Random Binary Choices that Satisfy Stochastic Betweenness*

Matthew Ryan
School of Economics
Auckland University of Technology

April 2017

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 we call *Stochastic Betweenness* – 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.

*An abbreviated version of this paper is forthcoming in the *Journal of Mathematical Economics*. I'm grateful to Pavlo Blavatskyy, Simon Grant and two anonymous referees 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) notes 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.

Decision theorists have taken two different approaches to accommodating such evidence. One is to treat it as evidence of a *preference to randomise*. This approach is advocated most famously by Machina (1985). Given two lotteries from which to choose, an individual whose lottery preferences are represented by a utility function that is "quasiconcave in probabilities" may strictly prefer a particular mixture of the two lotteries over either lottery individually, or indeed over any other mixture of the two (*ibid.*, Figure 2). The other approach, which is typical in mathematical psychology, is to suppose that an individual's preferences over lotteries are revealed by choices in a noisy fashion. Randomness is a manifestation of the psychophysical process that intervenes between "preference" and choice.

One might summarise these two approaches by saying that the first treats choice as "purposely noisy" while the second treats choice as "noisily purposeful". The latter approach is dominant in the literature – especially the experimental literature – and will also be taken here. (We make further comments on this modelling choice later in the paper.)

Since the mid-1990s, experimental evidence on expected utility (EU) has been increasingly viewed through the lens of these psychophysical models of probabilistic 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 binary choice problem, then P has a Fechner model if there is a utility function $u : A \to \mathbb{R}$ such that $P(a, b)$ is a non-decreasing function of $u(a) - u(b)$, with $P(a, b) = \frac{1}{2}$ whenever $u(a) = u(b)$.

It is natural to interpret u (in a Fechner model) as a representation of the decision-maker's "preferences". These preferences are expressed through choice but in a noisy fashion. If $u(a) > u(b)$, then the decision maker chooses a over b with probability at least

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

$\frac{1}{2}$. This probability is non-decreasing in $u(a) - u(b)$, so the greater the utility difference the more likely it is that the "more preferred" alternative is chosen. A Fechner model therefore describes a process of *noisy utility maximisation*. If the alternatives in A are lotteries, and u has the expected utility form, then we have a model of noisy *expected utility* maximisation – a model of random binary choice guided by EU preferences, or "EU with Fechnerian noise". Blavatskyy (2008, Theorem 1) and Dagsvik (2008, Theorem 4) provide axiomatic foundations for this model – conditions on P that are sufficient for the existence of a Fechner representation with u of the EU form (and necessary under some additional restrictions on the representation).[2] These two representation theorems are important benchmarks in the literature on binary stochastic choice. We discuss Blavatskyy's theorem in Section 2.2.

One can likewise imagine Fechner models in which u is *not* linear in probabilities (i.e., u is not within the EU class). For example, P might admit a Fechner model in which u has the rank-dependent expected utility form – a model of "rank-dependent expected utility with Fechnerian noise". A substantial body of experimental literature studies the descriptive merits of these various Fechnerian models of lottery choice.[3] Currently, however, axiomatic foundations are lacking for Fechner models in which u is restricted to some class of functions broader – or different – than the set of EU functions.

A rare exception is the *implicit expected utility (IEU)* class, which was introduced by Dekel (1986). Preferences described by an IEU function satisfy the *betweenness* property (Chew (1983)) but need not satisfy the more familiar *mixture independence* condition. Betweenness is, in fact, a restricted form mixture independence (see Proposition 1). It requires that the mixture independence condition holds only when all of the lotteries appearing in the condition are *collinear* – that is, they all sit on a common line when visualised in the probability simplex.

Given this fact, Blavatskyy (2008, pp.1052-3) conjectures that an axiomatisation of "IEU with Fechnerian noise" – necessary and sufficient conditions for P to possess a Fechner representation with u of the IEU form – can be obtained by a similar modification of his axioms for the "EU with Fechnerian noise model". Blavatskyy's conjectured axiomatisation involves restricting his stochastic analogue of the mixture independence property, which he calls *Common Consequence Independence (CCI)*, to collinear lotteries. We call this restricted form of CCI, *Stochastic Betweenness*.[4]

[2] See Section 2 for more precise statements of Blavatskyy's and Dagsvik's results.

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

[4] Blavatskyy (2008) uses the term *Betweenness*, but this invites confusion with the betweenness prop-

Despite its plausibility, Blavatskyy's conjecture is false. As we show in Section 3, there exist Fechner models with u of the IEU form that violate Stochastic Betweenness. In this sense, the Stochastic Betweenness axiom is too strong for the desired representation – Stochastic Betweenness is not necessary for P to be represented as "IEU with Fechnerian noise". In particular, the Stochastic Betweenness axiom cannot be used to test this model. In Section 4 we present a representation theorem for the "IEU with Fechnerian noise" model by suitably weakening Stochastic Betweenness (Theorem 3).

In preparation for our main results, we next recall some basic concepts from the theory of stochastic binary choice between lotteries, then review Dekel's (1986) implicit expected utility theory (Section 2.1) and Blavatskyy's (2008) conjecture (Section 2.2).

2 Stochastic IEU

We adopt the framework of Blavatskyy (2008) and Dagsvik (2008). Let A be the unit simplex in $\mathbb{R}^n$. Points in A will be interpreted as *lotteries* over a given set $X = \{x_1, ..., x_n\}$ of (pure) outcomes. If $a \in A$ then a_i is the probability with which lottery a delivers the outcome x_i. Hence $a_i \in [0, 1]$ for each $i \in \{1, 2, ..., n\}$ and $\sum_{i=1}^{n} a_i = 1$. We use $\delta^i \in A$ to denote the lottery that delivers outcome x_i with certainty. That is: $\delta^i_i = 1$ and $\delta^i_j = 0$ for any $j \neq i$. Following standard convention, if $a, b \in A$ and $\lambda \in [0, 1]$ then we use $a\lambda b$ to denote the convex combination $\lambda a + (1 - \lambda) b$.

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 which describe his or her stochastic choice function over binary choice (or "budget") sets. A *binary choice probability* is a mapping $P : A \times A \to [0, 1]$. 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 from the binary choice set $\{a, b\}$.[5] No behavioural interpretation is given to $P(a, b)$ when $a = b$, but it is conventional to define binary choice probabilities on the entire Cartesian product $A \times A$ for convenience. It is also conventional to set $P(a, a) = \frac{1}{2}$ for all $a \in A$.

---

erty of preference relations.

[5]In Machina's (1985) notation, if C denotes the decision-maker's stochastic choice function over sets of lotteries, then

$$C(\{a, b\}) = (P(a, b), P(b, a)).$$

Machina's definition of a stochastic choice function also requires that $P(b, a) = 1 - P(a, b)$. This is the *Completeness* condition on P – see Axiom 1 in Section 2.2.

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

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

The binary relations $\succ^P$ and $\sim^P$ are determined from $\succsim^P$ in the usual way. We call $\succsim^P$ the decision-maker's *weak stochastic preference relation*. Thus, a is weakly stochastically preferred to b if the probability of choosing a from the binary choice set $\{a,b\}$ is at least $\frac{1}{2}$. This is a slight abuse of terminology, since $\succsim^P$ need not be transitive without further restrictions on P, though it is complete by construction. However, we will later introduce an axiom that ensures the transitivity of $\succsim^P$ (Axiom 2).

