School of Economics

Working Paper Series

Random Binary Choices that Satisfy Stochastic Betweenness

Matthew Ryan

2017/01

Random Binary Choices that Satisfy Stochastic Betweenness*

Matthew Ryan
School of Economics
Auckland University of Technology

December 2016

Abstract

Experimental evidence suggests that the process of choosing between *lotteries* (risky prospects) is stochastic and is better described through choice probabilities than preference relations. Binary choice probabilities admit a *Fechner representation* if there exists a utility function u such that the probability of choosing a over b is a non-decreasing function of the utility difference $u(a) - u(b)$. The representation is *strict* if $u(a) \geq u(b)$ precisely when the decision-maker is at least as likely to choose a from $\{a, b\}$ as to choose b. Blavatskyy (2008) obtained necessary and sufficient conditions for a strict Fechner representation in which u has the expected utility form. One of these is the *common consequence independence (CCI)* axiom (*ibid.*, Axiom 4), which is a stochastic analogue of the mixture independence condition on preferences. Blavatskyy also conjectured that by weakening CCI to a condition he called *stochastic betweenness (SB)* – a stochastic analogue of the betweenness condition on preferences (Chew (1983)) – one obtains necessary and sufficient conditions for a strict Fechner representation in which u has the *implicit expected utility* form (Dekel (1986)). We show that Blavatskyy's conjecture is false, and provide a valid set of necessary and sufficient conditions for the desired representation.

*I'm grateful to Pavlo Blavatskyy and Simon Grant for very helpful comments on an earlier draft. Special thanks to John Hillas for "stress testing" the paper, which has significantly improved both the exposition and my understanding of the results. Remaining deficiencies are, of course, entirely my own responsibility.

1 Introduction

Experimentalists have long observed that subjects often make different choices in successive presentations of the same choice problem.[1] Loomes (2005) observes that "[t]his phenomenon has most frequently been reported for pairwise choices between lotteries, where as many as 30% of respondents may choose differently on each occasion" (*ibid.*, p.301). It is implausible to dismiss such a high level of variability as a manifestation of indifference.

If a decision-maker has anything that might be called *preferences* over lotteries, it would seem that these are only "revealed" by choices in a noisy fashion. Since the mid-1990s, experimental evidence on expected utility (EU) has therefore been viewed through the lens of *probabilistic* models of choice. Most commonly, this lens has been some variant on the classic *Fechner model* (Falmagne (2002)). A Fechner model is characterised by a utility function over some set, A, of alternatives together with an auxiliary function that converts *utility differences* into choice probabilities. If $P(a,b)$ denotes the probability with which the decision-maker chooses alternative $a \in A$ over alternative $b \in A$ in a (forced) binary choice, then a Fechner model for P takes the form

$$P(a,b) \;=\; f(u(a)-u(b)) \tag{1}$$

where $u : A \to \mathbb{R}$ is a utility function and $f : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function that satisfies $f(x) + f(-x) = 1$.

It is natural to interpret u as a representation of the decision-maker's preferences, and f as a description of the noise that mediates between preference and choice. According to (1), the probability of making a utility-maximising choice from the set $\{a,b\}$ is at least $\frac{1}{2}$ and this probability is weakly increasing in the magnitude of the utility difference, $|u(a)-u(b)|$. If

$$u(a) \geq u(b) \quad \Leftrightarrow \quad P(a,b) \geq \frac{1}{2} \tag{2}$$

for all $a, b \in A$, then we call (1) a *strict* Fechner model (Ryan, 2015). Note that any Fechner model satisfies the "$\Rightarrow$" partof (2); it is the converse implication that distinguishes a "strict" Fechner model.

If alternatives are *lotteries* and u has the expected utility form, then (1) gives a probabilistic version of EU – a model of random binary choice guided by EU preferences, or "EU with Fechnerian noise". Blavatskyy (2008, Theorem 1) and Dagsvik (2008, Theorem

[1] Mosteller and Nogee (1951) is an early example.

4) provide axiomatic foundations for this model when A is the unit simplex in $\mathbb{R}^n$, interpreted as the set of lotteries over a fixed set of n possible outcomes – sufficient conditions for P to possess a Fechner representation with a *linear* utility function.[2] Blavatskyy's axioms are also necessary if the representation is strict, while Dagsvik's are necessary if f is strictly increasing and continuous, which is a stronger restriction on a Fechner model.[3] These representation theorems are important benchmarks in the literature on binary stochastic choice.

Probabilistic versions of *generalised* EU models can likewise be constructed by restricting u to a larger class of functions. A substantial body of experimental literature evaluates the relative merits of these noisy models of lottery choice.[4] Currently, however, axiomatic foundations are lacking for the probabilistic versions of most generalised EU models.

A rare exception is the *implicit expected utility (IEU)* model of Dekel (1986). The preferences described by an IEU function satisfy betweenness but need not satisfy mixture independence (Chew (1983), Dekel (1986)). Betweenness imposes mixture independence only within "linear segments" of the simplex A. Mixture independence requires that[5]

$$a \succsim b \quad \Leftrightarrow \quad a\lambda c \succsim b\lambda c \tag{3}$$

for any $a, b \in A$ and any $\lambda \in (0, 1)$. Betweenness imposes (3) only when there exist $e, f \in A$ such that the lotteries $a, b, c \in A$ can all be expressed as mixtures of e and f. The IEU functions therefore include the EU functions as a proper subset. Blavatskyy (2008, pp.1052-3) proposes an axiomatisation of the "IEU with Fechnerian noise" model – necessary and sufficient conditions for P to possess a strict Fechner representation with u of the IEU form.

Blavatskyy conjectures that the desired axiomatisation can be obtained by replacing the *common consequence independence (CCI)* condition from his axiomatisation of "EU with Fechnerian noise" with a weaker variant called (by us)[6] *stochastic betweenness (SB)*. The CCI axiom is a probabilistic variant of mixture independence. It states that

$$P(a\lambda c, b\lambda c) = P(a\lambda d, b\lambda d) \tag{4}$$

[2] See also Dagsvik (2015).

[3] A Fechner model in which f is strictly increasing is called a *strong Fechner* (or *strong utility*) model. Any strong Fecner model is strict – see Ryan (2015).

[4] Hey (2014) is an excellent recent survey.

[5] As usual, $a\lambda c$ denotes the λ-mixture of a and c (with weight λ on a), and $b\lambda c$ is defined similarly. Formal definitions are given in the next section.

[6] Blavatskyy (2008) himself uses the term *betweenness* axiom, but this invites confusion with the betweenness property of preference relations.

for any lotteries $a, b, c, d \in A$ and any $\lambda \in (0,1)$. Stochastic betweenness only requires that (4) hold when there exist $e, f \in A$ such that the lotteries $a, b, c, d \in A$ can all be expressed as mixtures of e and f.

