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP = \int \left( \min_{P \in \mathcal{Q}} \int f dP \right) d\mu(\mathcal{Q}).$$

By the Measurable Maximum Theorem [1, Theorem 18.19], there is a measurable selection  $\phi, \phi(\mathcal{Q}) \in \operatorname{argmin}_{P \in \mathcal{Q}} \int f dP$  for each  $\mathcal{Q}$ . Then  $\bar{P} = \int \phi d\mu \in \int \mathcal{Q} d\mu(\mathcal{Q})$ , and

$$\int \left( \min_{P \in \mathcal{Q}} \int f dP \right) d\mu(\mathcal{Q}) = \int f d\bar{P}.$$

Step 3:

$$\bar{U}(f) = \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP, \text{ for every } f \in \mathcal{F}.$$

Argue as follows:

$$\int \mathcal{Q} d\mu(\mathcal{Q}) \subset \bar{\mu} \implies \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP \geq \bar{U}(f).$$

Next prove the reverse inequality. Since  $\bar{U}$  is regular, by Outer Regularity, given any  $f$  and  $\epsilon$ , there exists a simple lsc  $h$  such that

$$h \geq f \text{ and } \bar{U}(h) < \bar{U}(f) + \epsilon.$$

But then Step 2 and  $h \geq f \implies$

$$\inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP \leq \inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int h dP = \bar{U}(h) < \bar{U}(f) + \epsilon,$$

which proves the desired inequality

$$\inf_{P \in \int \mathcal{Q} d\mu(\mathcal{Q})} \int f dP \leq \bar{U}(f).$$

Combine Steps 1 and 3 to complete the proof. ■

## B.2. A Weakening of the Dominance Axiom

For the proof of sufficiency of the axioms, it is convenient (for reasons given shortly) to adopt a weakening of Dominance that is described here. Symmetry and Orthogonal Independence are assumed throughout.

Define the set of *strongly nonhedging* pairs of acts as the smallest set  $\mathcal{SNH} \subset \mathcal{F}_{fin} \times \mathcal{F}_{fin}$ , satisfying:

**SNH1**  $(f^* \cdot \theta^n f', f^* \cdot \theta^n f) \in \mathcal{SNH}$  for every  $f' \perp f$  and  $f^* \in \mathcal{F}_{\{1, \dots, n\}}$ ; and

**SNH2**  $(g^*, h^*) \in \mathcal{SNH} \implies (g^{**}, h^{**}) \in \mathcal{SNH}$ , if, for some  $F \in \mathcal{F}_{\{1, \dots, N\}}$ ,

$$g^{**} = \frac{1}{2}\theta^N g^* + \frac{1}{2}F \cdot \theta^N h^* \text{ and } h^{**} = \frac{1}{2}\theta^N h^* + \frac{1}{2}F \cdot \theta^N g^*. \quad (\text{B.2})$$

**Lemma B.4.** *Let  $(g^*, h^*) \in \mathcal{SNH}$ . Then:*

- (i)  $g^*$  and  $h^*$  are nonhedging (according to the definition in Section 3.3).
- (ii) If  $f' \perp f$ ,  $f', f \in \mathcal{F}_{\{1, \dots, n\}}$ , then

$$U\left(\frac{1}{2}f' \cdot \theta^n g^* + \frac{1}{2}f \cdot \theta^n h^*\right) = \frac{1}{2}U(f' \cdot \theta^n g^*) + \frac{1}{2}U(f \cdot \theta^n h^*).$$

**Proof.** (i) Let  $(g^*, h^*) \in \mathcal{SNH}$ ,  $f \in \mathcal{F}_{\{1, \dots, n\}}$ , and prove that

$$U\left(\frac{1}{2}f \cdot \theta^n g^* + \frac{1}{2}f \cdot \theta^n h^*\right) = \frac{1}{2}U(f \cdot \theta^n g^*) + \frac{1}{2}U(f \cdot \theta^n h^*).$$

This is obviously true, by OI, if  $g^* \perp h^*$ . It remains to show that if true for a pair  $(g^*, h^*)$ , then the equality is also true for  $(g^{**}, h^{**})$  defined by (B.2). Compute

$$\begin{aligned} U\left(\frac{1}{2}f \cdot \theta^n g^{**} + \frac{1}{2}f \cdot \theta^n h^{**}\right) &= U\left(f \cdot \left(\frac{1}{2}\mathbf{1} + \frac{1}{2}\theta^n F\right) \cdot \theta^{n+N} \left(\frac{1}{2}g^* + \frac{1}{2}h^*\right)\right) \\ (\text{by hypothesis for } (g^*, h^*)) &= \frac{1}{2}U\left(f \cdot \left(\frac{1}{2}\mathbf{1} + \frac{1}{2}\theta^n F\right) \cdot \theta^{n+N} g^*\right) \\ &\quad + \frac{1}{2}U\left(f \cdot \left(\frac{1}{2}\mathbf{1} + \frac{1}{2}\theta^n F\right) \cdot \theta^{n+N} h^*\right) \\ (\text{by OI}) &= \frac{1}{4}U\left(f \cdot \theta^{n+N} g^*\right) + \frac{1}{4}U\left(f \cdot \theta^n F \cdot \theta^{n+N} g^*\right) \\ &\quad + \frac{1}{4}U\left(f \cdot \theta^{n+N} h^*\right) + \frac{1}{4}U\left(f \cdot \theta^n F \cdot \theta^{n+N} h^*\right) \\ &= \frac{1}{2}\left[\frac{1}{2}U\left(f \cdot \theta^{n+N} g^*\right) + \frac{1}{2}U\left(f \cdot \theta^n F \cdot \theta^{n+N} h^*\right)\right] \\ &\quad + \frac{1}{2}\left[\frac{1}{2}U\left(f \cdot \theta^{n+N} h^*\right) + \frac{1}{2}U\left(f \cdot \theta^n F \cdot \theta^{n+N} g^*\right)\right] \\ (***) &= \frac{1}{2}U\left(\frac{1}{2}f \cdot \theta^{n+N} g^* + \frac{1}{2}f \cdot \theta^n F \cdot \theta^{n+N} h^*\right) \\ &\quad + \frac{1}{2}U\left(\frac{1}{2}f \cdot \theta^{n+N} h^* + \frac{1}{2}f \cdot \theta^n F \cdot \theta^{n+N} g^*\right) \\ &= \frac{1}{2}U\left(f \cdot \theta^n g^{**}\right) + \frac{1}{2}U\left(f \cdot \theta^n h^{**}\right). \end{aligned}$$

To see the justification for the starred equality, note that prior to it we established the additivity

$$\begin{aligned} &U\left(f \cdot \left(\frac{1}{2}\mathbf{1} + \frac{1}{2}\theta^n F\right) \cdot \theta^{n+N} \left(\frac{1}{2}g^* + \frac{1}{2}h^*\right)\right) \\ &= \frac{1}{4}U\left(f \cdot \theta^{n+N} g^*\right) + \frac{1}{4}U\left(f \cdot \theta^n F \cdot \theta^{n+N} h^*\right) + \frac{1}{4}U\left(f \cdot \theta^{n+N} h^*\right) + \frac{1}{4}U\left(f \cdot \theta^n F \cdot \theta^{n+N} g^*\right). \end{aligned}$$

Therefore, the starred equality follows from Lemma 2.3, which implies the additivity

$$\begin{aligned} &U\left(\frac{1}{2}f \cdot \theta^{n+N} g^* + \frac{1}{2}f \cdot \theta^n F \cdot \theta^{n+N} h^*\right) \\ &= \frac{1}{2}U\left(f \cdot \theta^{n+N} g^*\right) + \frac{1}{2}U\left(f \cdot \theta^n F \cdot \theta^{n+N} h^*\right). \end{aligned} \quad (\text{B.3})$$

- (ii) The proof is analogous to that of Lemma 3.9. ■

Denote by S-Dominance the axiom obtained from Dominance by replacing ‘‘hedging’’ by ‘‘strongly nonhedging.’’ Then Dominance implies S-Dominance. The latter is sufficient to derive the representation. We adopted the stronger axiom Dominance in the text because it is simpler to state and easier to interpret.