Blavatskyy (2008) and Dagsvik (2008) study the conditions under which P possesses a particular type of *Fechner representation*.

Definition 1. *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*

$$P(a,b) = F(u(a) - u(b)) \tag{2}$$

for any $a, b \in A$.[6]

Definition 2. *A* strict Fechner representation *for P is a Fechner representation (u,F) such that*

$$u(a) \geq u(b) \quad \Leftrightarrow \quad P(a,b) \geq \frac{1}{2} \tag{3}$$

for any $a, b \in A$.

The notion of a strict Fechner representation is due to Ryan (2015).[7] Condition (3) requires that u represents – in the usual sense – the weak stochastic preference relation $\succsim^P$. If (u,F) is a Fechner representation for P, then $F(0) = \frac{1}{2}$ (since $F(x) + F(-x) = 1$ for all $x \in \mathbb{R}$) and hence

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

[6]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 (2002, Definition 4.10).

[7]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).

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.

Definition 3. *A* strong Fechner representation *for P is a Fechner representation (u, F) such that F is strictly increasing on $\{u(a) - u(b) \mid a, b \in A\} \subseteq \mathbb{R}$.*

Ryan (2015) shows that 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 any strong Fechner representation is strict.

Blavatskyy (2008, Theorem 1) gives necessary and sufficient conditions for P to possess a strict Fechner representation (u, F) in which $u : A \to \mathbb{R}$ is linear (i.e., an EU function).[8] Dagsvik (2008, Theorem 4) gives an alternative set of axioms, which are necessary and sufficient for a strong Fechner representation (u, F) in which u is linear and F is continuous.

We focus on strict Fechner representations in the present paper. To avoid lengthy strings of qualifiers, we henceforth use "stochastic" as a synonym for "strict Fechner" when referring to representations of binary choice probabilities. For example, we say that P has a *stochastic EU representation* if it has a strict Fechner representation (u, F) in which u is linear. We likewise say that P has a *stochastic IEU representation* if it has a strict Fechner representation (u, F) in which u is an implicit expected utility function. The next subsection describes this latter class of functions.

2.1 Implicit Expected Utility

Let us briefly review Dekel's (1986) implicit expected utility theory.[9]

Consider a preference relation, $\succsim$, over the lotteries in A. We assume throughout this subsection that $\succsim$ is complete and transitive (i.e., a *weak order* on A). Define $\succ$ and $\sim$ from $\succsim$ in the usual way. The $\succsim$-ordering of $\{\delta^1, ..., \delta^n\} \subseteq A$ induces a weak order on X in the obvious fashion. (Recall that $\delta^i \in A$ is the degenerate lottery that gives outcome x_i

[8] In fact, Blavatskyy's theorem states that his axioms are necessary and sufficient for a Fechner representation with linear utility. As shown in Ryan (2015), they are in fact necessary and sufficient for a *strict* Fechner representation.

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

with certainty.) We use $\succsim$ to denote this latter weak ordering also. Finally, let $\overline{x}, \underline{x} \in X$ be such that $\overline{x} \succsim x_i \succsim \underline{x}$ for all $i \in \{1, 2, ..., n\}$. That is, outcome $\overline{x}$ is weakly preferred to any other outcome in X, and any outcome in X is weakly preferred to $\underline{x}$.

Definition 4 (**Dekel (1986)**). *The function $u : A \to [0,1]$ is an* implicit expected utility function *(or* implicit expected utility representation) *for $\succsim$ if there exists a function*

$$v : X \times [0,1] \to [0,1]$$

that is continuous in its second argument, with $v(\cdot, z)$ strictly increasing in the $\succsim$-ordering of X for any $z \in (0,1)$,[10] *such that $u(a)$ is the unique solution (in z) to*

$$zv(\overline{x}, z) + (1 - z)\, v(\underline{x}, z) = \sum_{i=1}^{n} a_i v(x_i, z) \tag{4}$$

for any $a \in A$.

To understand this definition, suppose that $v(x, z)$ is constant in z for any $x \in X$, and let $w(x) = v(x, z)$ for each x. Then $u(a)$ satisfies

$$u(a)\, w(\overline{x}) + (1 - u(a))\, w(\underline{x}) = \sum_{i=1}^{n} a_i w(x_i)$$

$$\Leftrightarrow \quad u(a) = \sum_{i=1}^{n} a_i \hat{w}(x_i)$$

where

$$\hat{w}(x_i) = \frac{w(x_i) - w(\underline{x})}{w(\overline{x}) - w(\underline{x})}$$

and we have used the fact that $\sum_{i=1}^{n} a_i = 1$. Since a_i is the probability that lottery $a \in A$ delivers outcome $x_i \in X$, we see that u has the expected utility form with associated Bernoulli utility function $\hat{w} : X \to \mathbb{R}$. When $v(x, z)$ is *not* constant in z for every $x \in X$, the mapping $\hat{v}(\cdot, u(a)) : X \to \mathbb{R}$ defined by

$$\hat{v}(x_i, u(a)) = \frac{v(x_i, u(a)) - v(\underline{x}, u(a))}{v(\overline{x}, u(a)) - v(\underline{x}, u(a))}$$

[10] 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, 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 (5) 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.

is a *local* Bernoulli utility function associated with the $\succsim$-indifference curve containing $a \in A$. That is, the $\succsim$-indifference curve through $a \in A$ is a contour of the EU function with associated Bernoulli utility function $\hat{v}(\cdot, u(a))$. It follows that each contour of an IEU function is linear, but these contours need not be parallel.

An IEU 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 – see Dekel (1986, pp.306-7). In particular, if an IEU representation 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 $\hat{v} = v$ and

$$u(a) \;=\; \sum_{i=1}^{n} a_i v(x_i, u(a)) \qquad (5)$$

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

Since the contours of an IEU function are linear but not necessarily parallel, preferences with an IEU representation must satisfy a *betweenness* property (Axiom A4 in Dekel (1986)) but need not satisfy mixture independence.

Definition 5. *Preferences $\succsim$ 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 $\lambda \in (0, 1)$.*

To see that betweenness is a restricted form of mixture independence, it is useful to introduce some notation for linear segments (or *intervals*) in the probability simplex. 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.[11] The following is well known but we give a proof in Appendix A to keep the paper self-contained.[12]

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 $\lambda \in (0, 1)$:*

$$a \succsim b \quad \Leftrightarrow \quad a\lambda c \succsim b\lambda c \qquad (6)$$

Preferences satisfy *mixture independence* if (6) holds for *all* $a, b, c \in A$. Betweenness requires only that (6) holds when $a, b, c \in A$ are collinear – when all lie on a common linear segment in the simplex. In other words, betweenness is the requirement that mixture independence is satisfied by the restriction of $\succsim$ to any linear segment of A.

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

[12] The proof makes use of the assumed completeness of $\succsim$, but not transitivity.

