Economic Theory (2021) 72:641–669
https://doi.org/10.1007/s00199-020-01307-8

RESEARCH ARTICLE

Stochastic expected utility for binary choice: a 'modular' axiomatic foundation

Matthew Ryan[1]

Received: 19 December 2019 / Accepted: 25 August 2020 / Published online: 31 August 2020

Abstract
We present new axiomatisations for various models of binary stochastic choice that may be characterised as "expected utility maximisation with noise". These include axiomatisations of simple scalability (Tversky and Russo in J Math Psychol 6:1–12, 1969) with respect to a scale having the expected utility (EU) form, and strong utility (Debreu in Econometrica 26(3):440–444, 1958) of the EU form. The latter model features Fechnerian "noise": choice probabilities depend on EU differences. Our axiomatisations complement the important contributions of Blavatskyy (J Math Econ 44:1049–1056, 2008) and Dagsvik (Math Soc Sci 55:341–370, 2008). Our representation theorems set all models on a common axiomatic foundation, with additional axioms added in modular fashion to characterise successively more restrictive models. The key is a decomposition of Blavatskyy's (2008) *common consequence independence* axiom into two parts: one (which we call *weak independence*) that underwrites the EU form of utility and another (*stochastic symmetry*) than underwrites the Fechnerian structure of noise. We also show that in many cases of interest (which we call *preference-bounded domains*) stochastic symmetry can be replaced with *weak transparent dominance (WTD)*. For choice between lotteries, WTD only restricts behaviour when choosing between probability mixtures of a "best" and a "worst" possible outcome.

Keywords Stochastic choice · Expected utility · Scalability · Fechner

JEL Classification D01 · D81

Previous drafts have circulated under the title"Stochastic Expected Utility for Binary Choice: New Representations".

✉ Matthew Ryan
mryan@aut.ac.nz

1 School of Economics, Auckland University of Technology, Auckland, New Zealand

1 Introduction

Stochastic generalisations of expected utility (EU) maximisation come in two flavours: models of random utility and models of noisy expected utility maximisation—the latter are also known as "single utility" models. Gul and Pesendorfer (2006) exemplify the former. There is a fixed probability distribution over Bernoulli utility functions (i.e., functions that attach utilities to *deterministic* outcomes). One such function is randomly assigned every time a choice must be made, and the expected value of this function is maximised in the process of choice. Prominent in the latter category (the single utility models) are Blavatskyy (2008) and Dagsvik (2008). In these models, the decision-maker has a fixed Bernoulli utility function, but is prone to error in assessing which option maximises its expected value.

Blavatskyy (2008) restricts attention to binary choice. He axiomatises a Fechnerian generalisation of EU in which the probability of choosing lottery α over lottery β is a *non-decreasing* function of the *difference* between the expected utility of α and the expected utility of β. The option with the higher expected utility is chosen at least 50% of the time but not necessarily with certainty. Blavatskyy's *common consequence independence (CCI)* axiom plays a dual role in his axiomatisation: it underwrites both the EU form (i.e., mixture-linearity) of the utility function over lotteries, and also the Fechnerian structure of decision "noise"—the dependence of choice probabilities on EU differences.

The key contribution of the present paper is to decompose CCI into two parts: Axiom 5 (*weak independence*), which ensures mixture-linearity of utility, and Axiom 6 (*stochastic symmetry*) which guarantees Fechnerian structure. By axiomatically decoupling linearity of utility from Fechnerian "noise", we demarcate the boundary between Fechnerian generalisations of EU and choice probabilities which are *simply scalable* (Tversky and Russo 1969) with respect to an EU scale. The latter requires only that the probability of choosing lottery α over lottery β is increasing in the expected utility of α and decreasing in the expected utility of β. This is a generalisation of the Fechner model. Our decomposition of CCI therefore allows us to better understand the restrictions of Fechnerian structure over and above those implied by scalability with respect to an EU scale. It also means that we can characterise a range of variant models in a coherent and modular fashion, replacing a menagerie of disparate axiomatisations. In particular, we provide a new axiomatisation for the continuous Fechnerian model of Dagsvik (2008, Theorem 4).

The following section reviews some basic ideas from binary stochastic choice. Our decomposition and representation results are given in Sect. 3, which is divided into several subsections. Section 3.1 reviews two classical models of binary stochastic choice. Section 3.2 gives our decomposition of CCI. In Sect. 3.3 we use this decomposition to provide an axiomatisations of EU embedded in several stochastic specifications: simple scalability, the classical Fechner model (Debreu 1958) and two Fechnerian models from Dagsvik (2008). Section 3.4 introduces the notion of a *preference-bounded domain* and shows that, for this special environment, stochastic symmetry can be replaced by a simple dominance condition. Section 3.5 contains brief comments on extensions to multinomial choice. Further discussion and interpretation of results is given in the concluding Sect. 4. Most proofs, and some supplementary material, are contained in the Appendix.