Despite its plausibility, Blavatskyy's conjecture turns out to be false, as we show in Section 3. In particular, there exist Fechner models with u of the IEU form that violate SB. In Section 4 we modify Blavatskyy's axioms to obtain a set of necessary and sufficient conditions for P to possess a strict Fechner representation in which u is an IEU function (Theorem 2).

In preparation for our main results, the next section reviews implicit expected utility theory and the axioms of Blavatskyy (2008).

2 Stochastic IEU

We adopt the framework of Blavatskyy (2008) and Dagsvik (2008). Let A be the unit simplex in $\mathbb{R}^n$, interpreted as the set of *lotteries* over a given set $X = \{x_1, ..., x_n\}$ of outcomes. We use $\delta^i \in A$ to denote the lottery that places probability 1 on x_i. Following standard convention, if $a, b \in A$ and $\lambda \in [0,1]$ then $a\lambda b$ will denote the convex combination $\lambda a + (1-\lambda)\, b$. It will also be useful to introduce the following notation for linear segments (intervals): for any $e, f \in A$, the closed interval with end points e and f is

$$[e, f] \equiv \{e\lambda f \ \mid \lambda \in [0,1]\}.$$

The open and half-open intervals (e, f), $(e, f]$ and $[e, f)$ are defined analogously.[7]

We consider binary choice problems in which pairs of alternatives are drawn from the set A. Choice behaviour may exhibit randomness, so each decision-maker will be characterised by a collection of choice probabilities rather than a preference relation. A *binary choice probability (BCP)* is a mapping

$$P : A \times A \to [0,1]$$

that satisfies the following *completeness* property:[8] for any $a, b \in A$

$$P(a, b) + P(b, a) = 1 \tag{5}$$

If $a \neq b$, the quantity $P(a, b)$ is the probability (or, in behavioural terms, the frequency) with which the decision-maker selects a when given the choice between a or b (abstention

[7] Note that $[e, f] = [f, e]$, $(e, f) = (f, e)$ and $(e, f] = [f, e)$.

[8] Also known as the *balancedness* condition (see, for example, Definition 4.9 in Falmagne, 1985).

not being an option). No behavioural interpretation is given to $P(a,b)$ when $a = b$, but it is conventional to define BCPs on the entire Cartesian product $A \times A$ for convenience. The completeness condition (5) implies that

$$P(a,a) = \frac{1}{2}$$

for every $a \in A$.

Given a binary choice probability P, we may define the following binary relation on A:

$$a \succsim^P b \quad \Leftrightarrow \quad P(a,b) \geq \frac{1}{2} \tag{6}$$

We call $\succsim^P$ the decision-maker's *stochastic preference relation*. The binary relations $\succ^P$ and $\sim^P$ are determined from $\succsim^P$ in the usual way.

A *Fechner representation* for P is a pair (u, f), where $u : A \to \mathbb{R}$ is a utility function and $f : \mathbb{R} \to \mathbb{R}$ is a non-decreasing function satisfying $f(x) + f(-x) = 1$ for all $x \in \mathbb{R}$, such that

$$P(a,b) \;=\; f(u(a) - u(b)) \tag{7}$$

for any $a, b \in A$.[9] If u also represents $\succsim^P$ (in the usual sense) then we say that (u, f) is a *strict Fechner representation* for P.[10] If f is strictly increasing, then (u, f) is a *strong Fechner representation*.

If (u, f) is a Fechner representation for P, then $f(0) = \frac{1}{2}$ and hence

$$u(a) \geq u(b) \quad \Rightarrow \quad P(a,b) \geq \frac{1}{2}$$

for all $a, b \in A$. The converse holds iff the representation is strict. In other words, if (u, f) is a non-strict Fechner representation for P, then there exist $a, b \in A$ such that $u(a) > u(b)$ but $P(a,b) = \frac{1}{2}$; there are utility differences that are not directly detectable from observation of choice probabilities. As noted in Ryan (2015), a Fechner representation (u, f) is strict iff $u(A)$ is a singleton or f is non-constant on any open neighbourhood of zero. It follows that a strong Fechner model is strict.

[9]The concept of a Fechner representation (or Fechner model), as we define it here, follows (*inter alia*) the terminology in Becker, DeGroot and Marschak (1963). However, some authors use slightly different definitions of a Fechner model; for example, by restricting the range of P to $(0, 1)$, as in Luce and Suppes (1965, Definition 17), or by requiring f to be strictly increasing, at least for points in the domain of f whose image is outside the set $\{0, 1\}$, as in Fishburn (1998, p.285) and Falmagne (1985, Definition 4.10).

[10]Our notion of a "strict Fechner representation" is that of Ryan (2015). It should not be confused with a "strict utility model" which is commonly used to refer to a Luce model for binary choice probabilities (Luce and Suppes, 1965, Definition 18).

We focus on strict Fechner representations. Blavatskyy (2008, Theorem 1) and Dagsvik (2008, Theorem 4) each give sufficient conditions for the existence of a strict Fechner representation with linear utility, the conditions of the former also being necessary (Ryan (2015)). To avoid lengthy strings of qualifiers, we henceforth use "stochastic" as a synonym for "strict Fechner" when defining specialisations of the strict Fechner model. For example, if A is a set of lotteries, we say that P has a *stochastic expected utility representation (SEUR)* if it has a strict Fechner representation (u, f) in which u has the EU form. We likewise say that P has a *stochastic implicit expected utility representation (SIEUR)* if it has a strict Fechner representation (u, f) in which u is of the IEU form.

2.1 Implicit Expected Utility

Let us briefly review implicit expected utility theory (Dekel (1986)).[11] Consider some preferences $\succsim \subseteq A \times A$ over lotteries. We assume throughout this section that $\succsim$ is a weak order (i.e., complete and transitive). From the $\succsim$-ordering of $\left\{\delta^1, ..., \delta^n\right\}$ we induce a weak order on X in the obvious fashion. We use $\succsim$ to denote this latter ordering also. Let $\overline{x}, \underline{x} \in X$ be such that $\overline{x} \succsim x_i \succsim \underline{x}$ for all $i \in \{1, 2, ..., n\}$. The function $u : A \to [0, 1]$ is an *implicit expected utility function* (or *implicit expected utility representation*) for $\succsim$ if there exists some $v : X \times [0, 1] \to [0, 1]$ that is continuous in its second argument with $v(\cdot, z)$ strictly increasing in the $\succsim$-ordering of X for any $z \in (0, 1)$,[12] such that $u(a)$ is the unique solution (in z) to

$$zv(\overline{x}, z) + (1 - z) v(\underline{x}, z) = \sum_{i=1}^{n} a_i v(x_i, z) \tag{8}$$

for any $a \in A$. The mapping $v(\cdot, u(a)) : X \to \mathbb{R}$ is the local Bernoulli utility function associated with the indifference class containing $a \in A$.