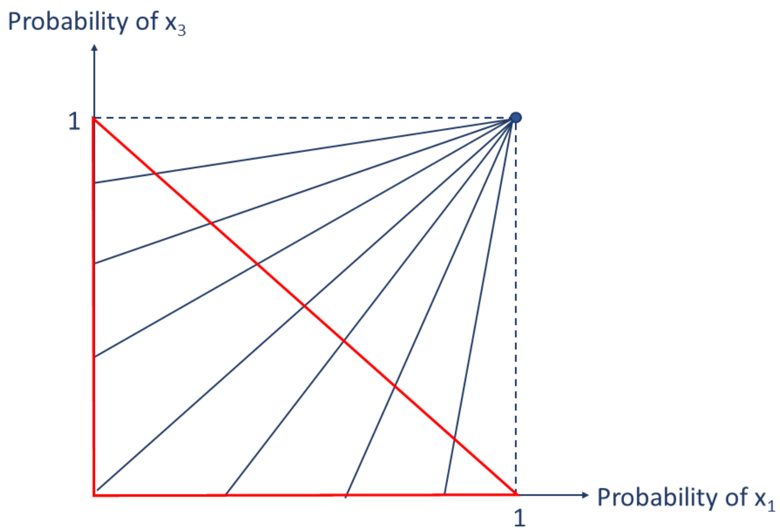

Figure 1: Preferences with an IEU representation

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

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

(i.e., $x_3 \succ x_2 \succ x_1$) 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. The point (a_1, a_3) in Figure 1 represents the lottery $(a_1, 1 - a_1 - a_3, a_3) \in A$. 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 (4) for any a. This is obvious if $a_1 = 1$. If $a_1 < 1$ we have:

$$zv(x_3, z) + (1 - z) v(x_1, z) \;=\; a_1 v(x_1, z) + a_2 v(x_2, z) + a_3 v(x_3, z)$$

$$\Leftrightarrow \qquad z \;=\; a_2 (1-z) + a_3$$

$$\Leftrightarrow \qquad z \;=\; \frac{a_2 + a_3}{1 + a_2} \;=\; \frac{1 - a_1}{1 + a_2} \;=\; \frac{1}{1 + x(a)}$$

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

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 satisfies the restricted form of linearity embodied in (4). 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 (4).[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, in particular, that any IEU function is continuous.)

2.2 Blavatskyy (2008)

When does a binary choice probability possess a stochastic EU representation? When does it have a stochastic IEU representation? Both of these questions are addressed by Blavatskyy (2008). He shows that the following axioms are sufficient for the existence of a stochastic EU representation:[14]

Axiom 1 (**Completeness**)**.** *For all $a, b \in A$,*

$$P(a,b) + P(b,a) = 1 \tag{7}$$

Axiom 2 (**Strong Stochastic Transitivity**)**.** *For all $a, b, c \in A$, if*

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

then

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

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

[14] The Completeness axiom is also known as the *Balance* condition (see, for example, Definition 4.9 in Falmagne, 2002). It implies that the decision-maker never abstains from making a choice. It also enforces the convention that $P(a,a) = \frac{1}{2}$ for any $a \in A$. (Recall, however, that these "diagonal" terms are not given any behavioural interpretation in the binary choice model.)

Axiom 3 (**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 4 (**Common Consequence Independence (CCI)**). *For any* $a, b, c, d \in A$ *and any* $\lambda \in [0,1]$

$$P(a\lambda c, b\lambda c) = P(a\lambda d, b\lambda d) \tag{8}$$

In fact, Blavatskyy claims that these axioms, plus one other that he calls *Interchangeability*, are necessary and sufficient for a Fechner representation with linear utility. His proof, however, establishes that his axioms suffice for a *strict* Fechner representation with linear utility (i.e., a stochastic EU representation). Ryan (2015) shows that the Interchangeability axiom is redundant – it is implied by the other axioms – and also verifies that Axioms 1-4 are *necessary* for a stochastic EU representation.

Theorem 1 (**Blavatskyy (2008); Ryan (2015)**). *Let* P *be a binary choice probability. Then* P *has a stochastic EU representation iff it satisfies Axioms 1-4.*

Note that CCI (Axiom 4) expresses the idea that choice probabilities respect a type of mixture independence property: the probability $P(a\lambda c, a\lambda c)$ is independent of the "common" lottery c – this probability would be unaffected by replacing c with some other lottery d. In particular, $a\lambda c \succsim^P a\lambda c$ iff $a\lambda d \succsim^P a\lambda d$.

Blavatskyy therefore makes the natural conjecture (*ibid.*, pp.1052-3) that P possess a stochastic IEU representation if (and only if) it satisfies Axioms 1-3 plus the following weakening of CCI:

Axiom 5 (**Stochastic Betweenness**). *For any* $e, f \in A$ *and any* $\alpha, \beta, \gamma, \mu \in [0,1]$ *with*

$$\alpha - \beta = \gamma - \mu \tag{9}$$

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

The fact that Stochastic Betweenness is weaker than CCI is not immediately obvious. We shortly re-state the Stochastic Betweenness condition in an equivalent form which makes this relationship more apparent (Lemma 1). First, however, let us clarify the role of condition (9) in Axiom 5. Observe that

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

for any $\alpha, \beta \in [0, 1]$ and any $e, f \in A$. Thus, if condition (9) holds, then

$$(e\alpha f) - (e\beta f) = (e\gamma f) - (e\mu f)\,.$$

Axiom 5 therefore says that for any $e, f \in A$ and any $a, b, a', b' \in [e, f]$ with

$$b - a = b' - a',$$

we must have $P(a, b) = P(a', b')$. In other words, if a, b, a', b' are collinear (i.e., all contained in some linear segment $[e, f] \subseteq A$) and $b - a = b' - a'$, then the probability of choosing a over b is the same as the probability of choosing a' over b' as Figure 2 illustrates.

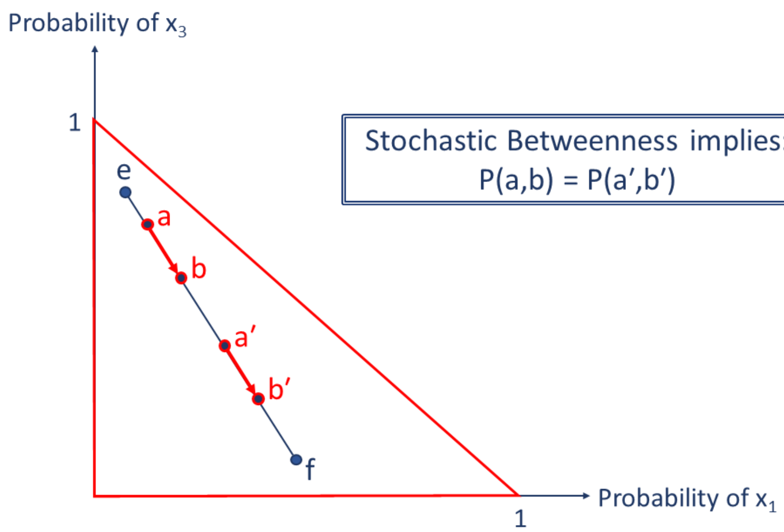

Figure 2: The red arrows have the same length

Lemma 1. *Let P be a binary choice probability. Then P satisfies Axiom 5 iff it satisfies the following for any $e, f \in A$, any $a, b, c, d \in [e, f]$ and any $\lambda \in [0, 1]$:*

$$P(a\lambda c, b\lambda c) = P(a\lambda d, b\lambda d) \tag{10}$$

Proof: Suppose P satisfies Axiom 5. Let $a, b, c, d \in [e, f]$ and $\lambda \in [0, 1]$. Since

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

we have $P(a\lambda c, b\lambda c) = P(a\lambda d, b\lambda d)$ by the argument above (just prior to the statement of the lemma).

