A PARSIMONIOUS THEORY OF SUBJECTIVE PROBABILITY

CHEW SOO HONG*, JACOB S. SAGI† AND WANG WENQIAN‡

FEBRUARY 6, 2021

We axiomatize subjective probabilities on finite domains without requiring richness in the outcome space or restrictions on risk preference using Event Exchangeability (Chew and Sagi, 2006), which has been implicit in the prior literature (Savage, 1954; Machina and Schmeidler, 1992; Grant, 1995). In three successively stronger theorems, we characterize a probability representation of the exchangeability relation, followed by characterizing a unique subjective probability, and finally endowing this subjective probability with the property of reduction consistency—acts inducing the same lottery are indifferent. This subjective probability can serve as foundation to derive expected utility and rank linear utility by imposing the sure-thing principle and its comonotonic variant. Moreover, our finite-domain setting reveals a novel possibility of state dependence arising from our subjective probability not being reduction consistent. Our axiomatic treatment can be further adapted to smaller collections of events and deliver small worlds probabilistic sophistication.

KEYWORDS: Subjective Probability, Finite State Spaces, Exchangeability, Probabilistic Sophistication, State Dependence, Source Preference.

1. INTRODUCTION

At the heart of decision making under uncertainty lies the concept of subjective probability. This has been an active direction of research since Ramsey (1926), de Finetti (1937), and Savage (1954) who derive subjective probability in conjunction with an underlying expected utility preference structure. The proliferation in the 1980s[1] of non-expected utility models has led to the question of how to disentangle the decision maker's belief represented by subjective probability from the underlying risk preference. Such a separation is first attempted by Machina and Schmeidler (1992). They characterize a "probabilistically sophisticated non-expected utility maximizer" whose preference over acts is represented by a utility functional over the corresponding act-induced lotteries through her subjective probability, requiring only the utility func-

*Southwestern University of Finance and Economics and National University of Singapore (chew.soohong@gmail.com)

†Kenan-Flagler Business School, University of North Carolina at Chapel Hill (sagi@unc.edu)

‡Society Hub, Hong Kong University of Science and Techonology (Guangzhou) (wqwang@ust.hk)

[1]See Quiggin (1982), Green and Jullien (1988), and Segal (1989) of non-expected utility models for decision making under risk for axiomatizations involving rank dependence and Chew (1983), Dekel (1986), and Chew (1989) for models which rely on the betweenness property.

tional to be continuous and monotone in the sense of stochastic dominance. This is weakened subsequently in Grant (1995) which partially relaxes monotonicity. A decade later, Chew and Sagi (2006) demonstrate how monotonicity and continuity can both be dispensed with by employing the notion of event exchangeability—two disjoint events are exchangeable if the decision maker is indifferent to exchanging their assigned outcomes regardless of how outcomes are assigned to other events. Event exchangeability provides a basic criterion for *probability-based choice* behavior, i.e., exchangeable events have the same subjective probability.

Significantly, research in the aforementioned literature tends to rely on a source of uncertainty that is infinitely divisible. At the same time, the development of subjective probability in finite state spaces has not evolved towards satisfactorily decoupling subjective probability from risk preference. Existing work typically requires some richness in the outcome space and accompanying restrictions on the underlying risk preference (Nakamura, 1990; Gul, 1992; Chew and Karni, 1994). Our primary contribution, adopting the exchangeability-based approach, is to provide a parsimonious framework for identifying subjective probability in finite state spaces with minimal requirements on both the outcome space and risk preference.

There is no dearth of problems involving decision making on a finite uncertainty domain without a rich outcome space. Which candidate will be elected in a primary? How many job offers will I receive in the next month? Will s(he) call tonight? Notice that the identification of the subjective probability in the aforementioned works typically requires an infinite enrichment with either irrelevant outcomes or unrelated states. In this regard, the following observation of Davidson and Suppes (1954) about the works of Ramsey (1926), de Finetti (1937), and Savage (1954) seems pertinent also for the other models of decision making discussed in the preceding paragraphs:

> While these theories may be satisfactory for normative purposes, from an empirical point of view they share the following disadvantage: for verification, and therefore for the derivation of measures, they require an infinite number of choices, yet no one can ever compare an infinite list of alternatives.

Consider Ellsberg's (1961) two-urn problem, which serves as a running example throughout this paper. This problem demonstrates a pattern of decision making that is not compatible with probability-based behavior; accordingly, we first illustrate how one of the characteristic properties of such behavior fails. Urn 1 contains 50 red and 50 black balls, while Urn 2 contains 100 red and black balls in total with unknown proportions. The state space is given by drawing one ball from each urn: $S = \{rr, rb, br, bb\}$, where the letter indicates the color of the ball (Red or Black) and its position the urn. For instance, rb is the state in which a red ball is drawn from Urn 1 and a black ball is drawn from Urn 2. The event $R_1 = \{rr, rb\}$ corresponds to drawing a red ball from the first urn; $B_1 = \{br, bb\}$ to drawing a black ball from the first urn; $R_2 = \{rr, br\}$ to drawing a red ball from the second urn; and $B_2 = \{rb, bb\}$ to drawing a black ball from the second urn. The commonly observed choice pattern of indifference between betting on R_1 and B_1 suggests that the two are exchangeable. Likewise, R_2 can be viewed as exchangeable with B_2. This can be depicted using a payoff matrix with outcomes x and y in which the two row events are exchangeable, as are the two column events:

$$\begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & x & x \\ B_1 & y & y \end{array} \sim \begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & y & y \\ B_1 & x & x \end{array} \qquad \begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & x & y \\ B_1 & x & y \end{array} \sim \begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & y & x \\ B_1 & y & x \end{array}$$

Consider the two families of exchangeable events $\mathcal{F}_1 = \{R_1, R_2\}$ and $\mathcal{F}_2 = \{B_1, B_2\}$. The collection of surplus state(s) in $\mathcal{F}_1$ relative to $\mathcal{F}_2$, i.e., state(s) whose number of appearances in the first family exceeds the number in the second family, consists only of rr, which appears twice more in $\mathcal{F}_1$. Likewise, the surplus states in $\mathcal{F}_2$ relative to $\mathcal{F}_1$ consist of bb, which appears twice more in $\mathcal{F}_2$. For any probability p such that

$$\begin{cases} p(R_1) = p(rr) + p(rb) = p(br) + p(bb) = p(B_1) \\ p(R_2) = p(rr) + p(br) = p(rb) + p(bb) = p(B_2) \end{cases},$$

it must be the case that $p(rr) = p(bb)$.

The prevailing choice pattern exhibited in experiments is that individuals prefer a bet of $x > 0$ on R_1 to a bet on B_2, as depicted below.

$$\begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & x & x \\ B_1 & 0 & 0 \end{array} \succ \begin{array}{c|cc} & R_2 & B_2 \\ \hline R_1 & 0 & x \\ B_1 & 0 & x \end{array}$$

In the above strict preference ordering, the two bets differ from each other by exchanging their outcomes on rr and bb. This is at odds with probability-based choice due to a violation of rr and bb being exchangeable with each other.

Our axiomatic treatment relies heavily on *exchangeable families*: same-sized (finitely indexed) families of events for which any pair of events with the same index are exchangeable, e.g., $\mathcal{F}_1$ and $\mathcal{F}_2$ in Ellsberg's two-urn example. The use of exchangeable families admits greater scope for investigating exchangeability in a finite setting, because a family of events can contain many instances of a particular state while each member event is limited to containing only one instance.

To characterize an exchangeability relation having a probability representation, the first condition we impose is Strong Event Non-satiation, which strengthens the corresponding axiom in Chew and Sagi (2006). Strong Event Non-satiation states that for any pair of exchangeable families, if there are no surplus states in the first family relative to the second, then any surplus states in the second family must be payoff irrelevant. The second axiom, termed Cancellation, is new. For any pair of exchangeable families, suppose that every surplus state in the first family relative to the second family appears k times more in the first family, and every surplus state in the second family relative to the first family also appears k times more in the second family. Cancellation requires these two collections of surplus states to be exchangeable, e.g., $\{rr\}$ and $\{bb\}$ to be exchangeable in Ellsberg's two-urn example. It is apparent that both Strong Event Non-satiation and Cancellation hold whenever the exchangeability relation has a probability representation. Far more nuanced and challenging to prove is the converse, a message delivered by Theorem 1 in this paper.

To arrive at an exchangeability-based likelihood comparison among events, Chew and Sagi (2006) consider an event to be at least as likely as another event whenever the former event contains a subevent that is exchangeable with the latter. In the presence of atoms, this definition is not generally applicable, because an atom can be strictly more likely than some other event, yet contain no subevents save for the empty set. To encompass likelihood comparisons between atoms, we can apply the idea behind the Cancellation axiom. For a pair of exchangeable families, let the first family be augmented with some extra events. Now, suppose that every surplus state in the first augmented family relative to the second family appears k times more, and every surplus state in the second family relative to the first augmented family also appears k times more. Relying on the idea of Cancellation, the collection of surplus states from the first augmented family would be more likely than the collection of surplus states from the second family. In our Theorem 2, a unique subjective probability follows from adding completeness of the extended comparative likelihood relation to Strong Event Non-satiation and Cancellation, encompassing all exchangeability relation from which a subjective probability can be identified on finite domains.

Our basic result in Theorem 2 sets up a parsimonious framework to axiomatize, successively, standard risk preferences, probabilistic sophistication and source preference. The existence of subjective probability leaves open the possibility that it could help simplify preference over acts to preference over their corresponding act-induced lotteries as in the case of decision making under risk. Should this be the case, we say that the decision maker exhibits *lottery-based choice* and that her subjective probability is reduction consistent. In the framework of Theorem 2 augmented by continuity and monotonicity, the sure-thing principle (STP) (Savage, 1954) and its comonotonic counterpart (Chew and Wakker, 1996) deliver lottery-based choice from the corresponding expected utility and rank linear utility (Green and Jullien, 1988; Segal, 1989) representations. It is worth noting that rank linear utility subsumes rank dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1992) as special cases.

Relying on the structural assumption of atomlessness, Savage (1954), Machina and Schmeidler (1992), and Grant (1995) all make use of a step showing the subjective probability to be the unique representation of an underlying exchangeability relation in their demonstration of reduction consistency. In this regard, our atomic setting reveals a novel possibility in which subjective probability may not be reduction consistent, pointing to the presence of state dependence. To arrive at lottery-based choice, we impose an additional axiom called Exchangeability Independence. Suppose the decision maker is indifferent between two acts inducing the same lottery under the uniquely identified subjective probability in Theorem 2. Exchangeability Independence requires indifference to be unaffected when a common outcome in both acts assigned to a pair of equally likely events is replaced with a new common outcome. From Exchangeability Independence, it emerges that all acts are treated as lotteries (Theorem 4).

We further adapt the reduction consistent subjective probability in Theorem 4 to model source preference in terms a "small world" in the decision theory literature (see, e.g., Savage 1954; Chew and Sagi 2008; Abdellaoui et al. 2011). Individuals who violate Cancellation in Ellsberg's two-urn problem exhibit a form of small worlds probability-based choice. The observation that R_i and B_i ($i = 1, 2$) are exchangeable pins down each of their subjective probabilities to 0.5 in both urns. In other words, the decision maker exhibits probability-based choice when making decisions in the finite space corresponding to each of the two urns. Probability-based choice fails only when the state space is enriched to incorporate both urns concurrently. This example illustrates the usefulness of our approach in the absence of global probability-based choice and points to an account of ambiguity aversion in terms of comparative risk aversion across the two sources of uncertainty.

The next section lays the groundwork for our exchangeability-based approach. Section 3 provides a characterization of exchangeability relation possessing probability representation with Strong Event Non-satiation and Cancellation on a finite domain. Section 4 characterizes a unique representing probability as the subjective probability with a Completeness axiom on the extended comparative likelihood based on exchangeability relation. Section 5 offers characterizations of standard models of decision making under risk within our framework. This is followed by a characterization in Section 6 of lottery-based choice via the addition of a new Exchangeability Independence axiom. In Section 7, we adapt the result of Section 6 to develop a result, which parallels Chew and Sagi (2008), concerning lottery-based choice on homogeneous domains of events in order to capture the idea of source preference. Finally, Section 8 summarizes our findings. Proofs are relegated to the Appendix.