The representation is unique up to transformations of v of the form $\alpha(z) v(x, z) + \beta(z)$ for some continuous functions α and β with $\alpha(a) > 0$ for all z. In particular, if an implicit

[11] Dekel's theory is actually more general than the one presented here. It requires only that the outcome set X be a compact metric space and Dekel's representation theorem (*ibid.*, Proposition 1) applies to preferences over all simple probability measures on the Borel subsets of this space.

[12] Dekel's (1986) definition actually requires $v(\cdot, z)$ to be strictly increasing in the $\succsim$-ordering of X for any $z \in [0, 1]$. However, his axioms only entail the weaker property, as a careful reading of Dekel's argument on p.313 reveals. (See also Dekel's intuitive discussion of his proof on p.309.) In our Example 1 below, which is easily seen to satisfy all of Dekel's axioms, $v(\cdot, 0)$ and $v(\cdot, 1)$ are not strictly increasing in the $\succsim$-ordering of X. It would be impossible to represent the preferences in Example 1 in the form (9) if we require $v(\cdot, 0)$ and $v(\cdot, 1)$ to be strictly increasing in the $\succsim$-ordering of X and also require v to be continuous in its second argument.

expected utility (IEU) function exists, we can always find one with $v(\overline{x}, z) = 1 - v(\underline{x}, z) = 1$ for all $z \in [0,1]$, so that

$$u(a) = \sum_{i=1}^{n} a_i v(x_i, u(a)) \tag{9}$$

for all $a \in A$. Hence the "implicit expected utility" terminology.

The contours of an IEU function are linear but not necessarily parallel – the associated preferences satisfy a *betweenness* property (Axiom A4 in Dekel (1986)) but need not satisfy mixture independence.

Definition 1. *Preferences $\succsim \subseteq A \times A$ satisfy* ***betweenness*** *if $a \succ b$ (respectively, $a \sim b$) implies $a \succ a\lambda b \succ b$ (respectively, $a \sim a\lambda b \sim b$) for any $a, b \in A$ and any $\lambda \in (0,1)$.*

The following is well known but we give a proof in Appendix A for completeness. Note that we make use of the completeness of $\succsim$ in the proof, though transitivity is not needed.

Proposition 1. *The preferences $\succsim$ satisfy betweenness iff the following holds for any $e, f \in A$, any $a, b, c \in [e, f]$ and any $\lambda \in (0,1)$:*

$$a \succsim b \quad \Leftrightarrow \quad a\lambda c \succsim b\lambda c \tag{10}$$

For complete preferences, betweenness is therefore the requirement that mixture independence is satisfied by the restriction of $\succsim$ to any linear segment of the simplex.

Example 1. *Suppose $X = \{x_1, x_2, x_3\}$. Let $\succsim \subseteq A \times A$ satisfy*

$$\delta^3 \succ \delta^2 \succ \delta^1$$

and have indifference classes as illustrated in the Machina Triangle of Figure 1, where one lottery is preferred to another if the former lies on an indifference curve obtained by a clockwise rotation – about the point $(1,1)$ – of the indifference curve containing the latter. An IEU representation for $\succsim$ may be constructed as follows. Let $v(x_1, z) = 0$, $v(x_2, z) = 1 - z$, $v(x_3, z) = 1$ and

$$u(a) = \begin{cases} (1 + x(a))^{-1} & \text{if } a_1 < 1 \\ 0 & \text{if } a_1 = 1 \end{cases}$$

for all $a \in A$, where

$$x(a) = \frac{1 - a_3}{1 - a_1}$$

is the slope of the indifference curve through the point (a_1, a_3) in Figure 1. Since

$$u(a) \geq u(a') \quad \Leftrightarrow \quad x(a) \leq x(a')$$

it is obvious that u represents $\succsim$. To verify that u has the IEU form we check that $z = u(a)$ solves (8) for any a. This is obvious if $a_1 = 1$. If $a_1 < 1$ we have:

$$\begin{aligned} zv(x_3, z) + (1 - z)v(x_1, z) &= a_1 v(x_1, z) + a_2 v(x_2, z) + a_3 v(x_3, z) \\ \Leftrightarrow \quad z &= a_2 (1 - z) + a_3 \\ \Leftrightarrow \quad z &= \frac{a_2 + a_3}{1 + a_2} = \frac{1 - a_1}{1 + a_2} = \frac{1}{1 + x(a)} \end{aligned}$$

It follows that u is an IEU representation for $\succsim$.

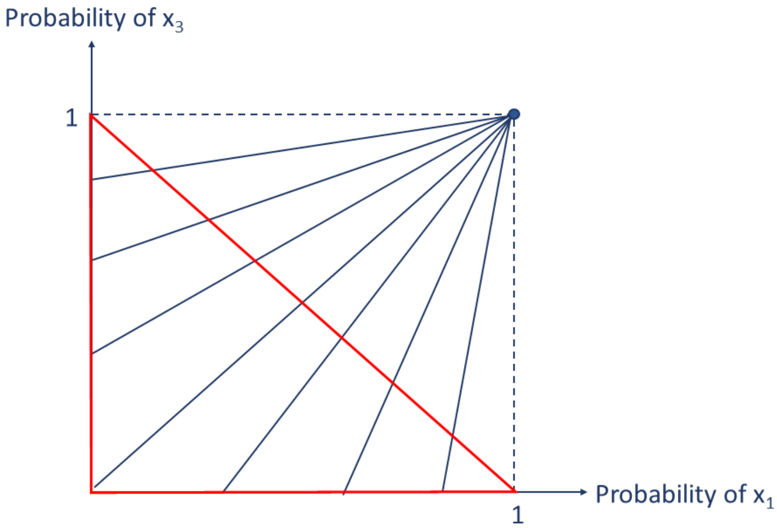

Figure 1: Preferences with an IEU representation

It is important to note that not all utility representations for the preferences in Example 1 are IEU functions, just as there exist non-linear representations for expected utility preferences. An IEU representation requires that u satisfy the restricted form of linearity embodied in (8). In particular, if u is an IEU representation for $\succsim$ and the elements of X are indexed such that $x_1 = \underline{x}$ and $x_n = \overline{x}$, with $\overline{x} \succ \underline{x}$, then we must have $u\left(\delta^n \lambda \delta^1\right) = \lambda$ for any $\lambda \in [0, 1]$, as is easily verified using (8).[13] In other words, if u is an IEU representation for $\succsim$, then $u(a)$ can be elicited as the value of λ that satisfies $a \sim \delta^n \lambda \delta^1$. (It follows that any IEU function is continuous.)