Conversely, suppose P satisfies (10) 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

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

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

$$P(e\alpha f, e\beta f) \;=\; P(e\gamma f, e\mu f) \tag{12}$$

If $k = 0$ this is immediate: set $\lambda = 0$, $c = e\alpha f$ and $d = e\gamma f$ in (10). If $\beta = \mu$, then (11) implies $\alpha = \gamma$ and (12) holds trivially. Therefore, suppose that $k > 0$ and $\beta < \mu$. Since $\gamma \leq 1$, we have $k \leq 1 - \mu$ from (11), 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 (10) 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 imposes the CCI condition (8) 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, Stochastic Betweenness is *not necessary* for a stochastic IEU representation.

3 Violating Stochastic Betweenness

Let u be the IEU function from Example 1 and define $F : [-1, 1] \to [0, 1]$ as follows:

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

Note that F is *strictly* increasing and satisfies $F(x) + F(-x) = 1$. (Its domain can be extended to $\mathbb{R}$ without disturbing these properties, but $[-1, 1]$ is the relevant domain for our purposes.) Now construct P from u and F using (2). Then (u, F) is a strong Fechner representation for P by construction, and u has the IEU form. It is therefore a stochastic IEU representation for P (since any strong Fechner model is strict). We will show that P violates Stochastic Betweenness.

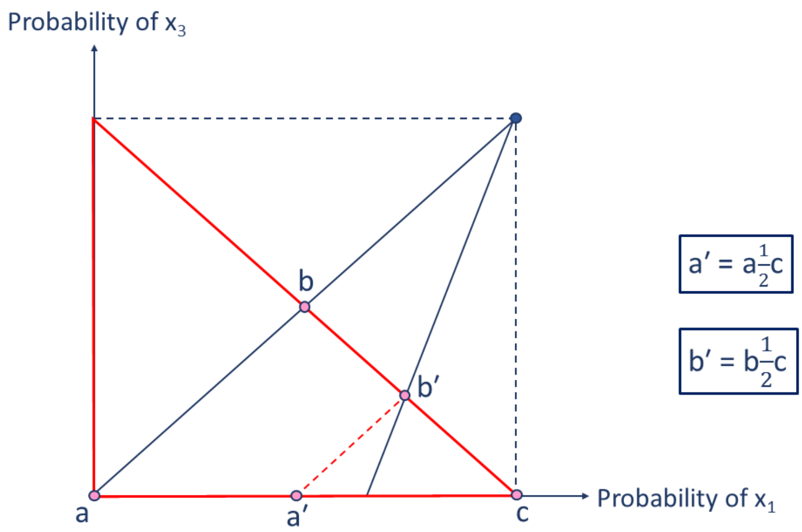

Figure 3: Constructing a violation of Stochastic Betweenness

Consider Figure 3. In this figure $b' = b\frac{1}{2}c$ and $a' = a\frac{1}{2}c$. Suppose, contrary to our claim, that P satisfies Stochastic Betweenness. Then we have:

$$\begin{aligned}
F\left(u\left(b\right)-u\left(b'\right)\right) &= P\left(b,b'\right) \\
&= P\left(b',c\right) && \text{(by Axiom 5)} \\
&= F\left(u\left(b'\right)-u\left(c\right)\right) \\
&< F\left(u\left(a'\right)-u\left(c\right)\right) && \text{(since } u\left(a'\right)>u\left(b'\right)\text{)} \\
&= P\left(a',c\right) \\
&= P\left(a,a'\right) && \text{(by Axiom 5)} \\
&= F\left(u\left(a\right)-u\left(a'\right)\right) \\
&= F\left(u\left(b\right)-u\left(a'\right)\right) && \text{(since } u\left(a\right)=u\left(b\right)\text{)} \\
&< F\left(u\left(b\right)-u\left(b'\right)\right) && \text{(since } u\left(a'\right)>u\left(b'\right)\text{)}
\end{aligned}$$

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 Stochastic Betweenness.

It is clear that there is nothing special about this example. If $\hat{P}$ is a binary choice probability and $\left(\hat{u},\hat{F}\right)$ is a stochastic IEU representation for $\hat{P}$, then we can construct a similar violation of Stochastic Betweenness provided $\hat{F}$ is strictly increasing and $\succsim^{\hat{P}}$ does not have an EU representation. In fact, with a more elaborate argument, one can obtain a violation of Stochastic Betweenness provided $\hat{F}$ is strictly increasing *in a neighbourhood of* 0. Given the strictness of the representation, this is equivalent to $\hat{F}$ being continuous at 0. Moreover, as the example suggests, we can detect violations of SB even when restricting attention to lotteries with no more than three outcomes, as in the experiments of Loomes and Sugden (1998).

Definition 6. *The* ***support*** *of* $a\in A$ *is the set* $\{i \mid a_i>0\}$.

Axiom 6. *A stochastic choice function* P *satisfies* ***Restricted Stochastic Betweenness*** *if, for any* $e,f\in A$ *such that the union of the supports of* e *and* f *contains no more than three elements, any* $a,b,c,d\in[e,f]$ *and any* $\lambda\in[0,1]$:

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

The proof of the following result may be found in Appendix B.

Theorem 2. *Let P be a binary choice probability that possesses a stochastic IEU representation (u, F) with F continuous at 0. If P satisfies Restricted Stochastic Betweenness then u has the EU form.*

Theorem 2 reveals the surprising strength of Stochastic Betweenness. Evidence that a decision-maker's choices violate Stochastic Betweenness does *not* imply that her choice behaviour is incompatible with a stochastic IEU representation. If P possesses a stochastic IEU representation with F continuous at 0, but does not possess a stochastic EU representation, then violations of Stochastic Betweenness are *inevitable* and such violations can (in principle) be detected even if experimental attention is restricted to lotteries whose outcomes are drawn from the same three-element set. In short, Stochastic Betweenness is unsuitable for testing the stochastic IEU model since it is not necessary for the existence of a stochastic IEU representation. The next section provides a set of necessary and sufficient conditions for P to possess such a representation.

However, before leaving the present section, it is important to observe that Theorem 2 would be false if we dropped the requirement that F be continuous at 0.

Example 2. *Let u be the utility function from Example 1 and define F as follows:*

$$F(x) = \begin{cases} 1 & \text{if } x > 0 \\ \frac{1}{2} & \text{if } x = 0 \\ 0 & \text{if } x < 0 \end{cases}$$

Consider the binary choice probability function P obtained from u and F using (2). Note that (u, F) is a stochastic IEU representation for P.

In this example, the decision-maker always makes a utility-maximising choice. She only chooses randomly when the two options have the same utility, choosing each option with probability $\frac{1}{2}$ in such cases. It is easily checked that P satisfies Axiom 6, using the fact that the preferences $\succsim$ satisfy the betweenness property. Since $\succsim^P = \succsim$ it is immediate that $\succsim^P$ does not have a linear representation and hence that P does not possess a stochastic EU representation.

4 A representation theorem

Consider the following pair of axioms:[15]

[15] When reading Axiom 8 recall that $\delta^i \in A$ is the degenerate lottery that gives outcome x_i with certainty.

Axiom 7. *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 8. *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 7 says that whenever mixing a with b produces a lottery that is no more (respectively, no less) stochastically desirable than a, then the complementary mixing of a with b produces a lottery that is no less (respectively, no more) stochastically desirable than b. Together with Axioms 1-3, Axiom 7 implies that $\succsim^P$ satisfies Dekel's (1986) Betweenness property – see Lemma 4 in Appendix C.