2 Binary choice probabilities

In Blavatskyy (2008) and Dagsvik (2008), alternatives are lotteries over a fixed, finite set of outcomes. With the exception of the analysis in Sect. 3.4, our results apply to any *mixture set*,[1] so we present our analysis in that (broader) context. This added generality is not without interest. For example, our analysis may be applied to models of subjective uncertainty as well as objective risk.[2]

Let A be a mixture set of alternatives. If $a, b \in A$ and $\lambda \in [0, 1]$ we use $a\lambda b$ to denote the λ-mixture of a and b. In particular, $a1b = a$ and $a0b = b$. For example, A might be the set of lotteries over some finite set $X = \{x_1, x_2, \ldots, x_n\}$ of outcomes, as in Blavatskyy (2008) and Dagsvik (2008). In this case, A is the unit simplex in $\mathbb{R}^n$ and if $a, b \in A$ and $\lambda \in [0, 1]$ then $(a\lambda b)_i = \lambda a_i + (1 - \lambda)\, b_i$ is the probability of receiving outcome x_i in lottery $a\lambda b$.

A function $u : A \to \mathbb{R}$ is *mixture-linear* if $u\,(a\lambda b) = \lambda u\,(a) + (1 - \lambda)\, u\,(b)$ for any $a, b \in A$ and any $\lambda \in [0, 1]$. When A is the set of lotteries over X, then u is mixture-linear if it has the expected utility form.[3] In this case, if $a \in A$ has $a_i = 1$ then $u\,(a)$ is the Bernoulli utility of outcome x_i. Note that $u\,(A)$ is an interval when u is mixture-linear (Ryan 2018a, Lemma 2); a fact of which we will make use in what follows.

The objects of analysis are functions $P : A^2 \to [0, 1]$. If $a \neq b$ then $P\,(a, b)$ is the probability that a given decision-maker chooses alternative a from the binary choice set $\{a, b\}$. If $a = b$ then no behavioural interpretation is given. We call such a function a *binary choice probability (BCP)*. It is natural that any BCP should satisfy $P\,(a, b) + P\,(b, a) = 1$ for any distinct $a, b \in A$, but we follow Blavatskyy (2008)

[1] See Herstein and Milnor (1953) or Fishburn (1982, Sect. 2.1) for a formal definition. For readers who are unfamilar with mixture sets, there is some loss of generality but no damage to comprehension from replacing “mixture set” with “convex subset of some Euclidean space”.

[2] Consider the set of all mappings $f : S \to \mathcal{D}$ from a given finite state space, S, to the set, $\mathcal{D}$, of distribution functions on $\mathbb{R}_+$. Thus, f is an *Anscombe-Aumann act* representing an action taken in an environment of uncertainty: act f delivers lottery $f\,(s)$ in state $s \in S$. Since $\mathcal{D}$ is convex, the set of Anscombe-Aumann acts is also convex, hence a mixture set.

[3] When A is a set of Anscombe-Aumann acts, the familiar *subjective expected utility* function is mixture-linear:

$$u\,(f) = \sum_s p\,(s)\, v\,(f\,(s))$$

where $p : S \to [0, 1]$ is a (subjective) probability mass function on S and v has the expected utility form. That is, for any $F \in \mathcal{D}$:

$$v\,(F) = \int_0^\infty \overline{v}\,(x)\;dF\,(x)$$

for some Bernoulli utility function $\overline{v} : \mathbb{R}_+ \to \mathbb{R}$.

and Dagsvik (2008) in imposing this as an axiomatic restriction rather than as part of the definition of a BCP.[4]

Axiom 1 (Balance) *For any* $a, b \in A$, $P(a, b) + P(b, a) = 1$.

Note that balance requires $P(a, a) = \frac{1}{2}$ for any $a \in A$. A binary choice probability that satisfies Axiom 1 will be called *balanced*.