[13] See Dekel (1986, p.313).

2.2 Blavatskyy's axioms

When does a BCP possess a stochastic expected utility representation (SEUR), or a stochastic implicit expected utility representation (SIEUR)?

Both of these questions are addressed by Blavatskyy (2008). He shows that the following are necessary and sufficient for a SEUR:

Axiom 1 (**Strong Stochastic Transitivity** [**SST**])**.** *For all $a, b, c \in A$, if*

$$\min\{P(a,b)\,,\ P(b,c)\} \ \geq \ \frac{1}{2}$$

then

$$P(a,c) \ \geq \ \max\{P(a,b)\,,\ P(b,c)\}\,.$$

Axiom 2 (**Continuity**)**.** *For any $a, b, c \in A$ the sets*

$$\left\{\lambda \in [0,1] \;\middle|\; P(a\lambda b, c) \geq \frac{1}{2}\right\}$$

and

$$\left\{\lambda \in [0,1] \;\middle|\; P(a\lambda b, c) \leq \frac{1}{2}\right\}$$

are closed.

Axiom 3 (**Common Consequence Independence** [**CCI**])**.** *For any $a, b, c, d \in A$ and any $\lambda \in [0,1]$*

$$P(a\lambda c, b\lambda c) = P(a\lambda d, b\lambda d)\,.$$

Theorem 1. ***[Blavatskyy (2008, Theorem 1) as modified by Ryan (2015)]*** *Let P be a binary choice probability. Then P has a stochastic expected utility representation iff it satisfies Axioms 1-3.*

Blavatskyy further conjectures (*ibid.*, pp.1052-3) that P possess a SIEUR if (and only if) it satisfies Axioms 1-2 plus the following weakening of CCI:

Axiom 4 (**Stochastic Betweenness** [**SB**])**.** *For any $e, f \in A$ and any $\alpha, \beta, \gamma, \mu \in [0,1]$ with*

$$\alpha - \beta = \gamma - \mu \tag{11}$$

we have $P(e\alpha f, e\beta f) = P(e\gamma f, e\mu f)$.

To clarify the role of condition (11), observe that

$$(e\alpha f)-(e\beta f) \;=\; (\alpha-\beta)(e-f)$$

for any $\alpha,\beta\in[0,1]$ and any $e,f\in A$. Axiom 4 therefore says that, for any $e,f\in A$,

$$P(a,b)=P(a',b')$$

for any $a,b,a',b'\in[e,f]$ with

$$b-a=b'-a'.$$

In other words, if the intervals $[a,b]$ and $[a',b']$ are both contained in some (larger) interval $[e,f]$, and if $b-a=b'-a'$, then the probability of choosing a over b is the same as the probability of choosing a' over b'.

The fact that SB is weaker than CCI is not immediately obvious. However, it can be re-stated in an equivalent form which makes this relationship apparent:

Lemma 1. *Let P be a BCP. The P satisfies Axiom 4 iff it satisfies the following for any $e,f\in A$, any $a,b,c,d\in[e,f]$ and any $\lambda\in[0,1]$:*

$$P(a\lambda c,b\lambda c)=P(a\lambda d,b\lambda d) \tag{12}$$

Proof: Suppose P satisfies Axiom 4. Let $a,b,c,d\in[e,f]$ and $\lambda\in[0,1]$. Since

$$a\lambda b-b\lambda c=\lambda(a-b)=a\lambda d-b\lambda d$$

we have $P(a\lambda c,b\lambda c)=P(a\lambda d,b\lambda d)$.

Conversely, suppose P satisfies (12) for any $e,f\in A$, any $a,b,c,d\in[e,f]$ and any $\lambda\in[0,1]$. Let $\alpha,\beta,\gamma,\mu\in[0,1]$ with

$$\alpha-\beta=\gamma-\mu=k \tag{13}$$

It is without loss of generality (WLOG) to assume $k\geq 0$ and $\beta\leq\mu$. We must show that

$$P(e\alpha f,e\beta f) \;=\; P(e\gamma f,e\mu f) \tag{14}$$

If $k=0$ this is immediate: set $\lambda=0$, $c=e\alpha f$ and $d=e\gamma f$ in (12). If $\beta=\mu$, then (13) implies $\alpha=\gamma$ and (14) holds trivially. Therefore, suppose that $k>0$ and $\beta<\mu$. Since $\gamma\leq 1$, we have $k\leq 1-\mu$ from (13), so $[\beta+k,\beta+1-\mu]$ is a non-empty subset of $(0,1)$. Fix some $\lambda\in[\beta+k,\beta+1-\mu]$ and define

$$\eta^{a}=\frac{\beta+k}{\lambda}$$

$$\eta^b = \frac{\beta}{\lambda}$$
$$\eta^c = 0$$
$$\eta^d = \frac{\mu - \beta}{1 - \lambda}.$$

Then $\eta^x \in [0,1]$ for all $x \in \{a,b,c,d\}$ and it is easily verified that

$$\alpha = \lambda\eta^a + (1-\lambda)\,\eta^c$$

$$\beta = \lambda\eta^b + (1-\lambda)\,\eta^c$$

$$\gamma = \lambda\eta^a + (1-\lambda)\,\eta^d$$

$$\mu = \lambda\eta^b + (1-\lambda)\,\eta^d$$

Let $x = e(\eta^x) f$ for each $x \in \{a,b,c,d\}$. Then, using (12) we have:

$$P(e\alpha f, e\beta f) \;=\; P(a\lambda c, b\lambda c) \;=\; P(a\lambda d, b\lambda d) \;=\; P(e\gamma f, e\mu f).$$

$\square$

Stochastic Betweenness thus imposes the CCI condition (4) within linear segments of the simplex, just as the betweenness property of preferences imposes mixture-independence within linear segments (Proposition 1). This observation lends credence to Blavatskyy's conjecture. Nevertheless, the conjecture is false. In particular, SB is *not necessary* for a stochastic IEU representation.

3 Violating SB

Let u be the IEU function from Example 1 and define $f : \mathbb{R} \to \mathbb{R}$ as follows:

$$f(x) = \frac{1}{2}(1+x).$$

Note that f is *strictly* increasing, with $f([-1,1]) = [0,1]$, and satisfies $f(x)+f(-x) = 1$. Now construct P from u and f using (7). Then (u,f) is a strong Fechner representation for P by construction, and u has the IEU form. It is therefore a SIEUR for P. We will show that P violates stochastic betweenness.