**Lemma B.5.**  $U$  satisfies S-Dominance iff, for all strongly nonhedging  $g^* \geq h^* \in \mathcal{F}_{\{1, \dots, m\}}$ , the function  $W : \mathcal{F}_{fin} \rightarrow \mathbb{R}$  defined by

$$W(f) = U(g^* \cdot \theta^m f) - U(h^* \cdot \theta^m f),$$

is increasing (satisfies Monotonicity) and concave.

**Proof.**  $W$  is monotone on  $\mathcal{F}_{fin}$  iff, for all  $n$  and  $f' \geq f \in \mathcal{F}_{\{1, \dots, n\}}$ ,

$$\begin{aligned} U(g^* \cdot \theta^m f') - U(h^* \cdot \theta^m f') &\geq U(g^* \cdot \theta^m f) - U(h^* \cdot \theta^m f) \iff \\ U(g^* \cdot \theta^m f') + U(h^* \cdot \theta^m f) &\geq U(g^* \cdot \theta^m f) + U(h^* \cdot \theta^m f') \iff \\ U(g^* \cdot \theta^{n+m} f') + U(h^* \cdot \theta^m f) &\geq U(g^* \cdot \theta^m f) + U(h^* \cdot \theta^{n+m} f') \iff^{***} \\ U\left(\frac{1}{2}g^* \cdot \theta^{n+m} f' + \frac{1}{2}h^* \cdot \theta^m f\right) &\geq U\left(\frac{1}{2}g^* \cdot \theta^m f + \frac{1}{2}h^* \cdot \theta^{n+m} f'\right) \iff \\ \frac{1}{2}g^* \cdot \theta^{n+m} f' + \frac{1}{2}h^* \cdot \theta^m f &\succeq \frac{1}{2}g^* \cdot \theta^m f + \frac{1}{2}h^* \cdot \theta^{n+m} f', \end{aligned}$$

which is case (i) of S-Dominance. The starred equivalence is implied by Lemma B.4(ii) and Symmetry.

Secondly,  $W$  is concave on  $\mathcal{F}_{fin}$  iff, for all  $n$  and  $F', F \in \mathcal{F}_{\{1, \dots, n\}}$ ,

$$\begin{aligned} &U(g^* \cdot \theta^m (\alpha F' + (1 - \alpha) F)) - U(h^* \cdot \theta^m (\alpha F' + (1 - \alpha) F)) \\ &\geq \alpha [U(g^* \cdot \theta^m F') - U(h^* \cdot \theta^m F')] + (1 - \alpha) [U(g^* \cdot \theta^m F) - U(h^* \cdot \theta^m F)] \iff \\ &[U(g^* \cdot \theta^m (\alpha F' + (1 - \alpha) F)) + \alpha U(h^* \cdot \theta^m F') + (1 - \alpha) U(h^* \cdot \theta^m F)] \\ &\geq [U(h^* \cdot \theta^m (\alpha F' + (1 - \alpha) F)) + \alpha U(g^* \cdot \theta^m F') + (1 - \alpha) U(g^* \cdot \theta^m F)] \iff \\ &[U(g^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F)) + \alpha U(h^* \cdot \theta^{n+m} F') + (1 - \alpha) U(h^* \cdot \theta^m F)] \\ &\geq [U(h^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F)) + \alpha U(g^* \cdot \theta^{n+m} F') + (1 - \alpha) U(g^* \cdot \theta^m F)] \iff \\ &[U(g^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F)) + U(h^* \cdot \theta^m (\alpha \theta^n F' + (1 - \alpha) F))] \\ &\geq [U(h^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F)) + U(g^* \cdot \theta^m (\alpha \theta^n F' + (1 - \alpha) F))] \iff \\ &U\left(\frac{1}{2}g^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F) + \frac{1}{2}h^* \cdot \theta^m (\alpha \theta^n F' + (1 - \alpha) F)\right) \\ &\geq U\left(\frac{1}{2}h^* \cdot \theta^{2n+m} (\alpha F' + (1 - \alpha) F) + \frac{1}{2}g^* \cdot \theta^m (\alpha \theta^n F' + (1 - \alpha) F)\right), \end{aligned}$$

which is case (ii) of S-Dominance. Once again, Lemma B.4(ii) is applied.  $\blacksquare$

**Lemma B.6.** If  $U$  satisfies S-Dominance then so does  $U^{**}$ , where, for some fixed  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$  with  $U(f^*) < 1$ ,

$$U^{**}(f) = \frac{U(f) - U(f^* \cdot \theta^m f)}{1 - U(f^*)}, \text{ for all } f \in \mathcal{F}.$$

**Proof.** Let  $(g^*, h^*) \in \mathcal{SNH}$ . Since the denominator  $1 - U(f^*)$  is not important, consider the function  $U_2$  defined by the numerator. Then

$$\begin{aligned}
U_2(g^* \cdot \theta^m f) - U_2(h^* \cdot \theta^m f) &= U(g^* \cdot \theta^m f) - U(f^* \cdot \theta^n (g^* \cdot \theta^m f)) \\
&\quad - [U(h^* \cdot \theta^m f) - U(f^* \cdot \theta^n (h^* \cdot \theta^m f))] \\
&= [U(g^* \cdot \theta^m f) + U(f^* \cdot \theta^n (h^* \cdot \theta^m f))] \\
&\quad - [U(h^* \cdot \theta^m f) + U(f^* \cdot \theta^n (g^* \cdot \theta^m f))] \\
&= [U(g^* \cdot \theta^m f) + U(f^* \cdot \theta^n h^* \cdot \theta^{n+m} f)] \\
&\quad - [U(h^* \cdot \theta^m f) + U(f^* \cdot \theta^n g^* \cdot \theta^{n+m} f)] \\
&= [U(\theta^n g^* \cdot \theta^{n+m} f) + U(f^* \cdot \theta^n h^* \cdot \theta^{n+m} f)] \\
&\quad - [U(\theta^n h^* \cdot \theta^{n+m} f) + U(f^* \cdot \theta^n g^* \cdot \theta^{n+m} f)] \\
(***) &= 2U\left(\left(\frac{1}{2}\theta^n g^* + \frac{1}{2}f^* \cdot \theta^n h^*\right) \cdot \theta^{n+m} f\right) \\
&\quad - 2U\left(\left(\frac{1}{2}\theta^n h^* + \frac{1}{2}f^* \cdot \theta^n g^*\right) \cdot \theta^{n+m} f\right) \\
&= 2[U(g^{**} \cdot \theta^{n+m} f) - U(h^{**} \cdot \theta^{n+m} f)],
\end{aligned}$$

where  $g^{**}$  and  $h^{**}$  are defined in (B.2). Both Symmetry and OI for  $U$  are used in deriving the above chain of equivalences. The starred equality follows from (B.3) and Symmetry. Note that  $(g^{**}, h^{**}) \in \mathcal{SNH}$  by construction of the latter. Also,  $g^{**} \geq h^{**}$ . Therefore, S-Dominance for  $U^{**}$  is implied by Lemma B.5.  $\blacksquare$

The Lemma shows that S-Dominance is inherited from  $U$  by  $U^{**}$ , which fact plays a key role below - see Lemma B.9 and the ensuing arguments. It is for this reason that we use S-Dominance - it is not clear to us how to prove that Dominance is similarly inherited.