2. PRELIMINARIES

Let the state space S be a finite set, X an arbitrary outcome space, and $\mathcal{F}$ the collection of all acts, i.e., functions from S to X. If E is a subset of S, fEg is the act that pays $f(s)$ in $s \in E$ and $g(s)$ in $s \in E^c$. Likewise, for two disjoint subsets E and F of S, the act $fEgFh$ pays $f(s)$ in $s \in E$, $g(s)$ in $s \in F$ and $h(s)$ in $s \in E^c \cap F^c$. We identify $x \in X$ with the constant act that pays x in every state of the world.

The decision maker's preference for acts is represented by a complete and transitive preference relation $\succcurlyeq$ over $\mathcal{F}$, with the corresponding indifference relation $\sim$ and strict preference $\succ$ being its symmetric and asymmetric parts, respectively. In addition, we require that preference not be degenerate, i.e., there exist $x, y \in X$ such that $x \succ y$. We say that an event E is *null* if $fEh \sim gEh$ for any $f, g, h \in \mathcal{F}$. Non-degeneracy implies that the whole state space S is not null.

We say that $p : S \to [0,1]$ is a *probability* on S if $\sum_{s \in S} p(s) = 1$. For a subset E of S, define $p(E) = \sum_{s \in E} p(s)$. The collection of *lotteries* with outcomes in X (probability distributions on X with finite support) is denoted by $\Delta(X)$. For a given probability p, an act f in $\mathcal{F}$ *induces* a lottery in $\Delta(X)$ with $\ell_f = \{(x, p_x) : x \in f(S) \text{ and } p_x > 0\}$ where $p_x = p(f^{-1}(x))$.

We use the event exchangeability relation in Chew and Sagi (2006), stated below, as a basis for equal likelihood.

DEFINITION 1: Two disjoint events E and F are said to be *exchangeable*, denoted as $E \approx F$, whenever $xEyFf \sim xFyEf$ for any $x, y \in X$ and any $f \in \mathcal{F}$.

For our main purpose of identifying subjective probability from the decision maker's preference, $\approx$ serves the purpose of the symmetric part of her yet-to-be-known qualitative probability (de Finetti, 1937). The following properties of $\approx$, summarized in the proposition below, are inherited from our assumptions over $\succcurlyeq$ and used throughout the paper.

PROPOSITION 1: *Suppose that $\succcurlyeq$ is a complete, transitive, and non-degenerate preference relation. Then,*

(i) $\varnothing \approx \varnothing$.
(ii) It is not the case that $\varnothing \approx S$.
(iii) $E \approx F$ *if and only if* $F \approx E$.
(iv) For any pairwise disjoint events E, F, E', F' *such that* $E \approx F$ *and* $E' \approx F'$*, we have* $E \cup E' \approx F \cup F'$.

We first seek probabilities that are compatible with $\approx$.

DEFINITION 2: We say that p is a probability *representing* $\approx$ if for any pair of disjoint events E and F, $E \approx F$ if and only if $p(E) = p(F)$.

Definition 2 describes a form of probability-based choice in which the decision maker is willing to exchange outcomes assigned to equally likely events according to p. Recall that Ellsberg's two-urn problem is not compatible with probability-based choice as discussed in the Introduction. At the other extreme, consider a setting in which no two non-null events are exchangeable. In this case, it seems vacuous to represent $\approx$

because it can be represented by any probability that assigns zero to null events, but for which no two non-null events have the same probability. It is evident that the value of a probability representing $\approx$ arises from having adequate pairs of events for which exchangeability holds. Should the exchangeability relation be sufficiently rich to pin down a unique probability representation of $\approx$, it would be the natural candidate for the decision maker's subjective probability.

DEFINITION 3: We say that p is a *subjective probability* if it is the unique probability representing $\approx$.

Our goal is to flesh out the necessary and sufficient conditions for the existence of a probability representation of $\approx$, and then to investigate its uniqueness, and then apply it to characterize expected utility along with rank linear utility, probabilistic sophistication, and source preference.

3. PROBABILITY REPRESENTATION OF EVENT EXCHANGEABILITY

An advantage of working with an infinitely divisible state space is that, under appropriate conditions on $\succcurlyeq$, one can add or subtract events (using the set union and set difference operations) at will until exchangeability is achieved. To obtain an analogous degree of flexibility in a finite state space, we work with *families of events* that contain a finite number of events. The families $\{E_1, \ldots, E_m\}$, $\{E_1, \ldots, E_m, \underbrace{e, \ldots, e}_{k}\}$ and $\{E_1, \ldots, E_m, e_1, \ldots, e_l\}$ will be abbreviated as $\{E_i\}$, $\{E_i, e^{(k)}\}$ and $\{E_i, e_j\}$ when it is unlikely to cause confusion. Previously, the axioms in Chew and Sagi (2006) are stated in terms of set unions and set differences. Our strategy is to develop similarly phrased axioms, applied to families of events, which can contain multiple copies of the same event and accommodate addition and subtraction at will.

We begin with an extension of set difference to an operation between families of sets. The *family difference* between $\{E_i\}$ and $\{F_j\}$, denoted as $\{E_i\} \setminus\!\!\setminus \{F_j\}$, consists of states whose occurrence in $\{E_i\}$ exceeds that in $\{F_j\}$. Formally, when $\{E_i\}$ and $\{F_j\}$ are of the equal size m, it is defined by

$$\{E_i\} \setminus\!\!\setminus \{F_j\} = \{s \in S : \forall \sigma, s \in \bigcup_{i=1}^{m} E_i \backslash F_{\sigma(i)}\}$$

where σ is a permutation of $\{1, \ldots, m\}$. When two families of events are of unequal size, we can augment the smaller family with some repetitions of $\varnothing$ to arrive at two families of equal size and then apply the previous definition. Note that when each family consists of a single event, family difference reduces to the set difference operation.

We say that $\{E_i\} \setminus\!\!\setminus \{F_j\} = e$ is *of order k* if $\{E_i\} \setminus\!\!\setminus \{F_j, e^{(k-1)}\} = e$ and $\{E_i\} \setminus\!\!\setminus \{F_j, e^{(k)}\} = \varnothing$. Applying this definition to Ellsberg's two-urn problem discussed in the Introduction yields that the two singleton events, $\{rr\} = \{R_1, R_2\} \setminus\!\!\setminus \{B_1, B_2\}$ and $\{bb\} = \{B_1, B_2\} \setminus\!\!\setminus \{R_1, R_2\}$, are both of order 2.

The following proposition draws on two characteristic properties of a probability—non-negativity and additivity.

PROPOSITION 2: *Let p be a probability over S, and $\{E_i\}_{i=1}^m$ and $\{F_i\}_{i=1}^m$ be two families of events.*

(i) Suppose that $p(E_i) \geq p(F_i)$ for any $i = 1, \ldots, m$ and $\{E_i\} \setminus\!\!\setminus \{F_i\} = \varnothing$, then $p(E_i) = p(F_i)$ and $p(\{F_i\} \setminus\!\!\setminus \{E_i\}) = 0$.

(ii) Suppose that $p(E_i) \geq p(F_i)$ for any $i = 1, \ldots, m$ and $\{E_i\} \setminus\!\!\setminus \{F_i\}$ and $\{F_i\} \setminus\!\!\setminus \{E_i\}$ are of the same order, then $p(\{E_i\} \setminus\!\!\setminus \{F_i\}) \geq p(\{F_i\} \setminus\!\!\setminus \{E_i\})$. In addition, if $p(E_i) > p(F_i)$ for some i, then $p(\{E_i\} \setminus\!\!\setminus \{F_i\}) > p(\{F_i\} \setminus\!\!\setminus \{E_i\})$.

We define two families $\{E_i\}_{i=1}^m$ and $\{F_i\}_{i=1}^m$ comprising pairwise exchangeable events, i.e., $E_i \approx F_i$ ($1 \leq i \leq m$), to be *exchangeable families*, denoted by $\{E_i\} \approx \{F_i\}$. As we may surmise from Proposition 2, the two axioms – Strong Event Non-satiation and Cancellation – introduced below, each stated in terms of event families, are necessary properties of a probability representation of $\approx$.

AXIOM E* (Strong Event Non-satiation): For any $\{E_i\} \approx \{F_i\}$, $\{E_i\} \setminus\!\!\setminus \{F_i\} = \varnothing$ implies that $\{F_i\} \setminus\!\!\setminus \{E_i\}$ is null.

If $\{E_i\} \setminus\!\!\setminus \{F_i\} = \varnothing$, then there are no surplus states in $\{E_i\}$ relative to $\{F_i\}$. If $\{E_i\} \approx \{F_i\}$, then Axiom E* requires surplus states in $\{F_i\}$ relative to $\{E_i\}$ to be null. This mimics the thinking behind Event Non-satiation from Chew and Sagi (2006), which is restated below.

AXIOM E (Event Non-satiation): For any pairwise disjoint E, F, e, if $E \approx F$ and e is non-null, then no subset of E is exchangeable with $F \cup e$.

To see that Strong Event Non-satiation implies Event Non-satiation, consider the case where the latter fails, i.e., when $E \approx F$ and $e' \approx e \cup E$ for some $e' \subseteq F$ where E, F, e are pairwise disjoint and e is non-null. Let $\{E_i\}$ be the family $\{e', E\}$ and $\{F_i\}$ be the family $\{e \cup E, F\}$. Then Strong Event Non-satiation also fails since $\{E_i\} \approx \{F_i\}$, $\{E_i\} \setminus\!\!\setminus \{F_i\} = \varnothing$ and $e \subseteq \{F_i\} \setminus\!\!\setminus \{E_i\}$ is not null.

AXIOM CN (Cancellation): For any $\{E_i\} \approx \{F_i\}$, if $\{E_i\} \setminus\!\!\setminus \{F_i\}$ and $\{F_i\} \setminus\!\!\setminus \{E_i\}$ are of the same order, then $\{E_i\} \setminus\!\!\setminus \{F_i\} \approx \{F_i\} \setminus\!\!\setminus \{E_i\}$.

Suppose that the differences between the two families, $\{E_i\}$ and $\{F_i\}$, are of the same order, say k. Then, the surplus states in $\{E_i\}$ relative to $\{F_i\}$ correspond to k copies of the event $\{E_i\} \setminus\!\!\setminus \{F_i\}$, while the surplus states in $\{F_i\}$ relative to $\{E_i\}$ correspond to k copies of $\{F_i\} \setminus\!\!\setminus \{E_i\}$. Cancellation requires these "residual" events to be exchangeable whenever they are generated by exchangeable families. As discussed in the Introduction, the observed choice in Ellsberg's two-urn problem typically violates Cancellation. Specifically, the residual events $\{rr\}$ and $\{bb\}$ are not exchangeable despite their coming from family differences of the same order (i.e., two) from the exchangeable families $\{R_1, R_2\}$ and $\{B_1, B_2\}$.

Our first main result is stated below.

THEOREM 1: *There exists a probability representing $\approx$ if and only if Cancellation and Strong Event Non-satiation hold.*

The non-uniqueness of probability representation points to the possibility of adopting additional criteria for selecting a probability representation or further shrinking its possible range, evoking criteria such as monotonicity or maxmin from the multiple-prior literature originating from Gilboa and Schmeidler (1989). Although Cancellation and Strong Event Non-satiation do not pin down the decision maker's subjective probability, they nonetheless serve to constrain subjective likelihood and provide a way of testing probability-based choice behavior.

4. CHARACTERIZING SUBJECTIVE PROBABILITY

Chew and Sagi (2006) offer the following definition of an exchangeability-based comparative likelihood $\succcurlyeq^c$.

DEFINITION 4: $E \succcurlyeq^c F$ if there exists $e \subseteq E \backslash F$ such that $e \approx F \backslash E$.

Completeness of $\succcurlyeq^c$ over finite state spaces will result in a uniform subjective probability. To see why it is the case, notice that for any two non-null singleton events, completeness of $\succcurlyeq^c$ requires them be exchangeable.[2] In a similar vein, to arrive at a Completeness axiom which accommodates a non-uniform subjective probability over finite state space, we begin with a generalization of Definition 4 using families of events.