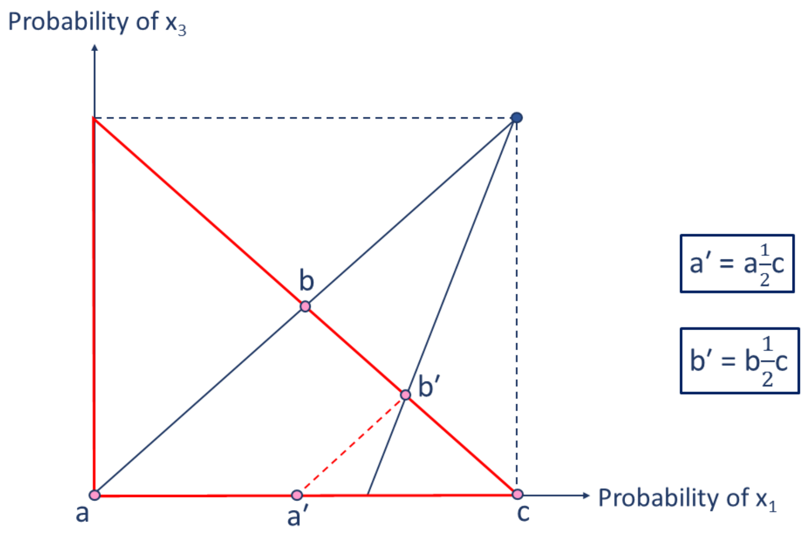

Figure 2: Constructing a violation of Stochastic Betweenness

Consider Figure 2. In this figure $b' = b\frac{1}{2}c$ and $a' = a\frac{1}{2}c$. Suppose, contrary to our claim, that P satisfies SB. Then we have:

$$\begin{aligned}
f\left(u\left(b\right)-u\left(b'\right)\right) &= P\left(b,b'\right) \\
&= P\left(b',c\right) && \text{(by SB)} \\
&= f\left(u\left(b'\right)-u\left(c\right)\right) \\
&< f\left(u\left(a'\right)-u\left(c\right)\right) && \text{(since } u\left(a'\right) > u\left(b'\right)\text{)} \\
&= P\left(a',c\right) \\
&= P\left(a,a'\right) && \text{(by SB)} \\
&= f\left(u\left(a\right)-u\left(a'\right)\right) \\
&= f\left(u\left(b\right)-u\left(a'\right)\right) && \text{(since } u\left(a\right) = u\left(b\right)\text{)} \\
&< f\left(u\left(b\right)-u\left(b'\right)\right) && \text{(since } u\left(a'\right) > u\left(b'\right)\text{)}
\end{aligned}$$

which is the desired contradiction.

In this example, the decision maker's choice probabilities have a stochastic IEU representation but it is impossible that both $P\left(a,a'\right) = P\left(a',c\right)$ and $P\left(b,b'\right) = P\left(b',c\right)$. Choice probabilities must violate SB.

It is clear that there is nothing special about this example. If $\hat{P}$ is a BCP and $\left(\hat{u},\hat{f}\right)$ is a SIEUR for P, then we can construct a similar violation of SB provided $\hat{f}$ is strictly increasing and $\succsim^{\hat{P}}$ does not have an EU representation.

In short, SB is not necessary for the existence of a SIEUR. The next section provides a set of necessary and sufficient conditions for P to possess a SIEUR.

4 A representation theorem

Consider the following pair of axioms:

Axiom 5. *For any $a, b \in A$ and any $\lambda \in (0,1)$,*

$$P(a, b\lambda a) \geq \frac{1}{2} \quad \Leftrightarrow \quad P(a\lambda b, b) \geq \frac{1}{2}$$

and

$$P(a, b\lambda a) \leq \frac{1}{2} \quad \Leftrightarrow \quad P(a\lambda b, b) \leq \frac{1}{2}$$

Axiom 6. *If $\underline{\delta}, \overline{\delta} \in \{\delta^1, ..., \delta^n\}$ are such that $\delta^i \succsim^P \underline{\delta}$ and $\overline{\delta} \succsim^P \delta^i$ for all i, then*

$$P\left(\overline{\delta}\alpha\underline{\delta}, \overline{\delta}\beta\underline{\delta}\right) = P\left(\overline{\delta}\lambda\underline{\delta}, \overline{\delta}\mu\underline{\delta}\right)$$

for any $\alpha, \beta, \lambda, \mu \in [0,1]$ with $\alpha - \beta = \lambda - \mu$.

Axiom 5 says that whenever mixing a with b produces a lottery that is less (respectively, more) stochastically desirable than a, then the complementary mixing of a with b produces a lottery that is more (respectively, less) stochastically desirable than b. Together with Axioms 1-2, Axiom 5 implies that $\succsim^P$ satisfies Dekel's (1986) Betweenness axiom – see Lemma 4 in Appendix B.

Axiom 6 says that if $\underline{\delta}$ and $\overline{\delta}$ are $\succsim^P$-worst and $\succsim^P$-best (respectively) amongst the degenerate lotteries (i.e., vertices of the simplex), and if $[a, b]$ and $[c, d]$ are sub-intervals of $\left[\underline{\delta}, \overline{\delta}\right]$ with $a - b = c - d$, then the probability of choosing a over b is the same as the probability of choosing c over d. Note that if P satisfies Axiom 1 (SST), then $\succsim^P$ is a weak order so $\succsim^P$-worst and $\succsim^P$-best vertices will exist.

Axioms 5 and 6 are both implied by stochastic betweenness (Axiom 4). This is obvious in the case of Axiom 6. To see that Axiom 5 is also weaker than SB, note that the latter implies

$$P(a\lambda a, b\lambda a) = P(a\lambda b, b\lambda b)$$

for any $a, b \in A$.

Axioms 1-2, 5 and 6 do not yet suffice for a SIEUR, since they permit $\succsim^P$ to violate FOSD-monotonicity.[14]

[14] They *do* suffice for the existence of a strict Fechner model with utility of the form described by Dekel's (1986) Proposition A.1. This is easily shown by adapting the proof of Theorem 2.

Example 2. *Suppose $X = \{x_1, x_2, x_3\}$. Let $\succsim \subseteq A \times A$ satisfy*

$$\delta^3 \succ \delta^2 \succ \delta^1$$

and have indifference classes as illustrated in the Machina Triangle of Figure 3, where one lottery is preferred to another if the former lies on an indifference curve obtained by a clockwise rotation – about the point $\left(\frac{2}{3}, \frac{2}{3}\right)$ – of the indifference curve containing the latter.

These preferences clearly satisfy betweenness. Define $u : A \to [0,1]$ such that $u\left(\delta^3 \lambda \delta^1\right) = \lambda$ for each $\lambda \in [0,1]$. Then u represesents the preferences $\succsim$. Now let P be the BCP obtained by combining u with some strictly increasing f. Then Axioms 1-2, 5 and 6 will all be satisfied, as is easily verified. However, the illustrated preferences violate Dekel's (1986) Monotonicity axiom, since $\delta^2 \succ \delta^1$ and $b \succ a$. It follows that they do not have an IEU representation.