### B.3. Sufficiency: The Hewitt-Savage Strategy Adapted

We turn to the sufficiency part of the theorem. Assume Symmetry, OI and S-Dominance (which is implied by Dominance).

We exploit heavily the homeomorphism between  $\mathcal{K}^c(\Delta(\Omega))$ , the space of sets of priors, and  $\mathcal{U} = \{U_{\mathcal{P}} : \mathcal{P} \in \mathcal{K}^c(\Delta(\Omega))\}$ , the space of (regular) multiple-priors utility functions. We pass freely between them. Recall also that  $\mathcal{K}^c(\Delta(\Omega))$ , and hence also  $\mathcal{U}$ , are compact metric.

The following preliminary results are straightforward.

**Lemma B.7.**  $\mathcal{P} \mapsto \min_{P \in \mathcal{P}} Pf$  is continuous for any continuous act  $f$ .

**Proof.** This is implied by the Maximum Theorem [1, Theorem 17.31].  $\blacksquare$

Define

$$\mathcal{U}^* = \{U \in \mathcal{U} : U \text{ satisfies Symmetry, OI and S-Dominance}\}.$$

**Lemma B.8.**  $\mathcal{U}^*$  is compact and convex, and  $\mathcal{V}$ , the subset of IID utility functions, is compact.

**Proof.** As noted,  $\mathcal{U}$  is compact. The further defining properties of  $\mathcal{U}^*$  and  $\mathcal{V}$  deal with finitely-based, and hence continuous, acts only. Therefore, the preceding lemma implies that each set is closed. Convexity of  $\mathcal{U}^*$  is obvious.  $\blacksquare$

The following lemma is the key to identifying the extreme points of  $\mathcal{U}^*$ . Much of the next subsection is concerned with proving the lemma. We continue here assuming the lemma is true.

**Lemma B.9.** For any  $U \in \mathcal{U}^*$  and  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$  satisfying  $U(f^*) \in (0, 1)$ , define the functions  $U^*$  and  $U^{**}$  by: for all  $f \in \mathcal{F}$ ,

$$\begin{aligned} U^*(f) &= \frac{U(f^* \cdot \theta^m f)}{U(f^*)}, \text{ and} \\ U^{**}(f) &= \frac{U(\theta^m f) - U(f^* \cdot \theta^m f)}{1 - U(f^*)}. \end{aligned}$$

Then  $U^*, U^{**} \in \mathcal{U}^*$ .

**Proposition B.10.** If  $U$  is an extreme point of  $\mathcal{U}^*$ , written  $U \in \text{ext}(\mathcal{U}^*)$ , then  $U \in \mathcal{V}$ .<sup>29</sup>

**Proof.** Let  $U \in \text{ext}(\mathcal{U}^*)$ . It suffices to show that

$$U(f^* \cdot \theta^m f) = U(f^*)U(f), \tag{B.4}$$

for every  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$  and  $f \in \mathcal{F}_{fin}$ .

Let  $\mathcal{P} \in \mathcal{K}^c(\Delta(S^\infty))$  be the set of priors corresponding to  $U$ . Consider three cases.

Case 1:  $U(f^*) = 0$ . Then  $\int f^* dP = 0$  for some  $P \in \mathcal{P}$ . Therefore,

$$f^* \geq 0 \implies f^*(\omega) = 0, \text{ } P\text{-a.s.} \implies (f^* \cdot \theta^m f)(\omega) = 0, \text{ } P\text{-a.s.},$$

which implies (B.4).

Case 2:  $U(f^*) = 1$ . Then  $\int f^* dP = 1$  for all  $P \in \mathcal{P}$ , and, again for all  $P$ ,

$$f^* \leq 1 \implies f^*(\omega) = 1, \text{ } P\text{-a.s.} \implies (f^* \cdot \theta^m f)(\omega) = \theta^m f(\omega), \text{ } P\text{-a.s.}$$

Therefore,

$$U(f^* \cdot \theta^m f) = U(\theta^m f) = U(f) = U(f^*)U(f),$$

where use has been made of the fact that Symmetry implies ‘‘shift invariance’’:

$$f \sim \theta f \text{ for every } f \in \mathcal{F}_{fin}.$$

Case 3:  $U(f^*) \in (0, 1)$ . For every  $f \in \mathcal{F}$ ,

$$U(f) = U(\theta^m f) = U(f^*)U^*(f) + (1 - U(f^*))U^{**}(f),$$

---

<sup>29</sup>We show later that the converse is also true -  $\text{ext}(\mathcal{U}^*) = \mathcal{V}$  - though we use only the fact that all extreme points lie in  $\mathcal{V}$ .

where  $U^*$  and  $U^{**}$  are defined in Lemma B.9. Thus

$$U = \alpha U^* + (1 - \alpha) U^{**}$$

where  $\alpha = U(f^*)$ . Since  $U$  is an extreme point of  $\mathcal{U}^*$  and  $U^*, U^{**} \in \mathcal{U}^*$ , we have  $U = U^*$ , and hence

$$U(f) = \frac{U(f^* \cdot \theta^m f)}{U(f^*)},$$

especially for  $f \in \mathcal{F}_I$  with finite  $I$ . This proves (B.4).  $\blacksquare$

We wish to apply the Choquet theorem [27, p.14]. For that purpose, note that  $\mathcal{U} \subset E \equiv \{\alpha U : \alpha \in \mathbb{R}, U \in \mathcal{U}\}$ , where  $E$  is a locally convex (vector) space under the topology generated by sets of the form  $\{\alpha U : a < \alpha < b, U \in G, G \text{ open in } \mathcal{U}\}$ . Now take  $U \in \mathcal{U}^*$ . Then Lemma B.8 and Choquet's theorem imply the existence of a Borel probability measure  $\mu$  on the set of extreme points of  $\mathcal{U}^*$  such that  $L(U) = \int L(V) d\mu(V)$  for every continuous linear functional  $L$  on  $E$ . Since  $\alpha U \mapsto \alpha U(f)$  is linear and continuous on  $E$  for every continuous  $f$ , it follows that

$$U(f) = \int V(f) d\mu(V) \tag{B.5}$$

for every continuous  $f$ . This, in fact, holds for any  $f \in \mathcal{F}$ : From the necessity proof, we know that  $f \mapsto \int V(f) d\mu(V)$  defines a utility function satisfying Regularity. In addition,  $U$  satisfies Regularity by assumption. Finitely-based acts are continuous since  $S$  is finite. Thus we can invoke the generalized Kolmogorov extension theorem in [15, Theorem D.2] to conclude that (B.5) holds for any  $f \in \mathcal{F}$ .

This completes the proof of sufficiency in Theorem 4.2, once we have proven Lemma B.9.