DEFINITION 5: $E \succcurlyeq^* F$ if there exist $\{E_i\} \approx \{F_i\}$ and a family of events $\{e_j\}_{j=1}^{l}$ such that $E \backslash F = \{E_i, e_j\} \setminus\!\!\setminus \{F_i\}$ and $F \backslash E = \{F_i\} \setminus\!\!\setminus \{E_i, e_j\}$ are of the same order.

We refer to an event e_j in the above definition as an *excess event*. To see how Definition 5 generalizes Definition 4, let e be a subset of $E \backslash F$ that is exchangeable with $F \backslash E$ (so

[2] By assuming that $\succcurlyeq^c$ is complete, we only need Event Non-satiation rather than Strong Event Non-satiation for existence of subjective probability, and Cancellation is implied.

that $E \succcurlyeq^c F$ in Definition 4). Define $\{E_i\}$ to contain only e and $\{F_i\}$ to contain only $F\backslash E$. Then, $\{E_i\} \approx \{F_i\}$. Let e_1 be the (single) excess event $(E\backslash F)\backslash e$. Note that $\{e, e_1\} \setminus\setminus \{F\backslash E\} = E\backslash F$ is of order one. Note further that $\{F\backslash E\} \setminus\setminus \{e, e_1\} = F\backslash E$ is also of order one. Thus a ranking of events under $\succcurlyeq^c$ of Definition 4 implies the same ranking under $\succcurlyeq^* F$ of Definition 5.

The Cancellation Axiom and Definition 5 only apply when residual events from family differences are of the same order. To see that this is essential, consider a uniform state space $S = \{s_1, s_2, s_3\}$ (i.e., all atoms are exchangeable). Let $\{E_i\} = \{\{s_1\}, \{s_1\}\}$ and $\{F_i\} = \{\{s_2\}, \{s_3\}\}$. Clearly, $\{E_i\}$ and $\{F_i\}$ are exchangeable families. Letting $\varnothing$ be a single excess event, note that $\{s_1\} = \{E_i, \varnothing\} \setminus\setminus \{F_i\}$ is of order two, and that $\{s_2, s_3\} = \{F_i\} \setminus\setminus \{E_i, \varnothing\}$ is of order one. Thus, the Cancellation Axiom and Definition 5 do not apply. This is sensible because it should not be the case that $\{s_1\} \approx \{s_2, s_3\}$ or that $\{s_1\} \succcurlyeq^* \{s_2, s_3\}$.

The following example of subjective probability shows how Definition 5 can deliver a complete comparative likelihood when Definition 4 does not.

EXAMPLE 1: Consider an urn containing balls of four different colors: purple (pl), red (r), orange (o) and teal (t). States correspond to the color of a ball drawn, $S = \{pl, r, o, t\}$. Suppose that a decision maker's rankings of acts yield $\{pl, r\} \approx \{o\}$, $\{pl, o\} \approx \{t\}$ and $\{pl, t\} \approx \{r, o\}$. There is a unique probability representation of the decision maker's exchangeability relation, $p = \langle \frac{1}{10}, \frac{2}{10}, \frac{3}{10}, \frac{4}{10} \rangle$. It should be clear, however, that $\succcurlyeq^c$ cannot capture all of the likelihood relationships implied by the representing probability, p. For instance, $\succcurlyeq^c$ does not order pl and r. By contrast, to verify that $\{r\} \succcurlyeq^* \{pl\}$, let $\{E_i\} = \{\{r, o\}, \{t\}\}$ and $\{F_i\} = \{\{pl, t\}, \{pl, o\}\}$ be the exchangeable families in Definition 5, and $e_1 = \{r\}$ be the excess event. Observe that $\{E_i, e_1\} \setminus\setminus \{F_i\} = \{r\}$ is of order two, as is $\{F_i\} \setminus\setminus \{E_i, e_1\} = \{pl\}$. The ranking of any other two pairs of events via $\succcurlyeq^*$ can similarly be established and shown to be consistent with the ranking according to the representing probability, p.

The next result establishes the relationship between $\succcurlyeq^*$ and probability representations of $\approx$ whose existence is guaranteed under Theorem 1.

PROPOSITION 3: *Suppose that $\approx$ has a probability representation. Then, $E \succcurlyeq^* F$ if and only if $p(E) \geq p(F)$ for any probability p representing $\approx$.*

The important message delivered by Proposition 3 is that a pair of events for which the decision maker's likelihood comparison can be determined is one on which all probabilities representing $\approx$ agree. As with the exchangeability relation, $\succcurlyeq^*$ is, in general, an incomplete binary relation over events. This proposition also helps relate $\succcurlyeq^*$ to the literature on incomplete qualitative probability, denoted by $\succcurlyeq^l$ (Nehring, 2009; Alon and Lehrer, 2014). There are differences between $\succcurlyeq^*$ and $\succcurlyeq^l$ that are rooted in the nature of the corresponding representing probabilities. Typically, if E and F cannot be ranked by $\succsim^I$, then there are different representing probabilities p_1, p_2 and p_3 such that $p_1(E) < p_1(F)$, $p_2(E) > p_2(F)$, and $p_3(E) = p_3(F)$. By contrast, for any representation p of $\approx$ used to define $\succsim^*$, $p(E) = p(F)$ implies that $E \approx F$, so that E and F are necessarily in the symmetric part of $\succsim^*$.

Proposition 3 also establishes that a (unique) subjective probability implies the completeness of $\succsim^*$.

AXIOM C* (Extended Completeness): $\succsim^*$ is complete.

The converse to the above yields the second key result of this paper.

THEOREM 2: *There exists a subjective probability if and only if Cancellation, Strong Event Non-satiation, and Extended Completeness hold.*

Under the axioms of Cancellation, Strong Event Non-satiation, and Extended Completeness, the derived $\succsim^*$ is a qualitative probability in the sense of de Finetti (1937) whose representation on a finite algebra has been a longstanding topic in the literature (Kraft et al., 1959; Scott, 1964; Fishburn and Roberts, 1989). Theorem 2 embodies the most encompassing construction of a finite comparative likelihood relation which accommodates all exchangeability relations possessing a unique probability representation. It is nothworthy that our construction does not rely on the decision maker's preference over sure outcomes.[3]

5. CHARACTERIZING STANDARD MODELS

Building on Theorem 2, we now illustrate how lottery-based choice can arise from incorporating the STP and its comonotonic variant to obtain expected utility and rank linear utility (Green and Jullien, 1988; Segal, 1989), which contains rank dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1992). We further show that a corresponding weakening of the STP to indifference sets is not sufficient to deliver a finite-state axiomatization of betweenness utility (Chew, 1983, 1989; Dekel, 1986) with non-uniform probabilities. In the next section, we show directly how lottery-based choice can arise from adding a new Exchangeability Independence axiom to Theorem 2. There is no loss of continuity in skipping the present section.

Consistent with common assumptions used in the literature, we begin by imposing continuity and monotonicity for the case where the outcome set is the real line. We further assume that there are $N \geq 3$ non-null states.

[3]Note that the constructions of comparative likelihood in Savage (1954), based on P4, and Machina and Schmeidler (1992), based on P4*, both utilize strict preference between sure outcomes.

AXIOM CT: $\succcurlyeq$ is *continuous*, i.e., $\{f \in \mathcal{F} : f \succcurlyeq h\}$ and $\{g \in \mathcal{F} : h \succcurlyeq g\}$ are both closed for all $h \in \mathcal{F}$.

AXIOM M: $\succcurlyeq$ is *monotone*, i.e., $x > y$ implies that $xEf \succ yEf$ when E is not null.

Continuity and monotonicity give rise to an overall representation of the decision maker's preference ordering by a continuous and monotone utility function $U : \mathcal{F} \to \mathbb{R}$. We next consider two versions of the STP (the first of which is the standard statement).

AXIOM S (Sure-thing Principle): $fEh \succcurlyeq gEh$ implies that $fEh' \succcurlyeq gEh'$ for any event $E \subseteq S$ and any $f, g, h, h' \in \mathcal{F}$.

To relate our framework to the literature on rank dependent preference, we adopt a weakening of the STP to comonotonic acts as in Chew and Wakker (1996). We say that two acts f and g are *comonotonic* if there do not exist $s_i, s_j \in S$ such that $f(s_i) \geq f(s_j)$ and $g(s_i) < g(s_j)$.

AXIOM SC (Comonotonic Sure-thing Principle): $fEh \succcurlyeq gEh$ implies that $fEh' \succcurlyeq gEh'$ for any event $E \subseteq S$ and any pairwise comonotonic fEh, gEh, fEh', and gEh'.

In their unifying axiomatization of expected utility, rank linear utility, and betweenness utility for decision making under risk, Chew and Epstein (1989) observe that the independence axiom and its comonotonic and betweenness variants deliver additive separability of the utility function on the corresponding domains. This motivates the following theorem which builds on the result of Theorem 2.

THEOREM 3: *Under the hypotheses of Theorem 2 and supposing that there are at least three non-null states,*

(i) $\succcurlyeq$ *on* $\mathcal{F}$ *satisfies continuity, monotonicity, and the STP if and only if it has an EU representation.*[4]

(ii) $\succcurlyeq$ *on* $\mathcal{F}$ *satisfies continuity, monotonicity, and the Comonotonic STP if and only if it has a rank linear utility (RLU) representation.*[5]

Furthermore, the EU and RLU representations above are with respect to the subjective probability identified in Theorem 2.

[4]A function $U : \mathcal{F} \to \mathbb{R}$ is an expected utility (EU) representation of $\succcurlyeq$ if there exist a probability p and $v : \mathbb{R} \to \mathbb{R}$ that is continuous and strictly increasing, such that $U(f) = \sum_{i=1}^{N} p(s_i)v(f(s_i))$.

[5]Consider a Savage act f written in the form of $\{x_1, A_1; \ldots; x_M, A_M\}$ where $\{A_i\}$ is a partition of S and $f(A_k) = x_k$ with $x_i \geq x_j$ if $i < j$. We say $W : \mathbb{R} \times [0,1] \to \mathbb{R}$ is an *outcome-dependent probability transformation* if (i) $W(0, \cdot) = 0$ and $W(\cdot, 0) = 0$, (ii) $W(x, 1) \neq 0$ for all $x \neq 0$ and (iii) $V(\cdot, p, q) = W(\cdot, p) - W(\cdot, q)$ is strictly increasing if $p > q$. We then define $U : \mathcal{F} \to \mathbb{R}$ to be an RLU representation of $\succcurlyeq$ if there exist a probability p and an outcome-dependent probability transformation W that is continuous in the first argument, such that $U(f) = \sum_{i=1}^{M} V(x_i, p(\cup_{k=1}^{i} A_k), p(\cup_{k=1}^{i-1} A_k))$.

Notwithstanding the results above, it is tempting to apply Theorem 2 together with a weakening of the STP to indifference sets and attempt a finite-state axiomatization of betweenness utility. This however is not to be. Consider a decision maker whose continuous, differentiable, and strictly increasing certainty equivalent c of an act (x_1, x_2, x_3) on a state space with three states is implicitly defined by the following equation:

$$F(x_1 - c) + F(x_2 - c) + G(x_3 - c) = 0$$

where $F(x) = x/2$ and $G(x) = \begin{cases} x + x^2, & x \geq 0 \\ 1/2 - \sqrt{1/4 - x}, & x < 0 \end{cases}.$

Notice that this certainty equivalent function is additively separable and is therefore consistent with restricting the STP to indifference sets. It is easy to see that the decision maker considers the first two states equally likely and to check that $(x, x, y) \sim (y, y, x)$ for all $x, y \in \mathbb{R}$. Applying Theorem 2, we can identify a non-uniform subjective probability given by $(\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$. This example serves to show that characterizing betweenness preference in a finite state space setting will require more structure than might have been anticipated.

6. CHARACTERIZING LOTTERY-BASED CHOICE

Complementing the development in the preceding section, we show how we can characterize lottery-based choice without relying on the auxiliary axioms of continuity, monotonicity, and the STP or its comonotonic variant. We begin with a formal definition of the quintessential property of subjective probability delivering lottery-based choice.