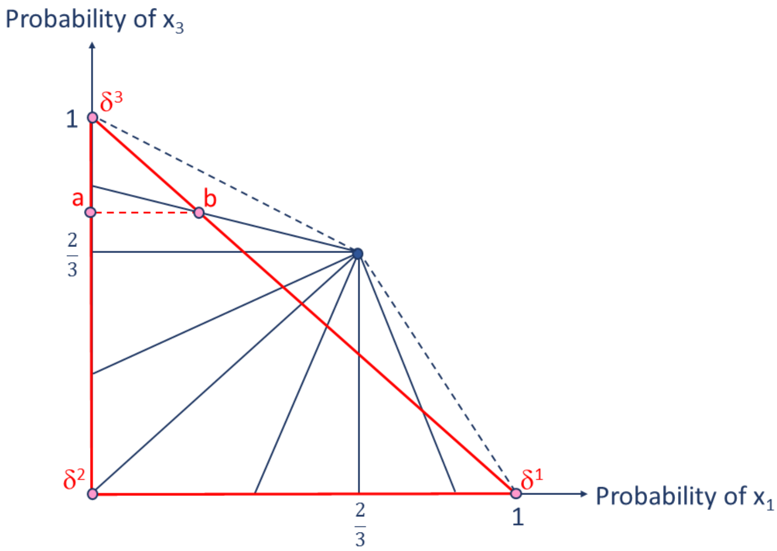

Figure 3: Preferences that violate FOSD-dominance

To avoid this problem we add the following stochastic version of the Monotonicity axiom from Dekel (1986).

Axiom 7 (Stochastic Monotonicity)**.** *If $\underline{\delta}, \overline{\delta} \in \left\{\delta^1, ..., \delta^n\right\}$ are such that $\delta^i \succsim^P \underline{\delta}$ and $\overline{\delta} \succsim^P \delta^i$ for all i, then for any $\boldsymbol{\delta} \in \left\{\underline{\delta}, \overline{\delta}\right\}$, any $\lambda \in (0,1)$ and any $\delta', \delta'' \in \left\{\delta^1, ..., \delta^n\right\}$*

$$P\left(\delta', \delta''\right) > \frac{1}{2} \quad \Rightarrow \quad P\left(\boldsymbol{\delta}\lambda\delta', \boldsymbol{\delta}\lambda\delta''\right) > \frac{1}{2}$$

and

$$P(\delta', \delta'') = \frac{1}{2} \quad \Rightarrow \quad P(\boldsymbol{\delta}\lambda\delta', \boldsymbol{\delta}\lambda\delta'') = \frac{1}{2}$$

Theorem 2. *Let P be a BCP. Then P has a SIEUR iff it satisfies Axioms 1-2 and 5-7.*

Theorem 2 is proved in Appendix B.

At this point, the reader may be wondering why Axiom 5 does not create problems like those illustrated in Section 3. Axiom 5 imposes the CCI condition (4) along the edge $[\underline{\delta}, \overline{\delta}]$ of the simplex A. Suppose P has a SIEUR (u, f) and let P^* be the restriction of P to $\Delta \times \Delta$, where $\Delta \subseteq A$ is a sub-simplex. It follows that (u^*, f) is a strict Fechner representation for P^* when u^* is the restriction of u to Δ. Provided u^* is an IEU representation for $\succsim^{P^*}$, Theorem 2 then implies that P^* satisfies the CCI condition along an edge of Δ joining "$\succsim^{P^*}$-best" and "$\succsim^{P^*}$-worst" vertices. Since this is also an edge of A, we may interate this logic to infer that the CCI condition holds along *any* edge of A, thereby creating the potential for the contradiction illustrated using Figure 2. However, to evade this apparent contradiction it suffices to observe that u^* will *not*, in general, be an IEU representation for $\succsim^{P^*}$ (though it will, of course, represent these preferences). For an IEU representation, the utility along a "$\succsim^{P^*}$-best-worst" edge must coincide with the weight on the "$\succsim^{P^*}$-best" vertex – recall the discussion following Example 1. This will certainly be the case if all $\succsim^{P}$-indifference surfaces are *parallel*, but may not be so otherwise.

5 Concluding remarks

Blavatskyy (2008) provides an important axiomatisation of the "EU plus Fechnerian noise" model. He further conjectures that weakening CCI to SB will provide an axiomatisation of the more general "IEU plus Fechnerian noise" model. As we have shown, this conjecture is false. A valid axiomatisation of the "IEU plus Fechnerian noise" model is obtained by replacing SB with the (regrettably, less elegant) triumvirate of Axioms 5-7 (Theorem 2).

Blavatskyy (2006) argues that there is *prima facie* evidence to suggest that much experimental evidence against the betweenness is nevertheless compatible with betweenness-satisfying preferences that are expressed with Fechnerian noise. This hypothesis has, to the best of our knowledge, yet to be formally tested. One of the attractive features of SB is that it is readily testable along the lines of Loomes and Sugden (1998), though we are not aware of such tests having been conducted. Unfortunately, this test is too strict. As we have shown, evidence that subjects violate SB should *not* be counted as evidence against the "IEU plus Fechnerian noise" model. The special cases of SB described by Axioms 5 and 6 impose significantly milder restrictions on binary choice probabilities.

Appendices

These Appendices contain proofs omitted from the text.

A Proof of Proposition 1

Proposition 1 is straightforward corollary of the following:

Lemma 2. *Suppose $\succsim \subseteq A \times A$ is complete and satisfies betweenness. If $e, f \in A$ and $\hat{e}, \hat{f} \in [e, f]$ are such that $\hat{e} - \hat{f} = k(e - f)$ for some $k > 0$, then $\hat{e} \succsim \hat{f}$ iff $e \succsim f$.*

Proof: If $e = f$ the result is trivial so assume otherwise. Let $\alpha, \beta \in [0, 1]$ be such that $\hat{e} = e\alpha f$ and $\hat{f} = e\beta f$. It follows that $\alpha > \beta$, since $\hat{e} - \hat{f} = (\alpha - \beta)(e - f)$. In other words, $\hat{e} \in \left[e, \hat{f}\right)$ and $\hat{f} \in (\hat{e}, f]$.

Suppose $e \succsim f$. Then betweenness implies $e \succsim \hat{e} \succsim f$. Applying betweenness to the preference $\hat{e} \succsim f$, we deduce $\hat{e} \succsim \hat{f}$.