#### B.4. Remaining Arguments Re Extreme Points of $\mathcal{U}^*$

The main objective in this section is to prove Lemma B.9, namely that the two functions  $U^*$  and  $U^{**}$  defined there lie in  $\mathcal{U}^*$ .

That  $U^* \in \mathcal{U}^*$  is straightforward. First, we show that it is regular.

**Lemma B.11.** *For any  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$  with  $U(f^*) > 0$ , the function  $U^* : \mathcal{F} \rightarrow [0, 1]$ , defined by*

$$U^*(f) = \frac{U(f^* \cdot \theta^m f)}{U(f^*)}, \quad f \in \mathcal{F},$$

*satisfies Regularity.*

**Proof.** Show Outer Regularity. Inner Regularity can be shown in the same way.

View  $f^*$  also as a function of  $(s_1, \dots, s_m) \in S^m$ . By Regularity for  $U$ , there exist  $h_n \in \mathcal{F}^\ell$  such that  $h_n \geq f^* \cdot \theta^m f$  and  $U(h_n) \searrow U(f^* \cdot \theta^m f)$ . Define

$$h'_n(\omega) = \min_{s'_1, \dots, s'_m \in S} \left\{ \frac{h_n(s'_1, \dots, s'_m, \omega)}{f^*(s'_1, \dots, s'_m)} : f^*(s'_1, \dots, s'_m) > 0 \right\}, \quad \omega \in S^\infty.$$

Then  $h'_n \in \mathcal{F}^\ell$  by [1, Lemma 17.30]. We will show that

$$h_n(\omega) \geq f^*(\omega) \cdot \theta^m h'_n(\omega) \geq (f^* \cdot \theta^m f)(\omega) \text{ for each } \omega \in S^\infty. \quad (\text{B.6})$$

Fix  $\omega$ . If  $f^*(\omega) = 0$ , the inequality is clear. Assume  $f^*(\omega) > 0$ .

The first inequality in (B.6) holds because

$$\begin{aligned} f^*(\omega) \cdot \theta^m h'_n(\omega) &= f^*(s_1, \dots, s_m) \cdot h'_n(s_{m+1}, \dots) \\ &\leq f^*(s_1, \dots, s_m) \cdot \frac{h_n(s_1, \dots, s_m, s_{m+1}, \dots)}{f^*(s_1, \dots, s_m)} \\ &= h_n(s_1, \dots, s_m, s_{m+1}, \dots). \end{aligned}$$

For the second inequality,  $f^*(s'_1, \dots, s'_m) \cdot f(s'_{m+1}, \dots) \leq h_n(\omega')$  for each  $\omega' = (s'_1, s'_2, \dots)$ . Therefore,

$$f(s_{m+1}, \dots) \leq \frac{h_n(s_1, s_2, \dots)}{f^*(s_1, \dots, s_m)}$$

whenever  $f^*(s_1, \dots, s_m) > 0$ , and

$$f(s_{m+1}, \dots) \leq \min_{s_1, \dots, s_m \in S} \frac{h_n(s_1, s_2, \dots)}{f^*(s_1, \dots, s_m)} = h'_n(s_{m+1}, \dots),$$

which completes the proof of (B.6).

Finally, since  $U$  is monotone,  $U(h_n) \geq U(f^* \cdot \theta^m h'_n) \geq U(f^* \cdot \theta^m f)$ . Thus,

$$[U(h_n) \searrow U(f^* \cdot \theta^m f)] \implies [U(f^* \cdot \theta^m h'_n) \searrow U(f^* \cdot \theta^m f)],$$

which proves Outer Regularity for  $U^*$ . ■

It is evident that  $U^*$  is monotone, concave, and that it (or the preference that it represents) satisfies the Gilboa-Schmeidler axioms. Symmetry is satisfied because  $U(f^* \cdot \theta^m f) = U(f^* \cdot (\theta^m(\pi f)))$  for any permutation  $\pi$ , by Symmetry for  $U$ . For Orthogonal Independence, let  $f^{**} \in \mathcal{F}_{I^{**}}$ ,  $f \in \mathcal{F}_I$ ,  $f' \in \mathcal{F}_{I'}$  with finite and disjoint  $I^{**}, I, I'$ . Then

$$\begin{aligned} &U(f^* \cdot \theta^m [\alpha(f^{**} \cdot f) + (1 - \alpha)(f^{**} \cdot f')]) \\ &= U(\alpha(f^* \cdot \theta^m f^{**}) \cdot \theta^m f + (1 - \alpha)(f^* \cdot \theta^m f^{**}) \cdot \theta^m f') \\ &= \alpha U((f^* \cdot \theta^m f^{**}) \cdot \theta^m f) + (1 - \alpha) U((f^* \cdot \theta^m f^{**}) \cdot \theta^m f') \\ &= \alpha U(f^* \cdot \theta^m (f^{**} \cdot f)) + (1 - \alpha) U(f^* \cdot \theta^m (f^{**} \cdot f')). \end{aligned}$$

This implies OI for  $U^*$ . S-Dominance is also immediate. Conclude that  $U^* \in \mathcal{U}^*$ .

Next prove that  $U^{**} \in \mathcal{U}^*$ . This is more difficult because, roughly speaking,  $U^{**}$  is a difference of two functions derived from  $U$ . The axiom (S-)Dominance enters here. For example, Lemma B.5 implies, by taking  $g^* = 1$  and  $h^* = f^*$ , that  $U^{**}$  is increasing and concave on  $\mathcal{F}_{fin}$ , and Lemma B.6 implies that  $U^{**}$  satisfies S-Dominance. Symmetry and OI are immediate: we showed above that they are satisfied by  $U^*$ , and hence they are satisfied also by  $f \mapsto U(\theta^m f) - U(f^* \cdot \theta^m f)$ , which, apart from a scalar multiple, is  $U^{**}$ .

It remains to prove regularity and also that monotonicity and concavity obtain on all of  $\mathcal{F}$ . For this purpose we exploit the regularity of  $U$ , as described in the following lemmas. As the surrounding arguments are routine, many details are omitted.

Let  $\mathcal{F}_{fin}^\ell = \mathcal{F}^\ell \cap \mathcal{F}_{fin}$ , the set of (simple) lsc acts that are finitely-based.<sup>30</sup>

**Lemma B.12.** *Let  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$ . Then, for any  $f', f \in \mathcal{F}$  and  $\alpha \in [0, 1]$ ,*

$$U(f^* \cdot \theta^m(\alpha f + (1 - \alpha)f')) = \inf_{\substack{f \leq h \in \mathcal{F}^\ell \\ f' \leq h' \in \mathcal{F}^\ell}} U(f^* \cdot \theta^m(\alpha h + (1 - \alpha)h')).$$