Associated with any binary choice probability, P, is the following binary relation on A:

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

That is, $a \succsim^P b$ iff the decision-maker is at least as likely to choose a as to choose b in a binary choice. The asymmetric and symmetric parts of $\succsim^P$, denoted $\succ^P$ and $\sim^P$ respectively, are defined in the usual way. We call $\succsim^P$ the *base relation* for P, by analogy with the theory of deterministic choice functions.[5] If P is balanced then

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

and the base relation is complete (i.e., for any $a, b \in A$, either $a \succsim^P b$ or $b \succsim^P a$), hence reflexive. However, it is not transitive unless P satisfies the following condition, known as *weak stochastic transitivity (WST)*:

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

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

Definition 1 (Marschak 1960) A weak utility for P is a function $u : A \rightarrow \mathbb{R}$ that represents the base relation: that is:

$$a \succsim^P b \quad \Leftrightarrow \quad u(a) \geq u(b)$$

for any $a, b \in A$.

The reader is warned that the term "weak utility" is ambiguous in the decision theory literature. Our usage follows that in the stochastic choice literature.[6]

[4] Blavatskyy (2008) calls Axiom 1 "completeness" but we adopt Dagsvik's (2008) terminology here.

[5] The reader is warned that this terminology is not standard for stochastic choice.

[6] Elsewhere in decision theory, a function $u : A \rightarrow \mathbb{R}$ is said to be a "weak utility" for an asymmetric binary relation, $\succ$, if it provides a *one-way* representation in the following sense:

$$a \succ b \quad \Rightarrow \quad u(a) > u(b).$$

3 Models and representations

3.1 Simple scalability and Fechner representations

Standard representations for BCPs relate choice probabilities to weak utilities. The two best known representations are the following:

Definition 2 (Tversky and Russo 1969) A binary choice probability P is simply scalable (or satisfies simple scalability) iff there is a weak utility function (or "scale") $u : A \to \mathbb{R}$ for P and a function $F : u(A) \times u(A) \to [0, 1]$ that is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument such that

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

for all $a, b \in A$. In this case, we say that (u, F) is a simple scale representation for P.

Definition 3 (Debreu 1958; Marschak 1960) A binary choice probability P has a Fechner representation iff there exists a weak utility $u : A \to \mathbb{R}$ for P and a strictly increasing function $G : \Gamma_u \to [0, 1]$, where $\Gamma_u = \{x - y \in \mathbb{R} \mid x, y \in u(A)\}$, such that

$$P(a, b) = G(u(a) - u(b)) \tag{3}$$

for all $a, b \in A$. In this case, we say that (u, G) is a Fechner representation for P.

Simple scalability captures the most basic sense in which a BCP might be said to describe a process of noisy utility maximisation. For a Fechner representation, we impose the additional constraint that this noise depend only on the utility difference between the alternatives: P has a Fechner representation iff there is a simple scale representation (u, F) for P with $F(x, y) = F(x', y')$ whenever $x - y = x' - y'$. It is also useful to note that P has a Fechner representation iff it has a weak utility $u : A \to \mathbb{R}$ which also satisfies

$$P(a, b) \geq P(c, d) \quad \Leftrightarrow \quad u(a) - u(b) \geq u(c) - u(d) \tag{4}$$

for any $a, b, c, d \in A$. A weak utility for P that additionally satisfies (4) is called a *strong utility* for P.

Neither Definition 2 nor Definition 3 would be affected if we dropped the requirement that the mapping $u : A \to \mathbb{R}$ is a weak utility for P, and it is conventional to drop it. The monotonicity properties of F and G (respectively) ensure that u must be a weak utility, so this fact could have been derived as an implication. However, we find it more convenient to include it explicitly in the definitions.

If P is balanced and (u, F) is a simple scale representation for P then F must satisfy

$$F(x, y) + F(y, x) = 1 \tag{5}$$

for any $x, y \in u(A)$. Conversely, if (u, F) is a simple scale representation for P and F satisfies (5) for any $x, y \in u(A)$, then P is balanced. Similarly, if (u, G) is a Fechner representation for P, then P is balanced iff

$$G(z) + G(-z) = 1 \tag{6}$$

for any $z \in \Gamma_u$. It is common to include these restictions on F and G in the definitions of simple scalability and Fechner representation respectively, but since balancedness is not part of our definition of a BCP it is more appropriate to exclude them here. We say that (u, G) is a *balanced Fechner representation* for P if it is a Fechner representation and G satisfies (6) for any $z \in \Gamma_u$. Likewise, we say that P satisfies *balanced simple scalability* if it has a simple scale representation (u, F) such that (5) holds for any $x, y \in u(A)$.

When does a BCP satisfy balanced simply scalability with respect to a mixture-linear scale? When does it possess a balanced Fechner representation with a mixture-linear weak utility function? Answers to these questions are given in Sect. 3.3 and summarised in Fig. 1. Blavatskyy (2008) and Dagsvik (2008, Theorem 4) already addressed questions of the second type. However, they characterise Fechner-like structures which are, respectively, slightly weaker and slightly stronger than Definition 3.