For the converse, we invoke completeness and prove the contrapositive. Therefore, let us suppose that $f \succ e$. If $\hat{f} = f$ then we have $\hat{f} \succ e$; otherwise, the same conclusion follows from betweenness, since $\hat{f} \in (\hat{e}, f]$. If $\hat{e} = e$ we have $\hat{f} \succ \hat{e}$ as required; otherwise, apply betweenness and the fact that $\hat{e} \in \left[e, \hat{f}\right)$ to reach the same conclusion. $\square$

Let $\succsim \subseteq A \times A$ be complete and satisfy betweenness. Suppose $e, f \in A$ and $a, b, c \in [e, f]$. If $e = f$ or $a = b$ then (10) is trivial. If $e \neq f$, $a \neq b$ and $\lambda \in (0, 1)$ then $a\lambda c, b\lambda c \in [e, f]$ and

$$a\lambda c - b\lambda c \;=\; \lambda(a - b).$$

Hence (10) follows by Lemma 2.

For the converse, let $a, b \in A$ and suppose $a \succ b$ (respectively $a \sim b$). By considering $c = a$ and $c = b$ in (10) we deduce that $a \succ a\lambda b \succ b$ (respectively $a \sim a\lambda b \sim b$) for any $\lambda \in (0, 1)$. (Note that $a, b, c \in [a, b]$ in this argument.)

This completes the proof of Proposition 1.

B Proof of Theorem 2

We will need the following useful result:

Lemma 3 (Davidson and Marschak, 1959)**.** *Let P be a BCP. Then P satisfies SST (Axiom 1) iff*

$$P(a,b) \geq \frac{1}{2} \quad \Rightarrow \quad P(a,c) \geq P(b,c) \tag{15}$$

for any $a, b, c \in A$.

Condition (15) is called the *weak substitutability* property.

Suppose that P satisfies Axioms 1-2 and 5-7. We start by showing that $\succsim^P$ has an IEU representation.

Axioms 1 ensures that $\succsim^P$ is a weak order. Since X is finite, $\succsim^P$ must satisfy Dekel's (1986) Axiom A1. We assume (WLOG) that $\delta^n \succsim^P \delta^{n-1} \succsim^P \cdots \succsim^P \delta^1$. If $\delta_1 \sim \delta_n$ the result is trivial, so we further assume that $\delta^n \succ^P \delta^1$. Our Axiom 7, and the fact that $\left\{\delta^1, ..., \delta^n\right\}$ is weakly ordered by $\succsim^P$, implies that $\succsim^P$ also satisfies Dekel's Axiom A3. The following two lemmata establish that $\succsim^P$ satisfies Dekel's (1986) Axioms A4 and A2 respectively.

Lemma 4. *For any $a, b \in A$ and any $\lambda, \mu \in [0,1]$ with $\lambda > \mu$,*

$$a \sim^P b \quad \Rightarrow \quad a\lambda b \sim^P b \tag{16}$$

and

$$a \succ^P b \quad \Rightarrow \quad a\lambda b \succ^P a\mu b \tag{17}$$

Proof. Suppose $a \sim^P b$. The following argument proves that $a \sim^P a\left(\frac{1}{2}\right) b$.

If $a \succ^P a\left(\frac{1}{2}\right) b$, then Axiom 5 gives $a\left(\frac{1}{2}\right) b \succ^P b$, so $a \succ^P b$ by transitivity of $\succsim^P$. This contradicts $a \sim^P b$. If $a\left(\frac{1}{2}\right) b \succ^P a$, then $b \succ^P a$ by a similar argument, which also contradicts $a \sim^P b$. Hence, $a \sim^P a\left(\frac{1}{2}\right) b$ by completeness of $\succsim^P$.

We may iterate this logic by continuing to subdivide the segment $[a,b]$. Thus, $a \sim^P b$ implies $a \sim^P a\lambda b$ for any *dyadic fraction* λ (i.e., any λ of the form $k/2^n$ for some $n \in \{1, 2, ...\}$ and some $k \in \{0, 1, ..., 2^n\}$). From Axiom 2 we know that the sets

$$\left\{\lambda \in [0,1] \;\middle|\; a\lambda b \succ^P a\right\}$$

and

$$\left\{\lambda \in [0,1] \;\middle|\; a \succ^P a\lambda b\right\}$$

are open. It follows that each set is empty. This proves (16).

We now prove (17). If $a \succ^P b$ then, by an argument similar to the one above, we can use Axiom 5 and SST to rule out the possibility that

$$b \succsim^P a\left(\frac{1}{2}\right)b$$

or

$$a\left(\frac{1}{2}\right)b \succsim^P a.$$

Hence:

$$a \succ^P a\left(\frac{1}{2}\right)b \succ^P b.$$

By iteration we have $a\lambda b \succ^P a\mu b$ for any dyadic fractions $\lambda, \mu \in [0,1]$ with $\lambda > \mu$. The following argument extends this to any $\lambda, \mu \in [0,1]$ with $\lambda > \mu$.

Let $\lambda, \mu \in [0,1]$ with $\lambda > \mu$. Since the dyadic fractions are dense in $[0,1]$, we may obtain λ as the limit of a sequence $\{x^m\}_{m=1}^{\infty} \subseteq ((\lambda+\mu)/2, 1)$ of dyadic fractions, and likewise obtain μ as the limit of a sequence $\{y^s\}_{m=1}^{\infty} \subseteq (0, (\lambda+\mu)/2)$ of dyadic fractions. Then

$$a(x^m)\,b \succ^P a(y^s)\,b$$

for all m and all s. By Axiom 2, $a\lambda b \succsim^P a(y^s)\,b$ for all s, and hence, applying Axiom 2 once more, $a\lambda b \succsim^P a\mu b$. If $a\lambda b \sim^P a\mu b$ then $a\lambda b \sim^P a\gamma b$ for any $\gamma \in [\mu, \lambda]$ by (16). But this is impossible, since we can find two distinct dyadic fractions in $[\mu, \lambda]$. Hence $a\lambda b \succ^P a\mu b$. □

Lemma 5. *If $a, b \in A$ with $a \succ^P b$, then for any $c \in A$ such that $a \succsim^P c \succsim^P b$ there exists a unique $\alpha \in [0,1]$ such that $c \sim^P a\alpha b$.*

Proof. Since $a \succ^P b$, (17) implies that

$$a \succsim^P a\lambda b \succsim^P b$$

for any $\lambda \in [0,1]$. By standard arguments, Axioms 1-2 imply that the set

$$S \;=\; \left\{\lambda \in [0,1] \;\middle|\; a\lambda b \succsim^P c\right\} \cap \left\{\lambda \in [0,1] \;\middle|\; a\lambda b \precsim^P c\right\}$$

is closed and non-empty. Lemma 4 implies that S must be a singleton. □

We have therefore shown that $\succsim^P$ satisfies Axioms A1-A4 of Dekel (1986). It follows that there exists an IEU representation for $\succsim^P$ with $v(x_1, z) = 1$ and $v(x_n, z) = 0$ for all

$z \in [0,1]$ (Dekel, 1986, Proposition 1). In particular, if $a \in A$ with $u(a) = u(\delta^1 \alpha \delta^n)$ then $u(a) = \alpha$. Moreoever, by Lemma 5, for every $a \in A$ there is a unique $\alpha \in [0,1]$ satisfying $u(a) = u(\delta^1 \alpha \delta^n)$.