**Proof.**  $\mathcal{P}$  denotes the set of priors corresponding to  $U$ . Note that

$$\begin{aligned} & U(\alpha(f^* \cdot \theta^m f) + (1 - \alpha)(f^* \cdot \theta^m f')) \\ &= \inf_{P \in \mathcal{P}} \left[ \alpha \int (f^* \cdot \theta^m f) dP + (1 - \alpha) \int (f^* \cdot \theta^m f') dP \right] \\ &= \inf_{P \in \mathcal{P}} \left[ \alpha \inf_{f \leq h \in \mathcal{F}^\ell} \int (f^* \cdot \theta^m h) dP + (1 - \alpha) \inf_{f' \leq h' \in \mathcal{F}^\ell} \int (f^* \cdot \theta^m h') dP \right] \\ &= \inf_{P \in \mathcal{P}} \inf_{\substack{f \leq h \in \mathcal{F}^\ell \\ f' \leq h' \in \mathcal{F}^\ell}} \int [\alpha(f^* \cdot \theta^m h) + (1 - \alpha)(f^* \cdot \theta^m h')] dP \\ &= \inf_{\substack{f \leq h \in \mathcal{F}^\ell \\ f' \leq h' \in \mathcal{F}^\ell}} \inf_{P \in \mathcal{P}} \int [\alpha(f^* \cdot \theta^m h) + (1 - \alpha)(f^* \cdot \theta^m h')] dP \\ &= \inf_{\substack{f \leq h \in \mathcal{F}^\ell \\ f' \leq h' \in \mathcal{F}^\ell}} U(\alpha(f^* \cdot \theta^m h) + (1 - \alpha)(f^* \cdot \theta^m h')). \end{aligned}$$

The second equality follows because  $f \mapsto \int f dP$ , for  $P \in \Delta(S^\infty)$ , is monotone and satisfies Regularity; hence Lemma B.11 implies  $\int f^* \cdot \theta^m f dP = \inf_{f \leq h \in \mathcal{F}^\ell} \int f^* \cdot \theta^m h dP$ . ■

**Lemma B.13.** *Let  $f^* \in \mathcal{F}_{\{1, \dots, m\}}$ .*

(a) *For any  $h \in \mathcal{F}^\ell$ , there exist  $h_n \in \mathcal{F}_{fin}^\ell$  such that  $h_n \leq h$ ,*

$$U(h_n) \nearrow U(h) \text{ and } U(f^* \cdot \theta^m h_n) \nearrow U(f^* \cdot \theta^m h).$$

(b) *For any  $f', f \in \mathcal{F}$  and  $\alpha \in [0, 1]$ , there exist  $h_n, h'_n \in \mathcal{F}^\ell$  such that*

$$\begin{aligned} & f \leq h_n, f \leq h'_n, U(h_n) \searrow U(f), U(h'_n) \searrow U(f'), \\ & U(f^* \cdot \theta^m h_n) \searrow U(f^* \cdot \theta^m h), U(f^* \cdot \theta^m h'_n) \searrow U(f^* \cdot \theta^m h'), \\ & U(\alpha h_n + (1 - \alpha)h'_n) \searrow U(\alpha f + (1 - \alpha)f') \text{ and} \\ & U(f^* \cdot \theta^m(\alpha h_n + (1 - \alpha)h'_n)) \searrow U(f^* \cdot \theta^m(\alpha f + (1 - \alpha)f')). \end{aligned}$$

<sup>30</sup>Since  $S$  is finite, every finitely-based act is continuous, hence lsc. However, we use the notation  $\mathcal{F}_{fin}^\ell$  in order to emphasize that we are using the lower semi-continuity of such acts, which would be important in any future generalization to infinite  $S$ .

**Proof.** (a) By Inner Regularity, there is a sequence  $g_n \in \mathcal{F}^u$  such that  $g_n \leq h$  and  $U(g_n) \nearrow U(h)$ . By [15, Lemma D.3], there exists  $h'_n \in \mathcal{F}_{fin}^\ell$  such that  $g_n \leq h'_n \leq h$ . Then  $U(h'_n) \nearrow U(h)$ . Similarly, by the regularity established in Lemma B.11, there exist  $h''_n \in \mathcal{F}_{fin}^\ell$  such that  $h''_n \leq h$  and  $U(f^* \cdot \theta^m h''_n) \nearrow U(f^* \cdot \theta^m h)$ . Define  $h_n = \max\{h'_n, h''_n\}$  and  $h_n$  does the job.

(b) By Regularity of  $U$ , there is  $\widehat{h}_n \in \mathcal{F}^\ell$  such that  $\widehat{h}_n \geq f$  and  $U(\widehat{h}_n) \searrow U(f)$ .

By the regularity established in Lemma B.11, there exist  $\widehat{\widehat{h}}_n \in \mathcal{F}^\ell$  such that  $\widehat{\widehat{h}}_n \geq f$  and  $U(f^* \cdot \theta^m \widehat{\widehat{h}}_n) \searrow U(f^* \cdot \theta^m f)$ . Define  $h_n \in \mathcal{F}^\ell$  by

$$h_n = \min \left\{ \widehat{h}_n, \widehat{\widehat{h}}_n \right\}.$$

Then  $h_n \in \mathcal{F}^\ell$ ,  $h_n \geq f$  and

$$U(h_n) \searrow U(f) \text{ and } U(f^* \cdot \theta^m h_n) \searrow U(f^* \cdot \theta^m f).$$

The preceding argument is readily extended to prove the remainder of (b), when combined with Lemma B.12.  $\blacksquare$

We can finally complete the proof of Lemma B.9.

*Monotonicity:*  $U^{**}(f') \geq U^{**}(f)$  if  $f' \geq f$  and  $f', f \in \mathcal{F}$ . We have observed that this is true if  $f'$  and  $f$  are finitely-based. The inequality is readily extended to all simple lsc acts, and then to arbitrary acts, by using Lemma B.13.

*Regularity:* Since  $U^{**}$  is increasing,  $U^{**}(f) \leq \inf \{U^{**}(h) : h \geq f, h \in \mathcal{F}^\ell\}$ . Lemma B.13 implies equality, which proves Outer Regularity. Inner Regularity can be shown similarly.

*Concavity:* We have to show that

$$U^{**}(\alpha f + (1 - \alpha) f') \geq \alpha U^{**}(f) + (1 - \alpha) U^{**}(f'), \text{ for all } f', f \in \mathcal{F}.$$

This has been shown for all finitely-based  $f'$  and  $f$ . The inequality is readily extended to all simple lsc acts, and then to arbitrary acts, by using Lemma B.13.

We offer a remark related to the proof. Above we showed that every extreme point of  $\mathcal{U}^*$  lies in  $\mathcal{V}$ . In fact, we can prove, using the representation, that the other direction is also true.

**Lemma B.14.**  $\mathcal{V}$  is the set of all extreme points of  $\mathcal{U}^*$ .

**Proof.** Let  $U \in \mathcal{V}$  and show that  $U$  is an extreme point of  $\mathcal{U}^*$ .

The proof of Theorem 4.2, specifically, application of Choquet's Theorem, implies that  $U(f) = \int V(f) d\mu(V)$  for some  $\mu$  that is supported by the set of extreme points of  $\mathcal{U}^*$  (and not only by its superset  $\mathcal{V}$ ). Therefore, for  $f \in \mathcal{F}_{\{1, \dots, m\}}$ ,

$$\begin{aligned} \left[ \int V(f) d\mu(V) \right]^2 &= [U(f)]^2 = U(f \cdot \theta^m f) \\ &= \int V(f \cdot \theta^m f) d\mu(V) = \int [V(f)]^2 d\mu(V). \end{aligned}$$

But  $[\int V(f) d\mu(V)]^2 = \int [V(f)]^2 d\mu(V)$  if and only if

$$V(f) \text{ is constant } \mu\text{-a.s.}[V].$$

The exceptional set depends on  $f$ . But since  $\mathcal{F}_1$  is separable, there exists a  $\mu$ -null set of  $V$ 's that works for all acts. Conclude that  $a.s.\text{-}\mu[V]$ ,  $V(\cdot) = U(\cdot)$  on  $\mathcal{F}_{\{1,\dots,m\}}$ . Since this is true for any  $m$ , the equality holds  $a.s.$  on all of  $\mathcal{F}$  by the generalized Kolmogorov extension theorem [15, Theorem D.2]. Thus,  $\mu$  is degenerate and  $U$  is an extreme point of  $\mathcal{U}^*$ . ■