DEFINITION 6: We say that a probability p is *reduction consistent* if $f \sim g$ whenever $\ell_f = \ell_g$ under p.

Arising from a reduction consistent subjective probability, lottery-based choice coincides with Grant's (1995) definition of probabilistic sophistication, which he distills from the behavior of a "probabilistic sophisticated non-expected utility maximizer" studied in Machina and Schmeidler (1992).[6] In both papers and also Savage (1954), an essential step in their characterizations of lottery-based choice is in demonstrating that the subjective probability p identified from the different systems of axioms satisfies Definition 3, i.e., p is the unique representation of an underlying exchangeability

[6]In Machina and Schmeidler (1992), a decision maker is a probabilistic sophisticated non-expected utility maximizer if there are a continuous and monotonic non-expected utility preference functional V and a probability measure p such that her preference over acts can be represented by V over their induced lotteries. It is clear that such a decision maker would be indifferent between two acts that induce the same lottery because they are indistinguishable in terms of V.

relation. Should p be atomless or uniformly atomic, it can be shown that it is necessarily reduction consistent. However, the validity of this procedure for finite domains is limited to the case of a small number of states according to the proposition below.

PROPOSITION 4: *If a state space has less than six states, a probability representing $\approx$ is necessarily reduction consistent.*

According to Proposition 4, probability-based behavior involving small numbers of states are inherently lottery based. The following example shows that this may not be the case when the state space has more than five states.

EXAMPLE 2: Consider a decision maker whose complete and transitive preference over acts on state space $S = \{s_1, ..., s_6\}$ admits an exchangeability relation that satisfies Strong Event Non-satiation, Cancellation, and Extended Completeness. Suppose that the decision maker's exchangeability relation contains the following instances: $\{s_3\} \approx \{s_4\}, \{s_5\} \approx \{s_1, s_2\}, \{s_6\} \approx \{s_1, s_4\}$, and $\{s_3, s_4\} \approx \{s_1, s_6\} \approx \{s_2, s_5\}$. This is sufficient to identify a unique subjective probability $p = \langle \frac{2}{24}, \frac{3}{24}, \frac{4}{24}, \frac{4}{24}, \frac{5}{24}, \frac{6}{24} \rangle$ on S. Consider the following two acts f and g, both inducing the same lottery $\frac{8}{24}\delta_x + \frac{7}{24}\delta_y + \frac{9}{24}\delta_z$:

	s_1	s_2	s_3	s_4	s_5	s_6
f	x	y	y	z	z	x
g	y	z	x	x	y	z

One cannot find a sequence of acts $h_0, h_1, \cdots, h_m$ such that $f = h_0$, $g = h_m$, and h_{i+1} differs from h_i by exchanging outcomes on a pair of exchangeable events. Thus, indifference between f and g cannot be established without an additional assumptions on $\succcurlyeq$.

We identify a property, called Exchangeability Independence, which when added to Theorem 2 delivers reduction consistency for the subjective probability. We say that f and g induce *exchangeable partitions* if $f^{-1}(x) \backslash g^{-1}(x) \approx g^{-1}(x) \backslash f^{-1}(x)$ for all $x \in X$.

AXIOM EI (Exchangeability Independence): For any acts $f, g \in \mathcal{F}$, and events $E, F \subseteq S$, if f and g induce exchangeable partitions, $f \sim g$, $f(E) = g(F)$, and $E \backslash F \approx F \backslash E$, then $xEf \sim xFg$ for any $x \in X$.

Consider the following acts f' and g' in the setting of Example 2.

	s_1	s_2	s_3	s_4	s_5	s_6
f'	x	y	y	x	x	x
g'	y	x	x	x	y	x

The indifference between f' and g' follows from observing that $\{s_1, s_5\} \approx \{s_2, s_3\}$. Exchangeability Independence rules out the possibility that replacing the common outcome x with z on exchangeable events $\{s_4, s_5\}$ and $\{s_2, s_6\}$, respectively, will lead to a strict preference between f and g.

We are now ready to present our final main result.

THEOREM 4: *There exists a reduction consistent subjective probability if and only if Strong Event Non-satiation, Cancellation, Extended Completeness, and Exchangeability Independence hold.*

It seems useful to relate Theorem 4 to the state dependence literature (see, e.g., Karni 1993). Consider the following behavioral property stated in terms of subjective probability and the exchangeability relation jointly: If a pair of events have the same subjective probability, then they are exchangeable with each other. This property, known as Event Independence in Grant (1995), is implied by our definition of subjective probability (Definition 3) being the unique representation of an underlying exchangeability relation.[7] This observation helps resolve a longstanding issue concerning the uniqueness of subjective probability in SEU in the presence of state dependent utility functions (see, e.g., Karni 2014), which is recently discussed in Chew and Wang (2020) and Karni (2020) for non-expected utility preferences.[8] Although Event Independence already conveys some flavor of state independence, it needs to work in tandem with Exchangeability Independence, as revealed by Example 2 and Theorem 4, to capture the overall picture of state independence more fully.

7. CHARACTERIZING SOURCE PREFERENCE

We next develop a result, which parallels Chew and Sagi (2008), concerning lottery-based choice on small worlds—domains of events capturing sources of uncertainty revealed in the decision maker's own preference. We say that a collection of events $\mathcal{A} \subseteq 2^S$ with $A = \bigcup \mathcal{A}$ is an *algebra* if $A \in \mathcal{A}$, $E \in \mathcal{A}$ implies that $A \backslash E \in \mathcal{A}$, and $E, F \in \mathcal{A}$ implies that $E \cap F \in \mathcal{A}$.[9]

[7] Event Independence is related to State Neutrality in Ok and Savochkin's (2020) study of when a decision maker fully yields to the power of suggestion through her preference over acts with suggested priors.

[8] As observed earlier, the characterizations of state-independent subjective probability p in the probabilistic sophistication literature rely on a step in their proofs which involve p representing an underlying exchangeability relation. This naturally differentiates it from an alternative notion of subjective probability q in the presence of state dependent utility functions, e.g., $\sum_{i=1}^n p(s_i)u(f(s_i)) = \sum_{i=1}^n q(s_i)u_{s_i}(f(s_i))$ with state-dependent utilities $u_{s_i}(\cdot)$ given by $[p(s_i)/q(s_i)]u(\cdot)$

[9] Chew and Sagi (2008) consider subjective probability on λ-systems. As a direct application of our previous results, we have made simplifying assumptions using algebras.

DEFINITION 7: $\mathcal{A} \subseteq 2^S$ is a *homogeneous collection* if it satisfies the following:
(i) $\mathcal{A}$ is an algebra.
(ii) $\approx$ satisfies Strong Event Non-satiation and Cancellation on $\mathcal{A}$.
(iii) $\succcurlyeq^*$ defined with events in $\mathcal{A}$ is complete on $\mathcal{A}$.[10]

To proceed, we re-introduce several definitions in the preceding analysis with minor modifications to adapt to the involvement of sources of uncertainty. A probability p on an algebra $\mathcal{A}$ with $A = \bigcup \mathcal{A}$ is a function from A to $[0,1]$ such that $p(A) = \sum_{s \in A} p(s) = 1$. An act f is *adapted* to $\mathcal{A}$ if $f^{-1}(x) \cap A \in \mathcal{A}$ for all $x \in X$ and induces a lottery $\ell_f^A = \{(x, p_x) : x \in f(A) \text{ and } p_x > 0\}$ where $p_x = p(f^{-1}(x) \cap A)$. Two acts f and g, which are both adapted to $\mathcal{A}$, induce exchangeable partitions $(f^{-1}(x) \cap A) \backslash (g^{-1}(x) \cap A) \approx (g^{-1}(x) \cap A) \backslash (f^{-1}(x) \cap A)$ for all $x \in X$.

DEFINITION 8: We say that a probability p on $\mathcal{A}$ is a *subjective probability* if it is the unique probability such that for all $E, F \in \mathcal{A}$, $p(E) \geq p(F)$ if and only if $E \succcurlyeq^* F$. Furthermore, we say that a subjective probability p on $\mathcal{A}$ is *reduction consistent* if $\ell_f^A = \ell_g^A$ implies that $fAh \sim gAh$ for all f, g adapted to $\mathcal{A}$ and all $h \in \mathcal{F}$.

Following Theorem 2, we can demonstrate the existence of subjective probability p on a homogeneous collection $\mathcal{A}$. To show that p is reduction consistent, we need a version of Exchangeability Independence for small worlds.

AXIOM EI* (Adapted Exchangeability Independence): For any acts $f, g \in \mathcal{F}$ that are adapted to $\mathcal{A}$, any act $h \in \mathcal{F}$, and events $E, F \in \mathcal{A}$, if f and g induce exchangeable partitions on $\mathcal{A}$, $fAh \sim gAh$, $f(E) = g(F)$, and $E \backslash F \approx F \backslash E$, then $xEf(A \backslash E)h' \sim xFg(A \backslash F)h'$ for any $x \in X$ and $h' \in \mathcal{F}$.

Our characterization of small worlds probabilistic sophistication is provided in the following corollary to Theorem 4.

COROLLARY: *Under Adapted Exchangeability Independence, if $\mathcal{A}$ is a homogeneous collection, then there exists a reduction consistent subjective probability on $\mathcal{A}$.*

In light of the corollary, Ellsberg's two-urn example provides two homogeneous collections—each generated by a single urn with an even-chance subjective probability. The commonly observed ambiguity averse behavior favoring bets on the first source (with known proportion) may be interpreted in terms of greater aversion to risks aris-

[10]In applying Definition 5, both the exchangeable families and the excess events are in $\mathcal{A}$.

ing from the second source (with unknown proportion). Such a source-based account applies naturally to the study of Chew et al. (2012).[11] They find that 40.4% of 325 subjects in Beijing favor a bet on whether the city's temperature is an odd or an even number paying RMB 11 rather than the corresponding bet on Tokyo's temperature paying RMB 13. As suggested in Fox and Tversky (1995), a preference for the familiar may underpin the equity home bias puzzle in international equity markets (Feldstein and Horioka, 1980; French and Poterba, 1991) which is replicated subsequently in the domestic U.S. equity market (Coval and Moskowitz, 2001; Huberman, 2001).

With this corollary, consider the following example, which injects the thinking behind Ellsberg's (1961) three-color problem into Example 1 through ambiguity in the teal-colored balls.

EXAMPLE 3: In an urn containing 150 colored balls, 15 of them are purple (pl), 30 are red (r), 45 are orange (o), and 60 are teal (t). Should a t ball be drawn, it may be closer to blue (b) or green (g) upon further scrutiny, but the proportion of b versus g is unknown. The overall state space is given by $\{pl, r, o, tb, tg\}$, where tb (tg) refers to the state of drawing t followed by finding out that it is closer to b (g). The event of drawing a teal ball t is given by $\{tb, tg\}$. When betting on $\{pl, r, o, t\}$, the decision maker exhibits the same pattern of exchangeability as in Example 1. Yet, when the betting involves tb or tg, she prefers to bet on t rather than $\{r, tb\}$ (or $\{r, tg\}$). At the same time, the decision maker is indifferent to exchanging outcomes associated with tb and tg. Consequently, tb and tg are exchangeable while r and tb (or tg) are not.

The choice behavior in the above example reveals that the decision maker does not exhibit lottery-based choice in the whole state space.[12] Here, we can identify two homogeneous collections in which the decision maker exhibits lottery-based choice individually. One of them is generated by $\{pl, r, o, t\}$ and the other by $\{tb, tg\}$. In line with the stylized behavior associated with Ellsberg's three-color problem, we anticipate a tendency to favor betting on r than tb (or tg) in our five-color example. As with Ellsberg's two-urn problem, this preference can be explained by the decision maker being less averse to risks arising from $\{pl, r, o, t\}$ than those from $\{tb, tg\}$.

[11]This is motivated by Fox and Tversky's (1995) finding of familiarity preference in which UC Berkeley undergraduate subjects assign on average higher willingness-to-pay for betting on San Francisco temperature being below 60 °F than betting on Istanbul temperature being above 60 °F, even though willingness-to-pay is higher for betting above than betting below for both cities.