We next use u to construct a suitable Fechner representation for P.

For any $a, b, c \in A$ we have

$$u(a) = u(b) \quad \Leftrightarrow \quad P(a,b) = \frac{1}{2} \quad \Rightarrow \quad P(a,c) = P(b,c) \quad \Leftrightarrow \quad P(c,a) = P(c,b)$$

where the first equivalence uses the fact that u represents $\succsim^P$, the middle implication uses weak substitutability (Lemma 3), and the final equivalence uses completeness of P. It follows that P is scalable: there exists a function $\pi : [0,1]^2 \to [0,1]$ such that $P(a,b) = \pi(u(a), u(b))$ for any $a, b \in A$. Weak substitutability and completeness further imply that π is non-decreasing in its first argument, non-increasing in its second and satisfies $\pi(x,y) = 1 - \pi(y,x)$.

We claim that $\pi(x,y)$ depends only on $x - y$. Suppose $x - y = x' - y'$. Let

$$a = x\delta_1 + (1-x)\,\delta_n$$

$$b = y\delta_1 + (1-y)\,\delta_n$$

$$a' = x'\delta_1 + (1-x')\,\delta_n$$

$$b' = y'\delta_1 + (1-y')\,\delta_n$$

so that $\pi(x,y) = P(a,b)$ and $\pi(x',y') = P(a',b')$. Axiom 6 implies that $P(a,b) = P(a',b')$ as required.

Thus, we may define $f : [-1,1] \to [0,1]$ by setting $f(k) = \pi(x,y)$ for any $(x,y) \in [0,1]^2$ with $x - y = k$. Note that f is non-decreasing in k since π is non-decreasing in its first argument and non-increasing in its second. Now extend f to $\mathbb{R}$ in any fashion that ensures f is non-decreasing and satisfies $f(x) + f(-x) = 1$. Then (u, f) is a SIEUR for P.

To prove the converse part of the Theorem, suppose that (u, f) is a SIEUR for P.

To see that P satisfies Axiom 1 (SST), we use the facts that u represents $\succsim^P$ and f is increasing: the former ensures $u(a) \geq u(b)$ whenever $P(a,b) \geq \frac{1}{2}$, and the latter implies

$$f(x+y) \geq \max\{f(x), f(y)\}$$

for all $x \geq 0$ and $y \geq 0$. To verify Axiom 2 we use the fact that u represents $\succsim^P$ to deduce

$$P(a\lambda b, c) \geq \frac{1}{2} \quad \Leftrightarrow \quad u(a\lambda b) \geq u(c)$$

and

$$P(a\lambda b, c) \leq \frac{1}{2} \quad \Leftrightarrow \quad u(a\lambda b) \leq u(c).$$

Axiom 2 therefore follows from the continuity of u. We next verify Axioms 5 and 7. Since u is an IEU representation for $\succsim^P$ it follows that $\succsim^P$ satisfies Dekel's Axioms A1-A4 (Dekel, 1986, Proposition 1): Axiom 7 follows directly from Dekel's Axiom A3 (and the fact that $\{\delta^1, ..., \delta^n\}$ is weak ordered by $\succsim^P$) while Axiom 5 is implied by the completeness of $\succsim^P$ and Dekel's Axiom A4 (Betweenness). Finally, we deduce Axiom 6 from the Fechner representation and the fact that $u(\overline{\delta}\alpha\underline{\delta}) = \alpha$ for any $\alpha \in [0, 1]$ – recall the discussion following Example 1.

This completes the proof of Theorem 2.

References

Becker, G. M., M. H. DeGroot and J. Marschak (1963), "Stochastic Models of Choice Behavior," *Behavioral Science* **8(1)**, 41–55.

Blavatskyy, P.R. (2006), "Violations of Betweenness or Random Errors?," *Economics Letters* **91(1)**, 34–38.

Blavatskyy, P.R. (2008), "Stochastic Utility Theorem," *Journal of Mathematical Economics* **44**, 1049–1056.

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(4)**, 1065-1092.

Dagsvik, J.K. (2008), "Axiomatization of Stochastic Models for Choice Under Uncertainty," *Mathematical Social Sciences* **55**, 341–370.

Dagsvik, J.K. (2015), "Stochastic Models for Risky Choice: A Comparison of Different Axiomatizations," *Journal of Mathematical Economics* **60**, 81–88.

Davidson, D. and J. Marschak. (1959), "Experimental Tests of a Stochastic Decision Theory," in C.W. Churchman and P. Ratoosh (eds), *Measurement: Definitions and Theories*. John Wiley and Sons: New York, NY.

Dekel, E. (1986), "An Axiomatic Characterization of Preferences under Uncertainty: Weakening the Independence Axiom," *Journal of Economic Theory* **40(2)**, 304-318.

Falmagne, J.C. (2002), *Elements of Psychophysical Theory*. Oxford University Press.

Fishburn, P.J. (1998), "Stochastic Utility," in S. Barberà, P.J. Hammond and C. Seidl (eds), *Handbook of Utility Theory.* Kluwer Academic: Dordrecht.

Hey, J.D. (2014), "Choice under Uncertainty: Empirical Methods and Experimental Results," in: M. J. Machina and W. Kip Viscusi (eds) *Handbook of the Economics of Risk and Uncertainty*, Vol 1, Oxford: North Holland, pp. 809-850.

Loomes, G. (2005), "Modelling the Stochastic Component of Behaviour in Experiments: Some Issues for the Interpretation of Data," *Experimental Economics* **8(4)**, 301-323.

Loomes, G. and R. Sugden (1998), "Testing Different Stochastic Specifications of Risky Choice," *Economica* **65**, 581–598.

Luce, R. and P. Suppes (1965), "Preference, Utility and Subjective Probability," in R.D. Luce, R.B. Bush and E. Galanter (eds), *Handbook of Mathematical Psychology, Volume III.* John Wiley and Sons: New York.

Mosteller, F. and P. Nogee (1951), "An Experimental Measurement of Utility," *Journal of Political Economy* **59(5)**, 371–404.

Ryan, M.J. (2015), "A Strict Stochastic Utility Theorem," *Economics Bulletin* **35(4)**, 2664-2672.