Blavatskyy (2008) characterises BCPs that possess a balanced representation of the form (3) with u mixture-linear and G *non-decreasing*. An important advantage of this weaker monotonicity requirement on G is that it accommodates behaviour in which the "better" option is chosen with certainty for a non-trivial range of utility differences. Classical Fechner representations—such as the familiar binary logit structure—exclude the possibility of choice being *certain* (i.e., $P(a, b) \in \{0, 1\}$) unless choosing between the most "extreme" alternatives.

Dagsvik (2008, Theorem 4) provides necessary and sufficient conditions for P to have a balanced Fechner representation with u mixture-linear and G *continuous*. The added continuity requirement is substantive. The following example illustrates that a balanced BCP may possess a mixture-linear strong utility yet not possess any Fechner representation (u, G) in which u is mixture-linear and G is continuous.

Example 1 Suppose $A = [0, 1]$ and

$$P(a, b) = \begin{cases} \frac{1}{4} - \frac{1}{4}(b - a) & \text{if } a < b \\ \frac{1}{2} & \text{if } a = b \\ \frac{3}{4} + \frac{1}{4}(a - b) & \text{if } a > b \end{cases}$$

Hence, the range of P is $\left[0, \frac{1}{4}\right) \cup \left\{\frac{1}{2}\right\} \cup \left(\frac{3}{4}, 1\right]$. If $u : A \to \mathbb{R}$ is mixture-linear, then $u(A)$ is an interval and so is Γ_u, which means that $G(\Gamma_u)$ must also be an interval for any continuous and strictly increasing G. It follows that P cannot have a Fechner representation (u, G) in which u is mixture-linear and G continuous. However, the identity function is a mixture-linear strong utility for P, as the reader may easily verify.

The case of continuous Fechner representations is obviously of interest. Indeed, some definitions of the Fechner representation (such as Definition 17 in Luce and Suppes 1965) require G to be a distribution function, hence right-continuous. In this case balance implies continuity of G, as noted by Dagsvik (2008, p. 359).

Blavatskyy (2008) and Dagsvik (2008, Theorem 4) not only characterise somewhat different Fechner-like stuctures, their respective representation theorems also employ very different sets of axioms.[7] Our objective is to set both representations on a common footing, and one that builds, in modular fashion, on a characterisation of scalability with respect to a mixture-linear scale.

3.2 Deconstructing Blavatskyy (2008)

We start by recalling Blavatskyy's (2008) representation theorem. It is based on the balance condition (Axiom 1) together with the following three axioms:[8]

Axiom 2 (Strong stochastic transitivity (SST)) *For any $a, b, c \in A$,*

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

Axiom 3 (Continuity) *For any $a, b, c \in A$ the following sets are closed*

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

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

Theorem 1 (Blavatskyy 2008 (modified by Ryan 2015)) *Let P be a BCP. Then P satisfies Axioms 1–4 iff there exists a mixture-linear weak utility $u : A \to \mathbb{R}$ and a non-decreasing function $G : \Gamma_u \to [0,1]$ satisfying (6) for any $z \in \Gamma_u$, such that (3) holds for all $a, b \in A$.*

As suggested in the Introduction, CCI does dual service in this representation result: it underwrites both a mixture-linear weak utility function and a Fechnerian noise structure. More precisely:

Proposition 1 *There exists a binary choice probability, P, that satisfies Axioms 1–3 but does not have a mixture-linear weak utility, nor a representation of the form (3) with u a weak utility and G non-decreasing.*

[7] See Dagsvik (2015) for some variant axiomatisations of the model in Dagsvik (2008, Theorem 4).

[8] Blavatskyy also included a fifth axiom, *interchangeability*, but this is implied by balance and strong stochastic transitivity—see Ryan (2015). Furthermore, Blavatskyy proves his result for the special case in which A is the unit simplex in $\mathbb{R}^n$. Our version is therefore more general. It is an immediate corollary of Theorem 3 below.

We wish to axiomatically disentangle these two roles played by CCI. We do so using the following pair of axioms:

Axiom 5 (Weak independence) *For any* $a, b, c \in A$*,*

$$P\left(a\frac{1}{2}c, b\frac{1}{2}c\right) > \frac{1}{2} \quad \Rightarrow \quad \min\left\{P\left(a, a\frac{1}{2}b\right), \; P\left(a\frac{1}{2}b, b\right)\right\} > \frac{1}{2}.$$

Axiom 6 (Stochastic symmetry) *For any* $a, b \in A$ *and any* $\lambda \in [0, 1]$*,*

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

Axiom 6 is a stochastic analogue of the *symmetry* axiom from SSB utility theory (Fishburn 1984), hence the name. It implies, in particular, that $a\frac{1}{2}b$ is a *stochastic mid-point* between a and b (Davidson and Marschak 1959).