Axiom 8 says that if $\underline{\delta}$ and $\overline{\delta}$ are $\succsim^P$-worst and $\succsim^P$-best (respectively) amongst the degenerate lotteries, 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 2 – Strong Stochastic Transitivity – then $\succsim^P$ is a weak order so $\succsim^P$-worst and $\succsim^P$-best degenerate lotteries will exist.)

Axioms 7 and 8 are both implied by Stochastic Betweenness (Axiom 5). This is obvious in the case of Axiom 8. To see that Axiom 7 is also weaker than Stochastic Betweenness, 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 1-3, 7 and 8 do not yet suffice for a stochastic IEU representation, since they permit $\succsim^P$ to violate Dekel's (1986) Monotonicity axiom. In particular, it is possible that $b \succ^P a$ even if lottery a dominates lottery b in the sense of first-order stochastic dominance.[16]

Example 3. *Suppose* $X = \{x_1, x_2, x_3\}$. *Let* $\succsim$ *satisfy*

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

[16] These axioms *do*, however, 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 3.

(i.e., $x_3 \succ x_2 \succ x_1$) and have indifference classes as illustrated in the Machina Triangle of Figure 4, 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 (Dekel, 1986, Axiom A4). Let $u : A \to [0,1]$ be the representation for $\succsim$ that satisfies $u\left(\delta^3 \lambda \delta^1\right) = \lambda$ for each $\lambda \in [0,1]$. (That is, label each indifference curve in Figure 4 with the probability on x_3 at the point where the indifference curve intersects the hypotenuse of the triangle.) Now let P be the binary choice probability obtained by combining u with some strictly increasing F satisfying $F(x) + F(-x) = 1$ for all x. Hence $\succsim = \succsim^P$. Then Axioms 1-3, 7 and 8 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 $b \succ a$. (Note that a dominates b in the sense of first-order stochastic dominance.) It follows that these preferences do not have an IEU representation.

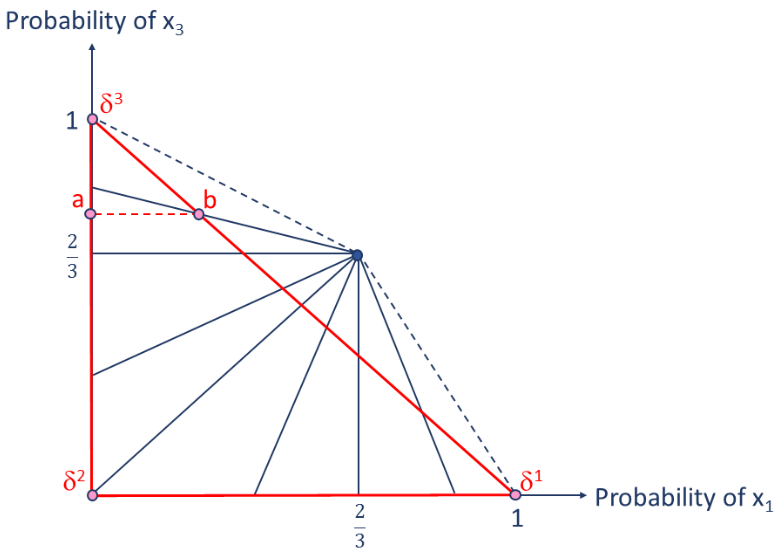

Figure 4: Preferences that violate FOSD-dominance

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

Axiom 9 (**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 3. *Let P be a binary choice probability. Then P has a stochastic IEU representation iff it satisfies Axioms 1-3 and 7-9.*

Theorem 3 is proved in Appendix C.

At this point, the reader may be wondering why Axiom 8 does not create problems like those illustrated in Section 3. Axiom 8 imposes the CCI condition (8) along the edge $[\underline{\delta}, \overline{\delta}]$ of the simplex A. Suppose P has a stochastic IEU representation (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 3 then implies that P^* satisfies the CCI condition along an edge of Δ joining "$\succsim^{P^*}$-best" and "$\succsim^{P^*}$-worst" vertices. If this logic holds for any sub-simplex, we deduce that the CCI condition holds along *any* edge of A, thereby creating the potential for the contradiction illustrated in Figure 3. 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 an edge joining a "$\succsim^{P^*}$-best" vertex to a "$\succsim^{P^*}$-worst" vertex 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 Discussion

Blavatskyy (2008) provides an important axiomatisation of the "EU plus Fechnerian noise" model. He further conjectures that weakening CCI to Stochastic Betweenness 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 Stochastic Betweenness with the (regrettably, less elegant) triumvirate of Axioms 7-9 (Theorem 3).

We end with some remarks about the interpretation of our results.

The axiomatic characterisation of a particular type of Fechner model may be useful for two reasons. First, the *necessity* of the axioms means that they are testable implications of the model. If Blavatskyy's (2008) conjecture were correct, then Stochastic Betweenness could be used to test the stochastic IEU model. Indeed, this axiom is readily testable along the lines of Loomes and Sugden (1998), though we are not aware of such tests having been conducted. Since Blavatskyy (2006) has argued that there is *prima facie* evidence in favour of the stochastic IEU model, such tests would be an attractive prospect for future research. Unfortunately, however, our results prove that Stochastic Betweenness is *not* necessary for the existence of a stochastic IEU representation, and hence Stochastic Betweenness *cannot* be used to test this model. On the other hand, Axioms 7 and 8 *are* necessary for the model and hence *could* be used for testing.

Second, the *sufficiency* of the axioms means that they may be used to assess the normative appeal, or perhaps the *a priori* "plausibility", of the model. For models of *stochastic* choice, this second role of an axiomatisation is less prominent. In models based on psychophysical theories, such as the ones considered here, it is not obvious how to articulate appropriate normative constraints on random choice behaviour. The underlying stochastic preferences, $\succsim^P$, could reasonably be assessed against standard normative criteria, but the randomness in the process of choice is "noise". One may nevertheless have reasons to think that the fitness of the decision-making organism demands (or at least is favoured by) certain structure to this noise. Fechner-type structures imply that noise varies inversely with the difference between two stimuli in a binary comparison. The axioms of Theorem 3 reduce the joint hypothesis that choice is based on IEU preferences expressed with Fechnerian noise to manageable properties of P that may be tested against our intuition.

The alternative approach to stochastic choice – that advocated by Machina (1985) – treats randomness as the *purposeful* outcome of choice: the decision-maker chooses to randomise if this maximises her deterministic utility function. This approach effectively treats the binary choice set $\{a, b\} \subseteq A$ as if it were the set $[a, b] \subseteq A$. Since utility is deterministic in this model, the axiomatic foundations of the model – were they to be ascertained – could be subjected to normative scrutiny in the usual fashion.[17] However, Machina's approach would not be useful for our purposes, since randomisation can only be

[17]Whatever their appeal might be, such axioms would certainly imply very different properties of P than those of Theorem 3. For example, adapting property (III) in Machina (1985, p.580) to our notation, if randomisation arises from the maximisation of a deterministic utility function, whatever its form, then the associated binary choice probability must satisfy the following condition: for all $a, b \in A$ and all $\underline{\lambda}, \overline{\lambda} \in [0,1]$ with $\underline{\lambda} < \overline{\lambda}$, if $P(a,b) \in [\underline{\lambda}, \overline{\lambda}]$ then

$$P\left(a\overline{\lambda}b, a\underline{\lambda}b\right) = \frac{P(a,b) - \underline{\lambda}}{\overline{\lambda} - \underline{\lambda}}.$$

optimal if the deterministic utility function is at least somewhere strictly quasiconcave (in probabilities). Our purpose, following Blavatskyy (2006, 2008), is to consider conditions under which stochastic choice behaviour might be guided by underlying preferences that satisfy a betweenness property. The utility representations of such preferences are nowhere strictly quasiconcave.