[12]Let the two exchangeable families $\{E_i\}$ and $\{F_i\}$ be given by $\{\{pl, r\}, \{tb\}, \{r, o\}\}$ and $\{\{o\}, \{tg\}, \{pl, t\}\}$, respectively. Violation of the Cancellation axiom follows from the non-exchangeability between $r = \{E_i\} \setminus\!\!\setminus \{F_i\}$ and $tg = \{F_i\} \setminus\!\!\setminus \{E_i\}$ as they are both of order two.

8. SUMMARY

Our main contribution in this paper is to provide a parsimonious framework for identifying subjective probability in finite state spaces with minimal requirements on both the outcome space and risk preference. We offer successively stronger definitions of a probability representing an underlying event exchangeability relation $\approx$. We first characterize in Theorem 1 the exchangeability relation $\approx$ that have a probability representation using a strengthening of Event Non-satiation in Chew and Sagi (2006) together with a new axiom of Cancellation which is unmasked in our finite-state setting. We then identify, in Theorem 2, a unique probability representing $\approx$ as subjective probability by incorporating completeness of an extended comparative likelihood which captures all exchangeability relations on finite domains from which a subjective probability can be identified. In Theorem 3 which additionally assumes continuity and monotonicity, we use the axioms of STP and comonotonic STP to axiomatize expected utility and rank linear utility on finite domains in a Savagian setting, delivering *inter alia* a reduction consistent subjective probability.

Intriguingly, the finite-state subjective probability characterized in Theorem 2 can exhibit state dependence in harboring the possibility that acts inducing the same lottery may not be indifferent. This prompts us to develop the Exchangeability Independence axiom which when added to Theorem 2 delivers state-independent lottery-based choice in Theorem 4 through a reduction consistent subjective probability. Along with the Cancellation axiom, note that Exchangeability Independence exemplifies an additional axiom that is masked by the commonly imposed structural assumptions of atomlessness or uniformity of atoms on the subjective probability. As observed in the preceding paragraph, Theorem 2 also yields lottery-based choice through the resulting utility representations in Theorem 3. Finally, Theorem 4 is adapted to smaller families of homogeneous events, giving rise to small worlds probabilistic sophistication (Chew and Sagi, 2008) through which we can account for the phenomena of ambiguity aversion and familiarity preference.

REFERENCES

ABDELLAOUI, M., A. BAILLON, L. PLACIDO, AND P. P. WAKKER (2011): "The Rich Domain of Uncertainty : Source Functions and Their Experimental Implementation," *American Economic Review*, 101, 695–723.

ALON, S. AND E. LEHRER (2014): "Subjective multi-prior probability: A representation of a partial likelihood relation," *Journal of Economic Theory*, 151, 476–492.

CHEW, S. H. (1983): "A generalization of the quasilinear mean with applications to the measurement of income inequality and decision theory resolving the Allais paradox," *Econometrica*, 51, 1065–1092.

——— (1989): "Axiomatic utility theories with the betweenness property," *Annals of Operations Research*, 19, 273–298.

CHEW, S. H., R. P. EBSTEIN, AND S. ZHONG (2012): "Ambiguity aversion and familiarity bias: Evidence from behavioral and gene association studies," *Journal of Risk and Uncertainty*, 44, 1–18.

CHEW, S. H. AND L. G. EPSTEIN (1989): "A Unifying Approach to Axiomatic Non- expected Utility Theories," *Journal of Economic Theory*, 49, 207–240.
CHEW, S. H. AND E. KARNI (1994): "Choquet expected utility with a finite state space: Commutativity and act-independence," *Journal of Economic Theory*, 62, 469–479.
CHEW, S. H. AND J. S. SAGI (2006): "Event exchangeability: Probabilistic sophistication without continuity or monotonicity," *Econometrica*, 74, 771–786.
——— (2008): "Small worlds: Modeling attitudes toward sources of uncertainty," *Journal of Economic Theory*, 139, 1–24.
CHEW, S. H. AND P. WAKKER (1996): "The comonotonic sure-thing principle," *Journal of Risk and Uncertainty*, 12, 5–27.
CHEW, S. H. AND W. WANG (2020): "On the robustness of indeterminacy in subjective probability," *Economics Letters*, 188, p.108939.
COVAL, J. D. AND T. J. MOSKOWITZ (2001): "The Geography of Investment: Informed Trading and Asset Prices," *Journal of Political Economy*, 109, 811–841.
DAVIDSON, D. AND P. SUPPES (1954): "A Finitistic Axiomatization of Subjective Probability and Utility," *Econometrica*, 24, 264–275.
DE FINETTI, B. (1937): "Foresight: Its Logical Laws, Its Subjective Sources," *Annales de l'Institut Henri Poincaré*, 6, 1–68; as translated (1964) in *Studies in Subjective Probability*, ed. by H. E. Kyburg and H. Smokler. New York: Wiley.
DEBREU, G. (1960): "Topological methods in cardinal utility theory," in *Mathematical methods in the social sciences*, ed. by K. J. Arrow, S. Karlin, and P. Suppes. Stanford, California: Stanford University Press.
DEKEL, E. (1986): "An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom," *Journal of Economic Theory*, 40, 304–318.
ELLSBERG, D. (1961): "Risk, Ambiguity, and the Savage Axioms," *The Quarterly Journal of Economics*, 75, 643–669.
FELDSTEIN, M. AND C. HORIOKA (1980): "Domestic Saving and International Capital Flows," *The Economic Journal*, 90, 314–329.
FISHBURN, P. C. AND F. S. ROBERTS (1989): "Axioms for Unique Subjective on Finite Sets," *Journal of Mathematical Psychology*, 33, 117–130.
FOX, C. R. AND A. TVERSKY (1995): "Ambiguity Aversion and Comparative Ignorance," *The Quarterly Journal of Economics*, 110, 585–603.
FRENCH, K. R. AND J. M. POTERBA (1991): "Investor Diversification And International Equity Markets," *American Economic Review*, 81(2), 222–226.
GILBOA, I. AND D. SCHMEIDLER (1989): "Maxmin expected utility with non-unique prior," *Journal of Mathematical Economics*, 18, 141–153.
GRANT, S. (1995): "Subjective Probability Without Monotonicity: or How Machina's Mom May Also be Probabilistically Sophisticated," *Econometrica*, 63, 159–189.
GREEN, J. R. AND B. JULLIEN (1988): "Ordinal independence in nonlinear utility theory," *Journal of Risk and Uncertainty*, 1, 355–387.
GUL, F. (1992): "Savage's Theorem with a Finite Number of States," *Journal of Economic Theory*, 57, 99–100.
HUBERMAN, G. (2001): "Familiarity Breeds Investment," *The Review of Financial Studies*, 14, 659–680.
KARNI, E. (1993): "A Definition of Subjective Probabilities with State-Dependent Preferences," *Econometrica*, 61, 187–198.
KARNI, E. (2014): "Axiomatic Foundations of Expected Utility and Subjective Probability," in *Handbook of the Economics of Risk and Uncertainty*, ed. by M. J. Machina and W. Kip Viscusi. Oxford: North Holland.
KARNI, E. (2020): "On the indeterminacy of the representation of beliefs by probabilities," *Economics Letters*, 196, p.109535.

KRAFT, C. H., J. W. PRATT, AND A. SEIDENBERG (1959): "Intuitive Probability on Finite Sets," *Annals of Mathematical Statistics*, 30, 408–419.

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

NAKAMURA, Y. (1990): "Subjective expected utility with non-additive probabilities on finite state spaces," *Journal of Economic Theory*, 51, 346–366.

NEHRING, K. (2009): "Imprecise probabilistic beliefs as a context for decision-making under ambiguity," *Journal of Economic Theory*, 144, 1054–1091.

Ok, E. A. and Savochkin, A. (2020). Believing in Forecasts, Uncertainty, and Rational Expectations. *working paper.*

QUIGGIN, J. (1982): "A Theory of Anticipated Utility," *Journal of Economic Behavior and Organization*, 3, 323–343.

RAMSEY, F. P. (1926): "Truth and Probability," in *The Foundations of Mathematics and Other Logical Essays*, ed. by R. B. Braithwaite, London: Routledge, chap. 7, 156—-198.

SAVAGE, L. J. (1954): *The Foundations of Statistics*, New York: Weily.

SCOTT, D. (1964): "Measurement structures and linear inequalities," *Journal of Mathematical Psychology*, 1, 233–247.

SEGAL, U. (1989): "Anticipated Utility: A Measure Representation Approach," *Annals of Operations Research*, 19, 359–373.

TVERSKY, A. AND D. KAHNEMAN (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty," *Journal of Risk and Uncertainty*, 5, 297–323.

PROOFS

We first prove propositions 1 and 2. We then introduce a vector representation of the exchangeability relation to facilitate the development of a proof of our last main result—Theorem 4—via a number of intermediate results including Theorem 1, Theorem 2, and Proposition 3. The proofs of Theorem 3 and Proposition 4 are given at the end. The proof of the corollary is omitted as a direct application of Theorem 4 to an algebra of events.

PROOF OF PROPOSITION 1: Property (i) follows from reflexivity of $\succcurlyeq$, because $f = x\varnothing y\varnothing f = y\varnothing x\varnothing f$ and $f \sim f$ for all $x, y \in X$ and $f \in \mathcal{F}$. Property (ii) follows from the non-degeneracy of $\succcurlyeq$, i.e., there exist $x, y \in X$ such that $xSy \succ ySx$. Property (iii) follows from the symmetry of $\sim$, given that $xEyFf \sim yExFf$ is equivalent to $xFyEf \sim yFxEf$ for all $x, y \in X$ and $f \in \mathcal{F}$. Property (iv) follows from the transitivity of $\succcurlyeq$, because $x(E \cup E')y(F \cup F')f = xExE'yFyF'f \sim yExE'xFyF'f \sim yEyE'xFxF'f = y(E \cup E')x(F \cup F')f$ for any $x, y \in X$ and $f \in \mathcal{F}$.

Q.E.D.

PROOF OF PROPOSITION 2: Let $\{e_1, \ldots, e_n\}$ be the partition of the state space generating the same algebra as is generated by $\{E_i\} \cup \{F_i\}$. We have $\sum_{i=1}^m p(E_i) = \sum_{i=1}^n a_i p(e_i)$ and $\sum_{i=1}^m p(F_i) = \sum_{i=1}^n b_i p(e_i)$, where a_i and b_i denote the number of events containing e_i in $\{E_i\}$ and $\{F_i\}$, respectively. If $\{E_i\} \setminus\!\!\setminus \{F_i\} = \varnothing$, then $a_i \leq b_i$ for all $1 \leq i \leq n$, which in turn implies that $\sum_{i=1}^m p(E_i) \leq \sum_{i=1}^m p(F_i)$. Because, by hypothesis, $p(E_i) \geq p(F_i)$ for any $1 \leq i \leq m$, it must be that $p(E_i) = p(F_i)$ for all $1 \leq i \leq m$, and $a_i = b_i$ for all $1 \leq i \leq n$. Consequently, $p(\{F_i\} \setminus\!\!\setminus \{E_i\}) = \sum_{i=1}^n (b_i - a_i)^+ p(e_i) = 0$, where $(x)^+$ denotes the non-negative part of x. This establishes Proposition 2(i).

To establish part (ii), we use the partition described above to write

$$0 \leq (<) \sum_{i=1}^{m} \Big(p(E_i) - p(F_i) \Big) = \sum_{i=1}^{n} (a_i - b_i) p(e_i)$$
$$= \sum_{i=1}^{n} \Big((a_i - b_i)^+ - (b_i - a_i)^+ \Big) p(e_i) = p\Big(\{E_i\} \setminus\!\!\setminus \{F_i\}\Big) - p\Big(\{F_i\} \setminus\!\!\setminus \{E_i\}\Big).$$

Q.E.D.

Henceforth, for a general element α of $\mathbb{R}^n$, we write $\alpha \geq \mathbf{0}$ if $\alpha_i \geq 0$ for all i, $\alpha \gg \mathbf{0}$ if $\alpha_i > 0$ for all i, and $\alpha > \mathbf{0}$ if $\alpha \geq \mathbf{0}$ and $\alpha \neq \mathbf{0}$; $|\alpha|$ is the vector of absolute values of the components of α and $\sum \alpha$ is the sum of the components of α. In the following lemmas and proofs, a finite collection of vectors $A = \{\alpha_1, \ldots, \alpha_m\}$ can be identified as column vectors in a matrix.