### B.5. Uniqueness

Let  $\mu'$  and  $\mu$ , Borel measures on the compact metric space  $\mathcal{V}$ , satisfy

$$\int V(f) d\mu' = \int V(f) d\mu \text{ for all } f \in \mathcal{F}.$$

We show that

$$\mu' = \mu.$$

Each finitely-based act  $f$  induces (by Lemma B.7) the continuous map  $\hat{f}: \mathcal{V} \rightarrow [0, 1]$ , given by

$$\hat{f}(V) = V(f).$$

Let  $\hat{\mathcal{F}}_{fin}$  be the set of all such maps and  $\mathcal{A} = sp(\hat{\mathcal{F}}_{fin})$ , the linear span of  $\hat{\mathcal{F}}_{fin}$  within  $C(\mathcal{V})$ , the set of continuous real-valued functions on  $\mathcal{V}$ . Then,

$$\int \hat{f}(V) d\mu' = \int \hat{f}(V) d\mu \text{ for all } \hat{f} \in \hat{\mathcal{F}}_{fin}.$$

This equality extends also to the linear span:

$$\int \phi(V) d\mu' = \int \phi(V) d\mu \text{ for all } \phi \in \mathcal{A}.$$

It is enough to show that

$$\int \phi(V) d\mu' = \int \phi(V) d\mu \text{ for all } \phi \in C(\mathcal{V}). \tag{B.7}$$

We do this by verifying the conditions of the Stone-Weierstrass Theorem, which implies that  $\mathcal{A}$  is sup-norm dense in  $C(\mathcal{V})$ , and hence also (B.7).

Obviously  $\mathcal{A}$  contains the constant functions and it separates points; in fact, since every IID utility is regular, if  $V' \neq V$ , then  $\phi(V') \neq \phi(V)$  for some  $\phi \in \hat{\mathcal{F}}_{fin} \subset \mathcal{A}$ . We need only show that

$$\phi', \phi \in \mathcal{A} \implies \phi'\phi \in \mathcal{A},$$

which follows from Steps 1 and 2.

*Step 1.* Any finite linear combination of elements in  $\hat{\mathcal{F}}_{fin}$  can be expressed as a linear combination of two such elements, that is,

$$\sum_i a_i \hat{f}_i = \kappa \hat{h} - \kappa' \hat{h}'. \tag{B.8}$$

Clearly,

$$\left(\widehat{\sum_i a_i f_i}\right)(V) = \sum_i a_i \widehat{f_i}(V) = \sum_i a_i V(f_i).$$

Suppose that every  $a_i$  is positive. We can shift each of the acts  $f_i$  so that they are mutually orthogonal and  $V$  is additive over them (since every IID utility satisfies OI). Because weights may not sum to 1, we obtain  $\kappa V(h)$  for some finitely based act  $h$  and  $\kappa > 0$ , that is,

$$\sum_i a_i \widehat{f_i} = \kappa \widehat{h}.$$

If one or more of the coefficients  $a_i$  is negative, then one can collect those acts having similarly signed weights, and derive (B.8).

*Step 2.* Verify that  $(a\widehat{f} + b\widehat{g})(a'\widehat{f}' + b'\widehat{g}') \in \mathcal{A}$ :

$$\begin{aligned} & [(a\widehat{f} + b\widehat{g})(V)] [(a'\widehat{f}' + b'\widehat{g}')(V)] \\ &= [aV(f) + bV(g)] [a'V(f') + b'V(g')] \\ &= aa'V(f \cdot \theta^n f') + ab'V(f \cdot \theta^n g') + ba'V(g \cdot \theta^n f') + bb'V(g \cdot \theta^n g') \\ &= \left( aa'(f \cdot \widehat{\theta^n f'}) + ab'(f \cdot \widehat{\theta^n g'}) + ba'(g \cdot \widehat{\theta^n f'}) + bb'(g \cdot \widehat{\theta^n g'}) \right) (V), \end{aligned}$$

where  $n$  is large enough so that all paired acts are orthogonal to one another. The last equality is derived by shifting each of the product acts so that they are mutually orthogonal, so that  $V$  is additive over them, and then applying shift invariance. Thus (B.8) implies

$$(a\widehat{f} + b\widehat{g})(a'\widehat{f}' + b'\widehat{g}') = \kappa \widehat{h} - \kappa' \widehat{h}' \in \mathcal{A}. \quad \blacksquare$$

## C. Appendix: Proof of Theorem 4.6

*Step 1:*  $V(g \cdot f) = V(g)V(f)$  for all  $g \in \mathcal{F}_{\{1, \dots, m\}}$  and  $f \in \mathcal{F}_{\{m+1, m+2, \dots\}}$ .

The equality is true by (4.1) if  $f$  is finitely-based. Extend it to all acts  $f$  indicated by applying Regularity.

*Step 2:* Fix  $A \in \Sigma^{tail}$  and define, (where  $A$  denotes  $\mathbf{1}_A$  and so on),

$$\mathcal{B} = \left\{ B \in \Sigma : V\left(\left(\frac{1}{2}B + \frac{1}{2}\right) \cdot \left(\frac{1}{2}A + \frac{1}{2}\right)\right) = V\left(\frac{1}{2}B + \frac{1}{2}\right) V\left(\frac{1}{2}A + \frac{1}{2}\right) \right\}.$$

Then  $\mathcal{B}$  is a monotone class.

(a) Assume  $B_n \in \mathcal{B}$ ,  $B_n \nearrow B$  and show  $B \in \mathcal{B}$ , that is,

$$V\left(\left(\frac{1}{2}B + \frac{1}{2}\right) \cdot \left(\frac{1}{2}A + \frac{1}{2}\right)\right) = V\left(\frac{1}{2}B + \frac{1}{2}\right) V\left(\frac{1}{2}A + \frac{1}{2}\right).$$

Let  $C_n = B \setminus B_n \searrow \emptyset$ , and define, for a fixed tail event  $A'$ ,

$$\begin{aligned} f &= \left(\frac{1}{2}B + \frac{1}{2}\right) \cdot \left(\frac{1}{2}A' + \frac{1}{2}\right), \\ f_n &= \left(\frac{1}{2}B_n + \frac{1}{2}\right) \cdot \left(\frac{1}{2}A' + \frac{1}{2}\right) \text{ and} \\ g_n &= \left(\frac{1}{4}, C_n; f_n, \Omega \setminus C_n\right). \end{aligned}$$

(a.i) If  $s \in C_n$ , then  $s \notin B_n$  and  $f_n(s) = \frac{1}{2} (\frac{1}{2}A' + \frac{1}{2})(s) \in \{\frac{1}{4}, \frac{1}{2}\}$ .

(a.ii) By (a.i),  $g_n \leq f_n$ . Therefore,  $V(g_n) \leq V(f_n)$ .

(a.iii)  $f_n(s) \neq f(s) \implies$

$$\begin{aligned} & [(\frac{1}{2}B_n + \frac{1}{2})(\frac{1}{2}A' + \frac{1}{2})](s) \neq [(\frac{1}{2}B + \frac{1}{2})(\frac{1}{2}A' + \frac{1}{2})](s) \\ & \implies (\frac{1}{2}B_n + \frac{1}{2})(s) \neq (\frac{1}{2}B + \frac{1}{2})(s) \implies s \in B \setminus B_n = C_n. \end{aligned}$$

Therefore,  $s \notin C_n \implies f_n(s) = f(s)$ .