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 that $\succsim$ 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 the assumed completeness of $\succsim$ 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. □

Let $\succsim$ satisfy betweenness. Suppose $e, f \in A$ and $a, b, c \in [e, f]$. If $e = f$ or $a = b$ then (6) 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 (6) 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 (6) 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 binary choice probability that satisfies Axiom 1. Then P satisfies Strong Stochastic Transitivity (Axiom 2) iff*

$$P(a,b) \geq \frac{1}{2} \quad \Rightarrow \quad P(a,c) \geq P(b,c) \tag{13}$$

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

Condition (13) is called the *Weak Substitutability* property.

Turning to the proof of Theorem 2, let P be a binary choice probability that satisfies Restricted Stochastic Betweenness and has a stochastic IEU representation (u, F) with F continuous at 0. Hence, the indifference classes for $\succsim^P$ are linear and u is an implicit expected utility representation for $\succsim^P$. In particular, u is continuous. We must show that (u, F) is a stochastic expected utility representation for P. The result is obvious if u is constant, so assume otherwise henceforth.

Our argument is rather lengthy, so we break it into several steps.

Step 1. *P satisfies Axiom 2 (Strong Stochastic Transitivity).*

This follows from the facts that u represents $\succsim^P$ and F is non-decreasing: 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 $y \geq 0$. Thus, if $u(a) - u(b) \geq 0$ and $u(b) - u(c) \geq 0$ then

$$\begin{aligned} P(a,c) &= F([u(a) - u(b)] + [u(b) - u(c)]) \\ &\geq \max\{F(u(a) - u(b)), F(u(b) - u(c))\} \\ &= \max\{P(a,b), P(b,c)\}. \end{aligned}$$

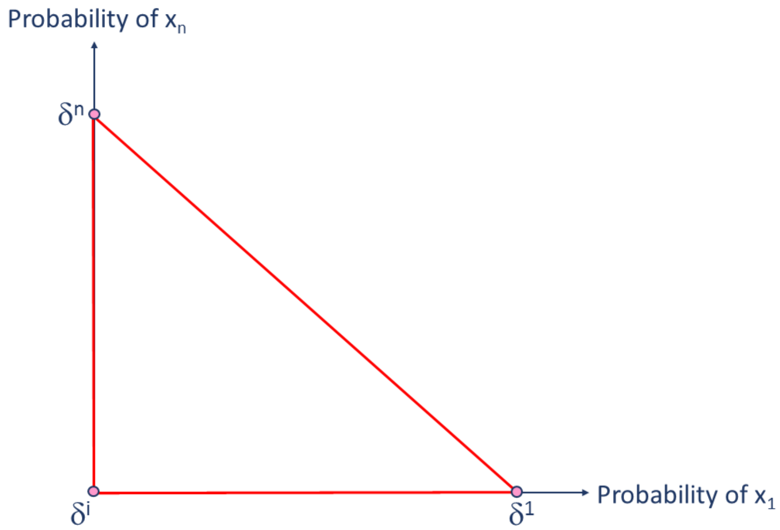

Figure 5: The Machina Triangle for Δ_i

Before proceeding to Step 2, we need some additional notation and definitions.

Since P satisfies completeness and SST, it is easy to see that $\succsim^P$ is complete and transitive. We therefore assume, without loss of generality, that the elements of X are indexed such that

$$\delta^n \succsim^P \delta^{n-1} \succsim^P \cdots \succsim^P \delta^1.$$

Since u is non-constant we must have

$$\delta^n \succ^P \delta^1.$$

Given any $i \in \{2, 3, ..., n-1\}$ let $\Delta_i \subseteq A$ denote the set of lotteries with outcomes in $\{x_1, x_i, x_n\}$. The Machina Triangle in Figure 5 depicts the set Δ_i. Let $\succsim_i^P \subseteq \Delta_i \times \Delta_i$ denote the restriction of $\succsim^P$ to Δ_i, with symmetric and asymmetric parts $\sim_i^P$ and $\succ_i^P$ respectively. By the betweenness property of $\succsim^P$ (Dekel (1986, Axiom A4)), every $\succsim_i^P$-indifference curve in the Machina Triangle (Figure 5) has a *unique* intersection with the hypotenuse. The collection of $\succsim_i^P$-indifference curves will also include at least one of the two singletons $\{\delta^1\}$ and $\{\delta^n\}$, since $\delta^n \succ^P \delta^1$.

Finally, given any $i \in \{2, 3, ..., n-1\}$ and any $b \in \Delta_i$ we define the binary relation $\succsim_i^b \subseteq \Delta_i \times \Delta_i$ as follows:

$$a \succsim_i^b c \quad \Leftrightarrow \quad P(a, b) \geq P(c, b).$$

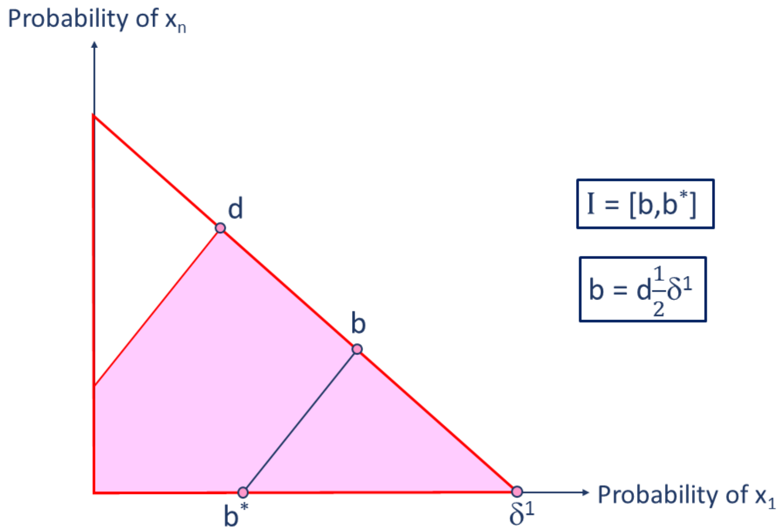

Figure 6: Points in the shaded region are squeezed by b

Since P satisfies Axiom 2, it follows from Lemma 3 that

$$\sim_i^P \subseteq \sim_i^b \tag{14}$$

where $\sim_i^b$ is the symmetric part of $\succsim_i^b$.

Given a point $b = \delta^1 \lambda \delta^n$ on the hypotenuse of the Machina Triangle in Figure 5 with $\lambda \geq \frac{1}{2}$ (respectively, $\lambda < \frac{1}{2}$), and its associated $\succsim_i^P$-indifference curve $I = [b, b^*]$, we say that $a \in \Delta_i$ is "squeezed by b" if a is contained some segment $[c, c']$ parallel to I,[18] with point c on the hypotenuse no further from b than is δ^1 (respectively, δ^n). Figure 6 illustrates the case $\lambda \geq \frac{1}{2}$. Any point in the shaded region of Figure 6 (including the boundaries of this region) is squeezed by b.