We now introduce indicator vectors for events and a corresponding representation for the exchangeability relation $\approx$ using N-dimensional vectors in $\mathcal{K} = \{-1, 0, 1\}^N$. For any fixed index of the states $S = \{s_1, \ldots, s_N\}$, an event $E \subseteq S$ can be identified by a subset I_E of $\{1, \ldots, N\}$ such that $E = \{s_i\}_{i \in I_E}$. We represent E in $\mathcal{K}$ using the vector $\delta_E \in \mathcal{K}$, whose i-th coordinate is defined to be 1 when $i \in I_E$, and 0 otherwise. The vector in $\mathcal{K}$ representing a singleton event, $\{s_i\}$, is denoted by δ_i. The relation, $E \approx F$, also has a vector representation: $\epsilon = \delta_E - \delta_F \in \mathcal{K}$. In particular, $\varnothing \approx \varnothing$ is represented by $\mathbf{0} = \langle 0, \ldots 0 \rangle$, meaning that $E \approx \varnothing$ is represented by δ_E.

Define $\mathcal{E} \equiv \{\epsilon \mid E \approx F, \text{ where } E, F \subseteq S\}$ to be the subset of $\mathcal{K}$ consisting of all vectors representing exchangeable events. By definition, $\mathcal{E}$ contains 0 and the reflexivity of $\approx$ implies that $\epsilon \in \mathcal{E}$ iff $-\epsilon \in \mathcal{E}$. After fixing an index for S, a probability p can be identified as a N-dimensional vector $p = \langle p_1, \ldots p_N \rangle$ such that $p_i \geq 0$ and $\sum_{i=1}^N p_i = 1$. If $p(E) = p(F)$ for $E \approx F$, we have $\epsilon^T p = 0$ for the vector ϵ representing $E \approx F$. Using this notation, Theorem 1 reduces to the statement that there exists $p > \mathbf{0}$ in $\mathbb{R}^N$ such that for any $v \in \mathcal{K}$, $v^T p = 0$ iff $v \in \mathcal{E}$. We now derive a series of results that are useful in proving Theorem 1.

LEMMA 1: *The Cancellation Axiom holds if and only if for any finite sequence $\epsilon_1, \cdots, \epsilon_m \in \mathcal{E}$, whenever $\sum_{i=1}^m \epsilon_i / k$ is an element of $\mathcal{K}$ for some positive integer k, then $\sum_{i=1}^m \epsilon_i / k$ is also in $\mathcal{E}$.*

PROOF: Assume that Cancellation holds and fix some finite sequence $\epsilon_1, \cdots, \epsilon_m \in \mathcal{E}$ such that $\sum_{i=1}^m \epsilon_i / k$ is an element of $\mathcal{K}$ for k, a positive integer. By definition, each ϵ_i represents a pair of exchangeable events, E_i and F_i. Let $\{E_i\} \approx \{F_i\}$ be the pair of exchangeable families of events defined by $\epsilon_1, \cdots, \epsilon_m$. Because $\sum_{i=1}^m \epsilon_i / k \in \mathcal{K}$, $\{E_i\} \setminus\!\!\setminus \{F_i\}$ and $\{F_i\} \setminus\!\!\setminus \{E_i\}$ are of the same order k. By Cancellation, $\{E_i\} \setminus\!\!\setminus \{F_i\} \approx \{F_i\} \setminus\!\!\setminus \{E_i\}$, and this implies that $\sum_{i=1}^m \epsilon_i / k \in \mathcal{E}$.

To prove necessity, consider any pair of exchangeable families $\{E_i\} \approx \{F_i\}$, and let ϵ_i represent $E_i \approx F_i$. If both $\{E_i\} \setminus\!\!\setminus \{F_i\}$ and $\{F_i\} \setminus\!\!\setminus \{E_i\}$ are of the same order k, then $\sum_{i=1}^m \epsilon_i / k \in \mathcal{K}$. If $\sum_{i=1}^m \epsilon_i / k$ is in $\mathcal{E}$ whenever it is in $\mathcal{K}$, then $\{E_i\} \setminus\!\!\setminus \{F_i\} \approx \{F_i\} \setminus\!\!\setminus \{E_i\}$, and Cancellation holds. *Q.E.D.*

LEMMA 2: *Strong Event Non-satiation holds if and only if for any finite sequence $\epsilon_1, \cdots, \epsilon_m \in \mathcal{E}$, if $\sum_{i=1}^m \epsilon_i \geq \mathbf{0}$, then $\delta_j \leq \sum_{i=1}^m \epsilon_i$ implies that $\delta_j \in \mathcal{E}$ for any $1 \leq j \leq N$ (i.e., δ_j is null).*

PROOF: Assume that Strong Event Non-satiation holds and that $\sum_{i=1}^m \epsilon_i \geq \mathbf{0}$ for a finite sequence $\epsilon_1, \cdots, \epsilon_m \in \mathcal{E}$. Let $\{E_i\} \approx \{F_i\}$ be the pair of exchangeable families of events represented by $\epsilon_1, \cdots, \epsilon_m$. Note that $\sum_{i=1}^m \epsilon_i \geq \mathbf{0}$ implies that there is no state that appears in the family $\{F_i\}$ more times than in the family $\{E_i\}$. Hence, $\{F_i\} \setminus\!\!\setminus \{E_i\} = \varnothing$, and by Strong Event Non-satiation, $\{E_i\} \setminus\!\!\setminus \{F_i\}$ is null. In particular, any singleton subevent $\{s_j\}$ of $\{E_i\} \setminus\!\!\setminus \{F_i\}$ is also null and therefore exchangeable with $\varnothing$. Note that $\delta_j \leq \sum_{i=1}^m \epsilon_i$ iff $\{s_j\} \in \{E_i\} \setminus\!\!\setminus \{F_i\}$. Consequently, $\delta_j \in \mathcal{E}$ for all $\delta_j \leq \sum_{i=1}^m \epsilon_i$.

To establish necessity, let ϵ_i represent $E_i \approx F_i$ for any $\{E_i\} \approx \{F_i\}$ such that $\{F_i\} \setminus\!\!\setminus \{E_i\} = \emptyset$. The latter implies that $\sum_{i=1}^{m} \epsilon_i \geq \mathbf{0}$ because there are no excess states in $\{F_i\}$ relative to $\{E_i\}$. If, for any $1 \leq j \leq N$, $\delta_j \in \mathcal{E}$ whenever $\delta_j \leq \sum_{i=1}^{m} \epsilon_i$, then any singleton event $s_j \in \{E_i\} \setminus\!\!\setminus \{F_i\}$ is a null event. Because $\{E_i\} \setminus\!\!\setminus \{F_i\}$ is the union of a finite collection of pairwise disjoint null events, it too is null.

Q.E.D.

In the following, for any $q \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and $a \in \mathbb{R}$, we use $[\alpha^T q = a] \subset \mathbb{R}^n$ to denote the hyperplane $\{\alpha \in \mathbb{R}^n : \alpha^T q = a\}$. A hyperplane $[\alpha^T q = a]$ strongly separates the sets B and C if there exists a real number, $c \neq 0$, such that $\alpha^T q < a - c$ for all $\alpha \in B$ and $\alpha^T q > a + c$ for all $\alpha \in C$.

LEMMA 3: *Suppose that the closed convex cone C generated by a finite subset $D \subseteq \mathbb{R}^n$, satisfying $\mathbf{0} \notin D$, is pointed,*[13] *then for any $\beta \notin C$ the set $Q = \{q \in \mathbb{R}^n : \beta^T q < 0 \text{ and } \forall \alpha \in D, \alpha^T q > 0\}$ is not empty and has internal points.*[14]

PROOF: Because D is finite, it should be clear that Q has internal points as long as it is not empty. It remains to prove that Q is non-empty. Because $\beta \notin C$, it can be strictly separated from C by some q, i.e., $\alpha^T q \geq 0$ for all $\alpha \in C$ and $\beta^T q < 0$. Index elements of D using $1 \leq i \leq |D|$. For any $\alpha_i \in D$, we have $-\alpha_i \notin C$, otherwise C would not be pointed. Let $-\alpha_i$ and C be strictly separated by q_i. Notice that $\alpha_i^T q_i > 0$ and that $\alpha_D^T q_i \geq 0$ for all $\alpha_D \in D$. Define $q^* = q + \frac{1}{|D|} \sum_{i=1}^{|D|} b_i q_i$ where $0 < b_i < |\beta^T q / \beta^T q_i|$. Then $\beta^T q^* < 0$ and $\alpha_D^T q^* > 0$ for all $\alpha_D \in D$.

Q.E.D.

LEMMA 4: *Suppose that A is a finite subset of $\mathbb{R}^n$ and Q is a subset of $\mathbb{R}^n$ that has internal points, then for any $\beta \notin A$, there exists $q \in Q$ such that $A \cap [\alpha^T q = \beta^T q] = \emptyset$.*

PROOF: Note that $A \cap [\alpha^T q = \beta^T q] = \emptyset$ if and only if $(\alpha - \beta)^T q \neq 0$ for all $\alpha \in A$. Let $A - \beta$ be the collection of vectors of the form $\alpha - \beta$ where $\alpha \in A$. Abusing notation, we also use $A - \beta$ to refer to the $n \times |A|$ matrix whose $|A|$ columns are the vectors in $A - \beta$, identified up to a permutation of the columns. Then, we can write that $A \cap [\alpha^T q = \beta^T q] = \emptyset$ if and only if $|(A - \beta)^T q| \gg \mathbf{0}$. Note that, because $\beta \notin A$, $A - \beta$ has no column vector equal to $\mathbf{0}$.

If $\text{rank}(A) = 1$ then, because Q has internal points and $A - \beta \neq \mathbf{0}$, there is some $q \in Q$ such that $(A - \beta)^T q \neq 0$. Proceeding inductively, assume that the result holds when $\text{rank}(A) = m$ and write $\hat{B} = (C \;\; d)$, where C has $m \geq 1$ non-zero columns and d is a single non-zero column vector in $\mathbb{R}^n$. Let $q' \in Q$ be an internal point solution to $|C^T q| \gg \mathbf{0}$. We construct a solution to $|\hat{B}^T q| \gg \mathbf{0}$ of the form $q' + bd \in Q$ for a sufficiently small scalar $b \geq 0$. Let $\max(|C^T d|)$ be the maximum of the absolute values of the components of $C^T d$, and $\min(|C^T q'|)$ be the minimum of the absolute values of the components of $C^T q'$. Then, $|C^T(q' + bd)| \gg \mathbf{0}$ for any $b < \frac{\min(|C^T q'|)}{\max(|C^T d|)}$ when $\max(|C^T d|) > 0$, and for any finite b in the case of $\max(|C^T d|) = 0$. That is, we have shown that $C^T(q' + bd) \neq 0$ for every b in some positive finite interval I. Because $d^T d > 0$, it cannot be that $d^T(q' + bd)$ is zero for all $b \in I$. Thus, because q' is internal, there is some $b \in I$ such that $q = (q' + bd)$ is in Q and $|\hat{B}^T q| \gg \mathbf{0}$. This establishes that, for any $\text{rank}(A)$, we can find $q \in Q$ such that $|(A - \beta)^T q| \gg \mathbf{0}$. *Q.E.D.*

[13] A closed convex cone C is pointed if $C \cap (-C) = \{\mathbf{0}\}$.

[14] A point q is an internal point of a subset Q of $\mathbb{R}^n$ if for any $q' \in \mathbb{R}^n$, there exists $b_0 > 0$ such that $q + bq' \in Q$ for all $0 < b \leq b_0$.

Let $A = \mathcal{K} \backslash \mathcal{E}$, B be the subspace generated by $\mathcal{E}$, and C be the convex cone generated by $D = \{\delta_1, \ldots, \delta_N\} \backslash \mathcal{E}$. The set A corresponds to the vector representation of pairs of events that are not exchangeable, while the set D corresponds to the non-null states in S. The following results relating A, C, and D are useful.