(a.iv)  $g_n = (\frac{1}{4}, C_n; f, \Omega \setminus C_n)$ . This is clear by (a.iii).

(a.v) By Monotone Continuity, for any  $\epsilon > 0$ , there exists  $N$  such that  $V(g_N) > V(f) - \epsilon$ .

Therefore, by (a.ii),  $V(f_N) > V(f) - \epsilon$ . But  $f_n \nearrow^n$ . Conclude that

$$V(f_n) \nearrow V(f). \tag{C.1}$$

We can now complete the proof of (a) and show that  $B \in \mathcal{B}$ : (C.1) implies that

$$\lim V((\frac{1}{2}B_n + \frac{1}{2}) \cdot (\frac{1}{2}A' + \frac{1}{2})) = V((\frac{1}{2}B + \frac{1}{2})(\frac{1}{2}A' + \frac{1}{2}))$$

for all  $A' \in \Sigma^{tail}$ . Thus

$$\begin{aligned} V((\frac{1}{2}B + \frac{1}{2}) \cdot (\frac{1}{2}A + \frac{1}{2})) &= \lim V((\frac{1}{2}B_n + \frac{1}{2})(\frac{1}{2}A + \frac{1}{2})) \quad (\text{set } A' = A) \\ &= \lim V(\frac{1}{2}B_n + \frac{1}{2}) V(\frac{1}{2}A + \frac{1}{2}) \quad (\text{since } B_n \in \mathcal{B}) \\ &= V(\frac{1}{2}B + \frac{1}{2}) V(\frac{1}{2}A + \frac{1}{2}). \quad (\text{set } A' = \Omega) \end{aligned}$$

(b) Assume  $B_n \in \mathcal{B}$ ,  $B_n \searrow B$  and show that  $B \in \mathcal{B}$ . The argument is similar to that in (a). We provide an outline for completeness.

Let  $C_n = B_n \setminus B \searrow \emptyset$ , and define, for a fixed tail event  $A'$ ,

$$\begin{aligned} f &= (\frac{1}{2}B + \frac{1}{2}) \cdot (\frac{1}{2}A' + \frac{1}{2}), \\ f_n &= (\frac{1}{2}B_n + \frac{1}{2}) \cdot (\frac{1}{2}A' + \frac{1}{2}) \text{ and} \\ g_n &= (1, C_n; f_n, \Omega \setminus C_n). \end{aligned}$$

(b.i) If  $s \in C_n$ , then  $s \in B_n$  and  $f_n(s) = (\frac{1}{2}A' + \frac{1}{2})(s) \in \{\frac{1}{2}, 1\}$ .

(b.ii) By (b.i),  $g_n \geq f_n$ . Therefore,  $V(g_n) \geq V(f_n)$ .

(b.iii)  $f_n(s) \neq f(s) \implies$

$$\begin{aligned} & [(\frac{1}{2}B_n + \frac{1}{2}) \cdot (\frac{1}{2}A' + \frac{1}{2})](s) \neq [(\frac{1}{2}B + \frac{1}{2}) \cdot (\frac{1}{2}A' + \frac{1}{2})](s) \\ & \implies (\frac{1}{2}B_n + \frac{1}{2})(s) \neq (\frac{1}{2}B + \frac{1}{2})(s) \implies s \in B_n \setminus B = C_n. \end{aligned}$$

Therefore,  $s \notin C_n \implies f_n(s) = f(s)$ .

(b.iv)  $g_n = (1, C_n; f, \Omega \setminus C_n)$ . This is clear by (b.iii).

(b.v)  $V(f_n) \searrow V(f)$ .

The rest of the argument is exactly as in (a).

*Step 3:* By Step 1,  $\cup_m \Sigma_{\{1, \dots, m\}} \subset \mathcal{B}$ . Thus the Monotone Class Lemma [1, p. 137] implies that  $\mathcal{B} = \Sigma$ , that is, for all  $A \in \Sigma^{tail}$  and  $B \in \Sigma$ ,

$$V\left(\left(\frac{1}{2}B + \frac{1}{2}\right) \cdot \left(\frac{1}{2}A + \frac{1}{2}\right)\right) = V\left(\frac{1}{2}B + \frac{1}{2}\right) V\left(\frac{1}{2}A + \frac{1}{2}\right).$$

In the same way we can show that, for all  $A \in \Sigma^{tail}$  and  $B \in \Sigma$ ,

$$V(1_A \cdot 1_B) = V(1_{A \cap B}) = V(A \cap B) = V(A)V(B). \quad (\text{C.2})$$

The rest of the proof uses these properties and not Monotone Continuity directly.

*Step 4:* Apply Step 4 to two tail events  $A$  and  $B$  to derive

$$\begin{aligned} V\left(\frac{1}{4}1_B \cdot 1_A + \frac{1}{4}1_A + \frac{1}{4}1_B + \frac{1}{4}\right) &= V\left(\left(\frac{1}{2}1_B + \frac{1}{2}\right) \cdot \left(\frac{1}{2}1_A + \frac{1}{2}\right)\right) \\ &= V\left(\frac{1}{2}1_B + \frac{1}{2}\right) V\left(\frac{1}{2}1_A + \frac{1}{2}\right) \\ &= \frac{1}{4}[V(A)V(B) + V(A) + V(B) + 1] \\ &= \frac{1}{4}[V(A \cap B) + V(A) + V(B) + 1]. \end{aligned}$$

By Lemma 2.3,

$$V\left(\frac{1}{2}1_A + \frac{1}{2}1_B\right) = \frac{1}{2}V(A) + \frac{1}{2}V(B).$$

*Step 5:*  $A \mapsto V(A)$  defines a finitely additive 0-1 valued measure (or charge) on  $\Sigma^{tail}$ . The 0-1 property follows from (C.2). For disjoint  $A, B \in \Sigma^{tail}$ , by Step 4,

$$\begin{aligned} V(A \cup B) &= V(1_{A \cup B}) = V(1_A + 1_B) \\ &= 2V\left(\frac{1}{2}1_A + \frac{1}{2}1_B\right) = V(A) + V(B). \end{aligned}$$

*Step 6:* Let  $\mathcal{P}$  be the set of priors corresponding to  $V$ . For  $A \in \Sigma^{tail}$ ,  $V(A) + V(\Omega \setminus A) = 1$ . Thus,  $V(A) = 0 \implies V(\Omega \setminus A) = 1$ . Further,  $V(A) = 1 \implies P(A) = 1$  for all  $P \in \mathcal{P}$ . Since  $V(A) = 0$  or  $1$ , it follows that

$$\{P(A) : P \in \mathcal{P}\} = \{0\} \text{ or } \{1\}.$$

*Step 7:* For each  $f \in \mathcal{F}_1$ , there is an exchangeable measure  $P^*$  that is minimizing for  $f$ . To see this, note that

$$\begin{aligned} V\left(\frac{1}{4}f \cdot \theta f + \frac{1}{4}f + \frac{1}{4}\theta f + \frac{1}{4}\right) &= V\left(\left(\frac{1}{2}f + \frac{1}{2}\right) \cdot \left(\frac{1}{2}\theta f + \frac{1}{2}\right)\right) \\ &= V\left(\frac{1}{2}f + \frac{1}{2}\right) V\left(\frac{1}{2}\theta f + \frac{1}{2}\right) \\ &= \frac{1}{4}[V(f \cdot \theta f) + V(f) + V(\theta f) + 1]. \end{aligned}$$

By Lemma 2.3, there is a common minimizing measure  $P$  for  $f$  and  $\theta f$ . Let  $\pi$  be the permutation that switches experiments 1 and 2. Then, using Symmetry,

$$(\pi P)f = P(\pi f) = P(\theta f) = V(\theta f) = V(f).$$

Therefore,  $P$  and  $\pi P$  are both minimizing for  $f$ . Finally,  $P^1 \equiv \frac{1}{2}P + \frac{1}{2}\pi P$  is also minimizing (it lies in  $\mathcal{P}$  because  $\mathcal{P}$  is convex) and it satisfies  $\pi P^1 = P^1$ .