Step 2. *Let $i \in \{2, 3, ..., n-1\}$ and let b be a point on the hypotenuse of Δ_i with associated $\succsim_i^P$-indifference curve $I = [b, b^*]$. If a' and a'' are distinct points squeezed by b, with $[a', a'']$ parallel to I, then $a' \sim_i^b a''$.*

If $[a', a''] \subseteq I$ the claim follows from (14), so assume otherwise. Consider Figure 7. Points a' and a'' are squeezed by b and $[a', a'']$ is parallel to I. The remaining points in Figure 7 are constructed from a', a'' and I as follows:

- Point a is where the line through a' and a'' hits the hypotenuse of the triangle.

[18] We treat singletons as parallel to any segment.

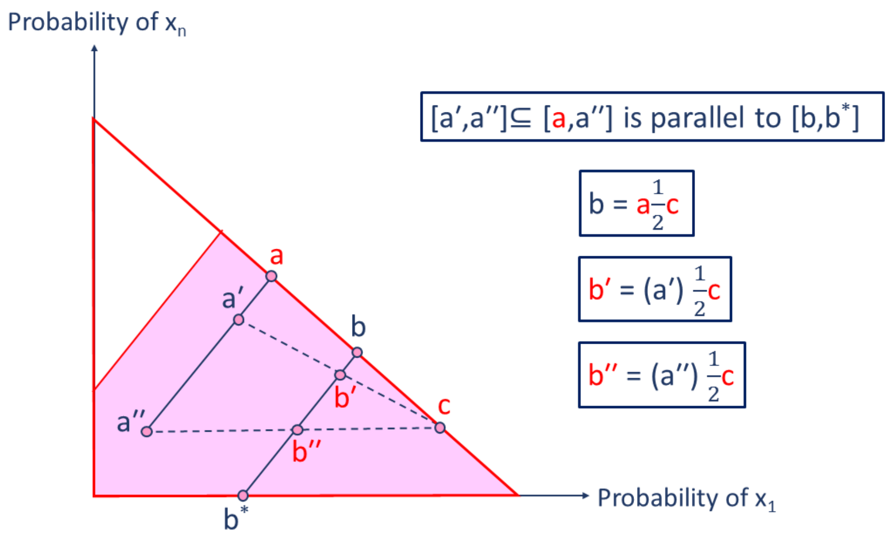

Figure 7: Construction to prove the Claim

- Point c is the point on the hypotenuse such that b is the mid-point of $[a, c]$. (Point c exists because a is squeezed by b.)
- Point b' is where the line joining a' to c intersects I, and b'' is the point where the line joining a'' to c intersects I.

By construction,

$$b' - c \; = \; a' - b'$$

and

$$b'' - c \; = \; a'' - b''.$$

Hence, using Restricted Stochastic Betweenness:

$$P(a', b') = P(b', c) \tag{15}$$

and

$$P(a'', b'') = P(b'', c) \tag{16}$$

Since $P(b', b'') = \frac{1}{2}$, the Weak Substitutability condition (13) implies

$$P(b', c) = P(b'', c) \tag{17}$$

Using (15)-(17) we have

$$P(a', b') = P(a'', b'') \tag{18}$$

Since $P(b', b) = P(b'', b) = \frac{1}{2}$, Weak Substitutability and Completeness allow us to replace b' and b'' with b in (18), which gives:

$$P(a', b) = P(a'', b) \quad \Leftrightarrow \quad a' \sim_i^b a''.$$

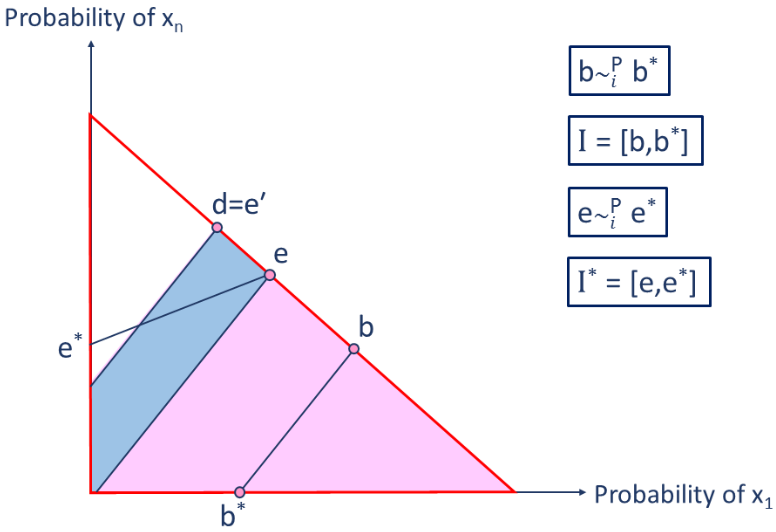

Figure 8: All points in the blue shaded area are $\succsim^b$-indifferent

Step 3. *Let $i \in \{2, 3, ..., n-1\}$ and let b and e be points on the hypotenuse of Δ_i with non-parallel $\succsim_i^P$-indifference curves $I = [b, b^*]$ and $I^* = [e, e^*]$ respectively. If a non-degenerate sub-interval $[e, e^{**}] \subseteq I^*$ is squeezed by b, then f is constant on an open neighbourhood of $u(e) - u(b)$.*

From (14) and the fact established in Step 2, there must be some $\succsim_i^b$-indifference class that contains a *non-degenerate* interval $[e, e'] \subseteq [\delta^1, \delta^n]$. This is illustrated in Figure 8, where $e' = d$ and the blue shaded region is contained within a single $\succsim_i^b$-indifference class. It follows that

$$P(\hat{e}, b) = F(u(\hat{e}) - u(b))$$

is *constant* for all $\hat{e} \in [e, e']$. Since u is an implicit expected utility representation for $\succsim^P$, and therefore $u(\delta^1 \lambda \delta^n)$ is strictly increasing in λ (Dekel (1986, Axiom A4)), we deduce that F must be constant on the non-degenerate interval

$$[u(e) - u(b), u(e') - u(b)].$$

Moreover, since u is continuous (and represents $\succsim^P$) points on the hypotenuse that are close enough to e will also occupy $\succsim_i^P$-indifference curves that are not parallel to I. It follows that F is constant on an *open neighbourhood* of $u(e) - u(b)$.

Step 4. *For any $i \in \{2, 3, ..., n-1\}$, the binary relation $\succsim_i^P \subseteq \Delta_i \times \Delta_i$ possesses an expected utility representation.*

Recall that F is non-decreasing, satisfies $F(x) + F(-x) = 1$ for any x, and is continuous at 0. Hence F must either be constant in some neighbourhood of 0 or else strictly increasing in some neighbourhood of 0.[19] The former scenario contradicts the assumption that u represents $\succsim^P$ (i.e., that the Fechner representation (u, F) is strict), so F must be strictly increasing in some neighbourhood of 0. Using this fact, together with the italicised observation in the previous paragraph, we will prove that all the non-singleton $\succsim_i^P$-indifference curves are parallel.