LEMMA 5: *For A, B, C, D, we have the following:*

(i) $D \subseteq A$

(ii) *C is a pointed convex cone*

(iii) *Cancellation implies that* $A \cap B = \varnothing$

(iv) *Strong Event Non-satiation implies that* $B \cap C = \{\mathbf{0}\}$

PROOF: Results (i) and (ii) are immediate. To establish $A \cap B = \varnothing$, we first show that $\mathcal{E} = B \cap \mathcal{K}$. Let $0 \leq i \leq |\mathcal{E}|$ index the vectors in $\mathcal{E}$, each of which is denoted as ε_i. If $\epsilon \in B \cap \mathcal{K}$, then $\epsilon = \sum_{i=1}^{|\mathcal{E}|} \mu_i \varepsilon_i$, where the μ_i's can be assumed to be rational numbers because ϵ and all of the ε_i's are integer-valued vectors. By choosing k to be the common denominator of the μ_i's, one can find a finite collection of vectors in $\mathcal{E}$, denoted $\{\epsilon_i\}_{i=1}^m$ such that $\sum_{i=1}^{|\mathcal{E}|} \mu_i \varepsilon_i = \sum_{j=1}^m \epsilon_j / k$.[15] By Lemma 1, Cancellation implies that $\epsilon \in \mathcal{E}$ and thus $B \cap \mathcal{K} \subseteq \mathcal{E}$. Equality of the two sets follows from the fact that $\mathcal{E}$ generates B and $\mathcal{E} \subseteq \mathcal{K}$, meaning that $\mathcal{E} \subseteq B \cap \mathcal{K}$. Now note that $\mathcal{E} = B \cap \mathcal{K} = B \cap (\mathcal{E} \cup A)$. Because $\mathcal{E}$ and A are disjoint and $\mathcal{E} \subseteq B$, it must be that $B \cap A = \varnothing$, which proves part (iii).

Next, we prove part (iv), that Strong Event Non-satiation implies that $B \cap C = \{\mathbf{0}\}$. Let $\hat{\mathcal{E}}$ be a $N \times M$ matrix of rank M whose columns span B. Then, any element of B can be written as $\hat{\mathcal{E}} u$ for some $u \in \mathbb{R}^M$. Correspondingly, for any $v \in C$, $v \geq 0$ and its i^{th} component must be zero whenever s_i is null. We can write $B \cap C = \{v \in C \mid \hat{\mathcal{E}} u = v,\ u \in \mathbb{R}^M\}$. Suppose that $v \in B \cap C \setminus \{\mathbf{0}\}$ has $L > 0$ non-zero components. The solution to $\hat{\mathcal{E}} u = v$ can be written as

$$\hat{\mathcal{E}}_P u = v_P \in \mathbb{R}^L$$
$$\hat{\mathcal{E}}_Z u = 0 \in \mathbb{R}^{N-L},$$

where v_P corresponds to the strictly positive elements of v, $\hat{\mathcal{E}}_P$ is the submatrix of $\hat{\mathcal{E}}$ with rows corresponding to v_P, and $\hat{\mathcal{E}}_Z$ is the matrix containing the remaining rows of $\hat{\mathcal{E}}$. Because $\hat{\mathcal{E}}_Z$ is a rational matrix, by Corollary 2 of Kraft et al. (1959), the solution space for $\hat{\mathcal{E}}_Z u = 0$ has a rational basis.[16] Because u is in this span, one can find $u_Q \in \mathbb{Q}^M$ in the same span and arbitrarily close to u so that $\hat{\mathcal{E}}_P(u - u_Q)$ is arbitrarily close to zero. This implies that if $B \cap C \setminus \{\mathbf{0}\}$ is non-empty, then so too is $\mathbb{Q}^M \cap B \cap C \setminus \{\mathbf{0}\}$. For any u in the latter set, one can write $\sum_{i=1}^M u_i \hat{\varepsilon}_i = v > 0$ with $v \in C$, and where $\hat{\varepsilon}_i \in \mathcal{E}$ denotes the ith column of $\hat{\mathcal{E}}$. Because $v > 0$ and C is a cone, one can rescale both the u_i's and v by a sufficiently large integer to guarantee that each u_i is an integer *and* that $\delta_j \leq v$ for some non-null s_j. As shown in proving part (iii), one can find a finite collection of vectors in $\mathcal{E}$, denoted $\{\epsilon_i\}_{i=1}^m$, such that $\delta_j \leq \sum_{i=1}^M u_i \hat{\varepsilon}_i = \sum_{i=1}^m \epsilon_i$ (where the u_i's have been appropriately rescaled). By Strong Event Non-satiation and Lemma 2, this implies that s_j is null—a contradiction. Thus $B \cap C = \{\mathbf{0}\}$. *Q.E.D.*

[15]The collection $\{\epsilon_i\}_{i=1}^m$ may contain repeated terms.

[16]The corollary says that, if the coefficients of a finite system of homogeneous linear equalities and inequalities are in $\mathbb{Q}$, then if it has solutions in $\mathbb{R}^M$, it has solutions in $\mathbb{Q}^M$.

PROOF OF THEOREM 1: Let $\lambda_1, \ldots, \lambda_M$ be the basis for B, and $\{\lambda_i\}_{i=M+1}^N$ be its orthogonal extension to a basis of $\mathbb{R}^N$. Let T be a square matrix with columns $\lambda_1, \ldots, \lambda_N$ and define $T^{-1}_{[M+1,N]} = \begin{pmatrix} 0 & I_{N-M} \end{pmatrix} T^{-1}$, where $\begin{pmatrix} 0 & I_{N-M} \end{pmatrix}$ is the $(N-M) \times N$ matrix with zeros in the first M columns and the $(M-N)$-identity matrix in the remaining columns. For any $x \in \mathbb{R}^N$, $T^{-1}_{[M+1,N]}x$ is the vector of coefficients of $\lambda_{M+1}, \ldots, \lambda_N$ in the $\{\lambda_i\}_{i=1}^N$-basis representation of x. By construction, $T^{-1}_{[M+1,N]}$ maps $\alpha \in \mathbb{R}^N$ to $\mathbf{0}$ if and only if $\alpha \in B$.

Define $A' = T^{-1}_{[M+1,N]}(A)$, $D' = T^{-1}_{[M+1,N]}(D)$, and $C' = T^{-1}_{[M+1,N]}(C)$. Because C is the convex cone generated by D, C' is a convex cone generated by D'. Observe that Lemma 5 (iii) implies that $\mathbf{0} \notin A'$ and, by Lemma 5 (i), we have $\mathbf{0} \notin D'$.

We now demonstrate that C' is pointed. Otherwise, there exist $c_1, c_2 \in C \setminus \{\mathbf{0}\}$ such that $T^{-1}_{[M+1,N]}c_2 = -\gamma T^{-1}_{[M+1,N]}c_1 \neq \{\mathbf{0}\}$ for some $\gamma > 0$. Note that $T^{-1}_{[M+1,N]}(c_2 + \gamma c_1) = \{\mathbf{0}\}$ implies that $c_2 + \gamma c_1 \in B$, while the convexity of C implies that $c_2 + \gamma c_1 \in C$. Lemma 5 (iv) therefore implies that $c_2 + \gamma c_1 = \{\mathbf{0}\}$, which means that C is not a pointed cone, thus contradicting Lemma 5 (ii).

D' is finite, does not contain $\{\mathbf{0}\}$, and generates C' (a pointed cone). The application of Lemma 3 to any $\beta \notin C'$ therefore implies that the collection $Q = \{q \in \mathbb{R}^{N-M} \mid \delta^T q > 0 \ \forall \delta \in D'\}$ is non-empty and has internal points. Moreover, because A' is finite and $\mathbf{0} \notin A'$, Lemma 4 implies that there exists $q^* \in Q$ defining a hyperplane through $\mathbf{0}$ that is disjoint from A'. Define $q^*_N = (0, (q^*)^T)^T \in \mathbb{R}^N$ as q^* with zeros for the first M coordinates. Then, $p = \frac{(T^{-1})^T q^*_N}{\sum (T^{-1})^T q^*_N}$ is a probability representing $\approx$. To see this, first recall that, for any $\epsilon \in \mathcal{E} \subset B$, the last $N-M$ components of $T^{-1}\epsilon$ are zero (by definition). Thus, by construction, $\epsilon^T p = \frac{(T^{-1}\epsilon)^T q^*_N}{\sum (T^{-1})^T q^*_N} = 0$. This means that exchangeable events have an equal probability under p. Likewise, because A' is disjoint from the hyperplane defined by q^*, $\alpha^T p \neq 0$ for any $\alpha \in A$. In other words, non-exchangeable events have distinct probabilities under p. Finally, because $q^* \in Q$, D' is non-empty and does not contain zero, we have $\delta_i^T p \propto (T^{-1}\delta_i)^T q^*_N \geq 0$ for all δ_i and the inequality is strict for any $\delta_i \in D$ (corresponding to a non-null state, s_i). Thus, p is a probability representing $\approx$. *Q.E.D.*

PROOF OF PROPOSITION 3: If $E \succcurlyeq *cF$, it should be clear that $p(E) \geq p(F)$ for all representations of $\approx$.

To demonstrate the converse, let A', C', D', and $T^{-1}_{[M+1,N]}$ be defined as in the proof of Theorem 1. For any two events, E and F, if $\beta = T^{-1}_{[M+1,N]}(\delta_E - \delta_F) \notin C'$, then Lemma 3 implies that we can find a set Q with internal points such that each $q \in Q$ strictly separates D' and β. One can then use Lemma 4 and follow the proof of Theorem 1 to construct a probability, p, representing $\approx$ using some $q^* \in \mathbb{R}^{N-M} \in Q$ and implying that $p(E) < p(F)$. Thus, $p(E) \geq p(F)$ for all representations of $\approx$ implies that $T^{-1}_{[M+1,N]}(\delta_E - \delta_F) \in C'$. The latter inclusion is equivalent to writing $\delta_E - \delta_F = b + c$, where $b \in B$ and $c \in C$.

Let $\hat{T}$ be an $N \times N$ matrix of column vectors, $\{\hat{\lambda}_i\}_{i=1}^N$, forming a basis for $\mathbb{R}^N$ and where the first M columns are in $\mathcal{E}$ while the remaining columns are in D. Define $\zeta = \hat{T}^{-1}(\delta_E - \delta_F)$ and note that ζ is in $\mathbb{Q}^N$.[17] Letting $\{u_i\}_{i=1}^M$ be the first M elements of ζ and $\{v_j\}_{j=1}^{N-M}$ be the remaining $N-M$ elements, $(\delta_E - \delta_F) = \hat{T}\zeta = \sum_{i=1}^N \zeta_i \hat{\lambda}_i = \frac{1}{w_{GCD}} \sum_{i=1}^N (w_{GCD}\zeta_i)\hat{\lambda}_i$, where w_{GCD} is the greatest common denominator of elements of ζ. Because $\hat{\lambda}_i \in \mathcal{E}$ if $i \leq M$ and otherwise $\hat{\lambda}_i \in D$, and noting that each $w_{GCD}\zeta_i$ is an

[17]By definition, $\hat{T}$ has rational elements. Thus, $\hat{T}^{-1}$ has rational elements because matrix inversion corresponds to a finite sequence of arithmetic operations on elements of $\hat{T}$.

integer, one can write

$$\delta_E - \delta_F = \frac{1}{w_{GCD}} \Big(\sum_{i=1}^{m} \epsilon_i + \sum_{i=1}^{l} \nu_i \Big),$$

where $\{\epsilon_i\}$ and $\{\nu_i\}$ are collections (allowing repetition) in $\mathcal{E}$ and D, respectively. Because, by definition, each $\epsilon_i = \delta_{E_i} - \delta_{F_i}$ corresponds to a pair of exchangeable events, $E_i \approx F_i$, one can construct the m-member exchangeable families $\{E_i\}_{i=1}^m$ and $\{F_i\}_{i=1}^m$. Identifying ν_i with its associated state, s_i, one can define $\{e_i\}_{i=1}^l = \{s_i\}_{i=1}^l$. This pins down $\{E_i\}, \{F_i\}$, and $\{e_i\}_{i=1}^l$ in Definition 5. It should be clear that the positive (resp. negative) projection of $\left(\sum_{i=1}^m \epsilon_i + \sum_{i=1}^l \nu_i \right)$ corresponds to $\delta_{E \setminus F}$ (resp. $\delta_{F \setminus E}$). Thus $\{E_i, e_j\} \Bbbk \{F_i\} = E \setminus F$, $\{F_i\} \Bbbk \{E_i, e_j\} = F \setminus E$, and both family differences are of the order w_{GCD}. Under Definition 5, this implies that $E \succcurlyeq^* F$.