Apply a similar argument to  $(\frac{1}{2}f + \frac{1}{2}) \cdot (\frac{1}{2}\theta f + \frac{1}{2}) \dots \cdot (\frac{1}{2}\theta^n f + \frac{1}{2})$  to deduce that there is a common minimizing measure  $P^n$  for  $\{f, \theta f, \dots, \theta^n f\}$  that satisfies  $\pi P^n = P^n$  for all  $\pi \in \Pi^n$ , the set of permutations on  $\{1, \dots, n\}$ . Since  $\mathcal{P}$  is compact, wlog (after relabelling),  $P^n \rightarrow P^* \in \mathcal{P}$ . Then  $P^*$  is exchangeable and minimizing for  $f$ .

*Step 8:* The measure  $P^*$  in Step 8 is i.i.d.: By Step 6,  $P^*$  is 0-1 valued on  $\Sigma^{tail}$ . But, using the de Finetti Theorem, it is straightforward to show that the only exchangeable measures with this property are i.i.d. measures.

*Step 9.*  $V(\alpha f' + (1 - \alpha)f) = \alpha V(f') + (1 - \alpha)V(f)$ , for all  $f', f \in \mathcal{F}_1$ .

Take i.i.d. measures  $P'$  for  $f'$  and  $P$  for  $f$ . Since both  $P'$  and  $P$  are i.i.d. measures, and they agree on tail events (Step 6), they must coincide. Thus, there is a common minimizing measure for  $f'$  and  $f$ .

*Step 10.*  $V(\alpha f' + (1 - \alpha)f) = \alpha V(f') + (1 - \alpha)V(f)$ , for all  $f', f \in \mathcal{F}$ .

For any  $n$ , view  $S^n$  as corresponding to one experiment and repeat the above to derive additivity for all  $f', f \in \mathcal{F}_{\{1, \dots, n\}}$ . Finally, apply Regularity to extend additivity to all acts. ■

## References

- [1] C.D. Aliprantis and K.C. Border, *Infinite Dimensional Analysis*, 3rd edition, Springer, 2006.
- [2] S. Arnold, Linear models with exchangeably distributed errors, *JASA* 74 (1979), 194-199.
- [3] K. Arrow, *Essays in the Theory of Risk-Bearing*, North-Holland, 1970.
- [4] J.M. Bernardo, The concept of exchangeability and its applications, *Far East J. Math. Sci.* 4 (1996), 111-121.
- [5] W. Brock and S.N. Durlauf, Growth empirics and reality, *The World Bank Review* 15 (2001), 229-272.
- [6] A. Castaldo, F. Maccheroni and M. Marinacci, Random correspondences as bundles of random variables, *Sankhya* (Series A) 80 (2004), 409-427.
- [7] A. Chateauneuf, F. Maccheroni, M. Marinacci and J.M. Tallon, Monotone continuous multiple priors, *Econ. Theory* 26 (2005), 973-982.
- [8] Y.C. Chen, Universality of the Epstein-Wang type structure, Northwestern, revised 2008.
- [9] D. Draper, Comment: on the exchangeability judgments in predictive modeling and the role of data in statistical research, *Statistical Science* 2 (1987), 454-461.

- [10] D. Draper, J. Hodges, C. Mallows and D. Pregibon, Exchangeability and data analysis, *J. Royal Statist. Soc. Series A* 156 (1993), 9-28.
- [11] L.E. Dubins, On Lebesgue-like extensions of finitely additive measures, *Ann. Prob.* 2 (1974), 456-463.
- [12] L.E. Dubins, Towards characterizing the set of ergodic probabilities, pp. 61-74 in *Exchangeability in Probability and Statistics*, G. Koch and F. Spizzichino eds., North Holland, Amsterdam, 1982.
- [13] L.G. Epstein and M. Schneider, IID: independently and indistinguishably distributed, *J. Econ. Theory* 113 (2003), 32-50.
- [14] L.G. Epstein and M. Schneider, Learning under ambiguity, *Rev. Ec. Studies* 74 (2007), 1275-1303.
- [15] L.G. Epstein and T. Wang, 'Beliefs about beliefs' without probabilities, *Econometrica* 64 (1995), 1343-1373.
- [16] 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.
- [17] P. Ghirardato, On independence for non-additive measure, with a Fubini theorem, *J. Econ. Theory* 73 (1997), 261-291.
- [18] I. Gilboa and D. Schmeidler, Maxmin expected utility with non-unique priors, *J. Math. Econ.* 18 (1989), 141-153.
- [19] P. Hall and C.C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, 1980.
- [20] E. Hendon, H.J. Jacobson, B. Sloth and T. Traanaes, The product of capacities and belief functions, *Math. Soc. Sc.* 32 (1996), 95-108.
- [21] E. Hewitt and L.J. Savage, Symmetric measures on Cartesian products, *Trans. Amer. Math. Soc.* 80 (1955), 470-501.
- [22] O. Kallenberg, *Foundations of Modern Probability*, Springer, 1997.
- [23] O. Kallenberg, *Probabilistic Symmetries and Invariance Principles*, Springer, 2005.
- [24] J.F.C. Kingman, Uses of exchangeability, *Ann. Probab.* 6 (1978), 183-197.
- [25] D.M. Kreps, *Notes on the Theory of Choice*, Westview, 1988.
- [26] F. Maccheroni and M. Marinacci, A strong law of large numbers for capacities, *Ann. Prob.* 33 (2005), 1171-1178.

- [27] R.R. Phelps, *Convex Functions, Monotone Operators and Differentiability*, Springer Verlag, Berlin, 1989.
- [28] F. Philippe, G. Debs and J.Y. Jaffray, Decision making with monotone lower probabilities of infinite order, *Math. Oper. Res.* 24 (1999), 767-784.
- [29] E. Regazzini, Observability and probabilistic limit theorems, in *Proceedings of the International Conference "The Notion of Event in Probabilistic Epistemology,"* Trieste, 27-29 Maggio 1996.
- [30] M.J. Schervish, *Theory of Statistics*, Springer, 1995.
- [31] D. Schmeidler, Subjective probability and expected utility without additivity, *Econometrica* 57 (1989), 571-587.
- [32] F. Terkelsen, Some minmax theorems, *Math. Scand.* 31 (1972), 405-413.
- [33] P. Walley, *Statistical Reasoning with Imprecise Probabilities*, Chapman and Hall, 1991.
- [34] P. Walley and T.L. Fine, Towards a frequentist theory of upper and lower probability, *Ann. Statist.* 10 (1982), 741-761.