To do so, given any $\lambda \in (0,1)$, let us define $h^i(\lambda)$ to be the slope of the $\succsim_i^P$-indifference curve through the point $\delta^1 \lambda \delta^n \in \Delta_i$ in the Machina Triangle. Note that $h : (0,1) \to \mathbb{R}$ is a continuous function, since u is continuous. We must prove that it is constant. Suppose, to the contrary, that h is non-constant on $(0,1)$. From the continuity of h, it follows that there exists some $\hat{\lambda} \in (0,1)$ and some sequence $\{\lambda^m\}_{m=1}^{\infty} \subseteq (0,1)$ such that $\lambda^m \to \hat{\lambda}$ as $m \to \infty$ and $h(\lambda^m) \neq h\left(\hat{\lambda}\right)$ for each m.[20] For m sufficiently large, the $\succsim_i^P$-indifference curve through $\delta^1 \lambda^m \delta^3$ will have a

[19] Let $\overline{\varepsilon} = \inf\{\varepsilon > 0 \mid F(\varepsilon) = F(\varepsilon') \text{ for some } \varepsilon' > \varepsilon\}$. If $\overline{\varepsilon} > 0$ then F must be strictly increasing on $(-\overline{\varepsilon}, \overline{\varepsilon})$. If $\overline{\varepsilon} = 0$ then continuity of F at 0 implies that $F(0) = F(\varepsilon')$ for some $\varepsilon' > 0$, and hence F is constant on $[-\varepsilon', \varepsilon']$.

[20] Let $\lambda^*, \lambda^{**} \in (0,1)$ with $h(\lambda^*) < h(\lambda^{**})$. We assume $\lambda^* < \lambda^{**}$. (The argument for the other case is symmetric.) Now define

$$\overline{\lambda} \ = \ \max\{\lambda \in [\lambda^*, \lambda^{**}] \mid h(\lambda) \leq h(\lambda^*)\}.$$

Note that $\overline{\lambda}$ exists and $h\left(\overline{\lambda}\right) = h(\lambda^*)$ by the continuity of h. For each $\varepsilon \in \left[0, 1 - \overline{\lambda}\right)$, let

$$x(\varepsilon) \ = \ \sup\left\{h(\lambda) \mid \overline{\lambda} \leq \lambda < \overline{\lambda} + \varepsilon\right\}.$$

Thus, $x(\varepsilon) \geq h\left(\overline{\lambda}\right)$ for all ε and $x(\varepsilon)$ is non-decreasing in ε. If $x(\varepsilon) = h\left(\overline{\lambda}\right)$ for some $\varepsilon > 0$ we obtain a contradiction to the definition of $\overline{\lambda}$. Thus, $x(\varepsilon) > h\left(\overline{\lambda}\right) = h(\lambda^*)$ for all $\varepsilon > 0$. Thus, we can choose $\hat{\lambda} = \overline{\lambda}$ and for every $m \in \{1, 2, ...\}$ choose some

$$\lambda^m \in \left(\hat{\lambda}, \hat{\lambda} + \left(\frac{1 - \hat{\lambda}}{m+1}\right)\right)$$

with $h(\lambda^m) > h\left(\hat{\lambda}\right)$.

non-degenerate portion that is squeezed by $\delta^1\hat{\lambda}\delta^3$. For any such m, the function f is constant in some open neighbourhood of

$$z_m \equiv u\left(\delta^1\lambda^m\delta^3\right) - u\left(\delta^1\hat{\lambda}\delta^3\right).$$

Since $z_m \to 0$ as $m \to \infty$ it follows that F cannot be strictly increasing in any neighbourhood of 0. This is the desired contradiction. We therefore conclude that h is constant on $(0, 1)$. By the continuity of u, it follows that all non-singleton $\succsim_i^P$-indifference curves in the Machine Triangle are parallel. This completes Step 2.

Step 5. *The pair (u, F) is a SEUR for P.*

We must show that u is an expected utility representation for $\succsim^P$. Since u is a continuous representation for $\succsim^P$ of the implicit expected utility form (4), it suffices to prove that $v(x_i, z)$ is constant in $z \in (0, 1)$ for every $i \in \{2, 3, ..., n-1\}$. Fix some $i \in \{2, 3, ..., n-1\}$ and consider the Machina Triangle in Figure 5. For any $z \in (0, 1)$, the $\succsim_i^P$-indifference curve through the point $\delta^1 z \delta^n$ is a non-singleton with normal vector

$$(1 - v(x_i, z),\ -v(x_i, z)).$$

By Step 2, these normal vectors must be colinear in z. This is only possible if $v(x_i, z)$ is the same for every $z \in (0, 1)$:

$$(1 - v(x_i, z),\ -v(x_i, z)) \;=\; (\alpha - \alpha v(x_i, z'),\ -\alpha v(x_i, z')) \quad \Rightarrow \quad \alpha = 1.$$

This completes the proof of Theorem 2.

C Proof of Theorem 3

Suppose that P satisfies Axioms 1-3 and 7-9. We start by showing that $\succsim^P$ has an IEU representation.

Axiom 2 ensures that $\succsim^P$ is a weak order. Since X is finite, $\succsim^P$ must satisfy Dekel's (1986) Axiom A1. We assume (without loss of generality) 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 9, and the fact that the set $\left\{\delta^1, ..., \delta^n\right\}$ 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 $\lambda > \mu$,*

$$a \sim^P b \quad \Rightarrow \quad a\lambda b \sim^P b \tag{19}$$

and

$$a \succ^P b \quad \Rightarrow \quad a\lambda b \succ^P a\mu b \tag{20}$$

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 7 gives $a\left(\frac{1}{2}\right)b \succ^P b$, so $a \succ^P b$ by the 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 the 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 3 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 (19).

We now prove (20). If $a \succ^P b$ then, by an argument similar to the one above, we can use Axioms 2 and 7 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 3, $a\lambda b \succsim^P a(y^s)b$ for all s, and hence, applying Axiom 3 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 (19). 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 $a \succ^P b$, (20) implies that

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

for any $\lambda \in [0,1]$. By standard arguments, Axioms 2-3 imply that the set

$$S = \{\lambda \in [0,1] \mid a\lambda b \succsim^P c\} \cap \{\lambda \in [0,1] \mid a\lambda b \precsim^P c\}$$

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 $u(a) = \alpha$. Moreover, 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 8 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 stochastic IEU representation for P.

To prove the converse part of the Theorem, suppose that (u, F) is a stochastic IEU representation for P. Since $F(x) + F(-x) = 1$, it follows that Axiom 1 is satisfied. To see that P satisfies Strong Stochastic Transitivity (Axiom 2), recall Step 1 in the proof of Theorem 2. To verify Axiom 3 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 3 therefore follows from the continuity of u. We next verify Axioms 7 and 9. 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 9 follows directly from Dekel's Axiom A3 (and the fact that $\{\delta^1, ..., \delta^n\}$ is weakly ordered by $\succsim^P$) while Axiom 7 is implied by the completeness of $\succsim^P$ and Dekel's Axiom A4 (Betweenness). Finally, we deduce Axiom 8 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 3.

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.

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.

MACHINA, M.J. (1985), “Stochastic Choice Functions Generated from Deterministic Preferences Over Lotteries,” *The Economic Journal* **95(379)**, 575–594.

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.

RYAN, M.J. (2016), “Representation of Binary Choice Probabilities. Part I: Scalability,” *AUT Economics Working Paper 16/04*.

Ryan, M.J. (2017), "Uncertainty and Binary Stochastic Choice," *Economic Theory* (forthcoming).