Q.E.D.

PROOF OF THEOREM 2: Suppose that Cancellation, Strong Event Non-satiation, and Extended Completeness hold. By Theorem 1, there exists a probability representing $\approx$. Let p and p' be two distinct probabilities representing $\approx$. Define $\bar{\eta} = \max_{s_i \in D} \frac{p(s_i)}{p'(s_i)}$, $E = \operatorname{argmax}_{s_i \in D} \frac{p(s_i)}{p'(s_i)}$, $\underline{\eta} = \min_{s_i \in D} \frac{p(s_i)}{p'(s_i)}$ and $F = \operatorname{argmin}_{s_i \in D} \frac{p(s_i)}{p'(s_i)}$. Because $p \neq p'$, it must be that $E \cap F = \varnothing$ and $\underline{\eta} < 1 < \bar{\eta}$. For any $\varepsilon \in (0, 1)$, define

$$p_\varepsilon = \frac{p - (1-\varepsilon)\underline{\eta} p'}{\sum_i \big(p(s_i) - (1-\varepsilon)\underline{\eta} p'(s_i) \big)}, \; p'_\varepsilon = \frac{p' - (1-\varepsilon)\frac{1}{\bar{\eta}} p}{\sum_i \big(p'(s_i) - (1-\varepsilon)\frac{1}{\bar{\eta}} p(s_i) \big)}.$$

Note that p_ε and p'_ε are probabilities that represent $\approx$. Moreover, by construction, one can find some ε such that $p_\varepsilon(F) < p_\varepsilon(E)$ and $p'_\varepsilon(E) < p'_\varepsilon(F)$. Proposition 3 implies that E and F are not comparable via $\succcurlyeq^*$, which is inconsistent with Extended Completeness. The probability representing $\approx$ must therefore be unique.

When a representation, p, of $\approx$ exists, Theorem 1 implies Cancellation and Strong Event Non-satiation. If p is unique, Extended Completeness of $\succcurlyeq^*$ is a trivial application of Proposition 3.

Q.E.D.

PROOF OF THEOREM 4: Suppose that p is the unique probability representing $\approx$ and that it is reduction consistent. Theorem 2 implies Cancellation, Strong Event Non-satiation, and Extended Completeness. Exchangeability Independence is satisfied because every indifference relation hypothesized in the statement of the axiom is a trivial consequence of reduction consistency.

To prove the converse, note first that Cancellation, Strong Event Non-satiation, and Extended Completeness imply the existence of a unique representing probability, p. To complete the proof, we need only show that p is reduction consistent.

We say that two m-partitions of the state space, $\{E_i\}_{i=1}^m$ and $\{F_i\}_{i=1}^m$, are exchangeable whenever either $E_i \approx F_i$ or $E_i = F_i$ for all $1 \leq i \leq m$. For any exchangeable m-partitions, $\{E_i\}$ and $\{F_i\}$, if $x_1 E_1 \cdots x_{m-1} E_{m-1} x_j \sim x_1 F_1 \cdots x_{m-1} F_{m-1} x_j$ for every $x_1, \ldots, x_{m-1} \in X$ and j arbitrary in $\{1, \cdots, m-1\}$, then Exchangeability Independence implies that $x_1 E_1 \cdots x_{m-1} E_{m-1} y \sim x_1 F_1 \cdots x_{m-1} F_{m-1} y$ for every $y \in X$. We will use this observation to complete the proof.

First consider all two-outcome acts, $f = xEy$ and $g = xFy$ such that $E \sim^* F$, and thus $p(E) = p(F)$. Note that one can write $f = x(E \setminus F)y(F \setminus E)x$ and $g = x(F \setminus E)y(E \setminus F)x$. Because $p(E \setminus F) = p(F \setminus E)$, it must be that $E \setminus F \approx F \setminus E$, which yields $f \sim g$. Note that this holds for any $(x, y) \in X^2$. Proceed now by induction and consider any m-outcome acts ($m \geq 3$), f and g, that induce the same lottery under p. Note that one can write $f = x_1 E_1 \cdots x_{m-1} E_{m-1} x_m$ with $E_i = f^{-1}(x_i)$ for some distinct $(x_1, \ldots, x_m) \in X^m$.

Likewise, write $g = x_1F_1 \cdots x_{m-1}F_{m-1}x_m$ with $F_i = g^{-1}(x_i)$. Because f and g induce the same lottery, $p(E_i) = p(F_i)$, which by Proposition 3 implies that $E_i \sim^* F_i$ for each $i \in \{1, \ldots, m\}$. Under the induction hypothesis, which we confirmed in the two-outcome case, any $m-1$ outcome acts of the form x_iE_mf and x_iF_mg are indifferently ranked, for any $(x_1, \ldots, x_{m-1}) \in X^{m-1}$. Exchangeability Independence then implies that $f \sim g$.

Q.E.D.

PROOF OF THEOREM 3: The verification of the necessity of our axioms is straightforward. Note that the existence of either representation implies that p is reduction consistent. It remains to demonstrate that the STP and the comonotonic STP imply an EU and RLU representation respectively.

Case 1: Sure-thing principle (STP)

According to Debreu (1960), under the STP, there exists v_i such that preference is represented by $U(\boldsymbol{x}) = \sum_{i=1}^{N} v_i(x_i)$. By subtracting $U(\boldsymbol{0})$ from the representation, we can without loss of generality, assume that $v_i(0) = 0$ for each i. For a pair of disjoint events E and F with equal probabilities and a pair of outcomes x and y, $E \approx F$ implies that

$$(1) \qquad \sum_{s_i \in E} v_i(x) + \sum_{s_i \in F} v_i(y) = \sum_{s_i \in F} v_i(x) + \sum_{s_i \in E} v_i(y)$$

Because this equation holds for all pairs of outcomes, it must be that $\sum_{s_i \in E} v_i(x) - \sum_{s_i \in F} v_i(x) = 0$ for all $x \in \mathbb{R}$. From the proofs of Theorems 1 and 2, we know that rank $\mathcal{E}$ is $N-1$, meaning that, for each outcome x, there are $N-1$ independent equations like Equation (1) in the N unknowns $v_i(x)$. Because these equations are identical to those determining p, the unique probability representation for $\approx$, it must be that $v_i(x) = p_i v(x)$ for some $v(x)$, which in turn must be continuous and monotonic. This establishes the existence of an EU representation.

Case 2: Comonotonic STP

Each permutation $\boldsymbol{n}$ of $\{1, \ldots, N\}$ defines a comonotonic cone $\mathcal{F}_{\boldsymbol{n}} = \{\boldsymbol{x} \in \mathbb{R}^N : x_{\boldsymbol{n}_i} \geq x_{\boldsymbol{n}_j} \text{ if } i < j\}$. To proceed, we say that a permutation concentrates on event E if $s_{\boldsymbol{n}_i}, s_{\boldsymbol{n}_j} \in E$ implies that $s_{\boldsymbol{n}_k} \in E$ for all k such that $i < k < j$. Suppose that $\boldsymbol{n}$ concentrates on E with $s_{\boldsymbol{n}_k} \in E$ and $s_{\boldsymbol{n}_{k-1}} \notin E$, we say $A = \{s_{\boldsymbol{n}_i}\}_{i<k}$ is the event preceding E in $\boldsymbol{n}$.

Chew and Wakker (1996) have shown that there exists a unifying $v_{\boldsymbol{n},i}$ satisfying $v_{\boldsymbol{n},i}(0) = 0$ for all $\boldsymbol{n}$ and i such that $U(\boldsymbol{x}) = \sum_{i=1}^{N} v_{\boldsymbol{n},i}(x_{\boldsymbol{n}_i})$ if $\boldsymbol{x} \in \mathcal{F}_{\boldsymbol{n}}$. This implies that $\sum_{s_{\boldsymbol{n}_i} \in E} v_{\boldsymbol{n},i}(\cdot) = \sum_{s_{\boldsymbol{m}_i} \in E} v_{\boldsymbol{m},i}(\cdot)$ for $\boldsymbol{n}$ and $\boldsymbol{m}$ such that both $\boldsymbol{n}$ and $\boldsymbol{m}$ concentrate on E and $A = B$ where A precedes E in $\boldsymbol{n}$ and B precedes E in $\boldsymbol{m}$.

Now, suppose that $p(A) = p(B)$ and denote $D = A \cap B$, $E = A \backslash B$ and $F = B \backslash A$. It is then the case that $E \approx F$. Consider the following two acts, $f = \{D, w; E, x; F, y; (A \cup B)^c, z\}$ and $g = \{D, w; F, x; E, y; (A \cup B)^c, z\}$, where $w > x > y > z$. They are indifferent because of exchangeability, which implies that the following equation,

$$\sum_{s_i \in E} v_{i,\boldsymbol{n}}(x) + \sum_{s_i \in F} v_{i,\boldsymbol{n}}(y) = \sum_{s_i \in F} v_{i,\boldsymbol{m}}(x) + \sum_{s_i \in E} v_{i,\boldsymbol{m}}(y)$$

,
holds for all $\boldsymbol{n}$ and $\boldsymbol{m}$ such that they both concentrate on E, F, and $E \cup F$, such that $E \cup D$ precedes F in $\boldsymbol{n}$ and $F \cup D$ precedes E in $\boldsymbol{m}$.

Because this equation holds for all pairs of outcomes, we can derive that $\sum_{s_i \in E} v_{i,\boldsymbol{n}}(x) - \sum_{s_i \in F} v_{i,\boldsymbol{m}}(x) = 0$ for all $x \in \mathbb{R}$ and all $\boldsymbol{n}$ and $\boldsymbol{m}$ that concentrate on E and F, respectively, and the same event D precedes E and F in $\boldsymbol{n}$ and $\boldsymbol{m}$, respectively. This gives $\sum_{s_i \in A} v_{i,\boldsymbol{n}}(\cdot) = \sum_{s_i \in B} v_{i,\boldsymbol{m}}(\cdot)$ for all $p(A) = p(B)$, $\boldsymbol{n}$ and $\boldsymbol{n}$ concentrate A and B, respectively, and A and B have no preceding events in $\boldsymbol{n}$ and $\boldsymbol{m}$. Notice that $W(\cdot, A) = \sum_{s_{\boldsymbol{n}_i} \in A} v_{\boldsymbol{n},i}(\cdot)$ if $\boldsymbol{n}$ concentrates on A and A has no preceding event and this establishes the existence of an RLU representation.

Q.E.D.

PROOF OF PROPOSITION 4: Consider a state space S with n states. Suppose that there are two acts f and g, both inducing the same lottery under a probability representation p of exchangeability relation $\approx$.

Case (i): f and g have two distinct outcomes x, y. It follows that $f \sim g$ because $E = \{s \in S : f(s) = x, g(s) = y\} \approx F = \{s \in S : f(s) = y, g(s) = x\}$.

Case (ii): There are n distinct outcomes. Let $T = \{s \in S : f(s) \neq g(s)\}$, then $p(s) = p(t)$ for all $s, t \in T$. It follows that p is uniform on T and further that $f \sim g$.

Case (iii): f and g have $n-1$ distinct outcomes. Let x be the outcome assigned by f on two distinct states s_1 and s_2. Without loss of generality, consider four sub-cases: (i) $g(s_1) = g(s_2) = x$; (ii) $g(s_1) = g(s_2) = y$ for another outcome y; (iii) $g(s_1) = x$ and $g(s_2) = y$ for another outcome y; (iv) $g(s) = y$ and $g(s_2) = z$ for two different outcomes y and z. For each sub-case, $f \sim g$ follows from the same reasoning as that used in Case (ii).

Combining all three cases, the only remaining possibility for a probability representation of $\approx$ to not satisfy reduction consistency is when f and g both have three outcomes on a state space with five states. We omit further details in the demonstration that this possibility cannot arise.

Q.E.D.