Weak independence imposes the following restriction on the base relation for P:

$$a\frac{1}{2}c \succ^P b\frac{1}{2}c \quad \Rightarrow \quad \left[a \succ^P a\frac{1}{2}b \text{ and } a\frac{1}{2}b \succ^P b\right]$$

for any $a, b, c \in A$. While this is not a standard restriction on preferences, it can be shown that weak independence, together with balance and WST (which is implied by SST), suffice for $\succsim^P$ to satisfy a classical mixture-independence condition:[9]

Proposition 2 *Let* P *be a BCP that satisfies balance, WST and weak independence. Then the following holds for any* $a, b, c \in A$*:*

$$a \sim^P b \quad \Rightarrow \quad a\frac{1}{2}c \sim^P b\frac{1}{2}c \tag{7}$$

The weak independence axiom is also closely related to (and its name motivated by) the independence axiom of Dagsvik (2008), which, when stated in contrapositive form, says that:

$$a\lambda c \succ^P b\lambda c \quad \Rightarrow \quad a \succ^P b \tag{8}$$

for any $a, b, c \in A$ and any $\lambda \in (0, 1)$. It is obvious that (8) implies (7). Furthermore, given balance, WST and continuity, (7) implies weak independence: see Fishburn (1982; proof of H4 from B1–B3 on pp. 16–17). Hence, given balance, WST and continuity, Dagsvik's (2008) independence property implies weak independence.

Stochastic symmetry is implied by CCI as follows:

$$P(a, a\lambda b) = P(a(1-\lambda)a, b(1-\lambda)a) = P(a(1-\lambda)b, b(1-\lambda)b) = P(b\lambda a, b)$$

[9] Condition (7) is the independence condition of Herstein and Milnor (1953), though they did not call it such. It appears, under the "independence" nomenclature, as Axiom B2 in Fishburn (1982).

where CCI is used for the middle equality. Weak independence also follows from CCI. To see this note that

$$P\left(a, a\frac{1}{2}b\right) = P\left(a\frac{1}{2}a, b\frac{1}{2}a\right) = P\left(a\frac{1}{2}c, b\frac{1}{2}c\right) = P\left(a\frac{1}{2}b, b\frac{1}{2}b\right) = P\left(a\frac{1}{2}b, b\right)$$

(where the second and third equalities use CCI) and hence

$$P\left(a\frac{1}{2}c, b\frac{1}{2}c\right) = P\left(a, a\frac{1}{2}b\right) = P\left(a\frac{1}{2}b, b\right).$$

Thus, CCI implies the conjunction of Axioms 5 and 6. In fact, the latter conjunction is *strictly* weaker than CCI.

Proposition 3 *Let P be a BCP. If P satisfies CCI then it satisfies Axioms 5 and 6 but not conversely.*

The next two theorems verify that Axioms 5 and 6 provide the desired decomposition of CCI.

Theorem 2 *Let P be a BCP. Then the following are equivalent:*

(i) *P satisfies Axioms 1–3 and 5.*
(ii) *There exists a mixture-linear weak utility $u : A \to \mathbb{R}$ and a function*

$$F : u(A) \times u(A) \to [0, 1]$$

that is non-decreasing (respectively, non-increasing) in its first (respectively, second) argument and which satisfies (5) for any $x, y \in u(A)$, such that (2) holds for all $a, b \in A$.

Theorem 3 *Let P be a BCP. Then P satisfies Axioms 1–3 and 5–6 iff there exists a mixture-linear weak utility $u : A \to \mathbb{R}$ and a non-decreasing function $G : \Gamma_u \to [0, 1]$ satisfying (6) for any $z \in \Gamma_u$, such that (3) holds for all $a, b \in A$.*

Since we already observed that CCI implies Axioms 5–6 but not conversely, Theorem 1 is implied by Theorem 3 and the latter shows that CCI can be relaxed in Blavatskyy's (2008) result.

Theorem 2 shows that Axioms 1–3 and 5 suffice for scalability with respect to a mixture-linear scale. However, it is not *simple* scalability (Definition 2) since F is only weakly monotone.[10] Adding stochastic symmetry gives Fechnerian structure to the noise (Theorem 3).

If continuity of G is important, it may be guaranteed by adding one further axiomatic piece. Debreu's (1958) solvability condition is sufficient for this purpose.

[10] It corresponds to *strict scalability* as defined in Ryan (2018b), which is slightly stronger than the classical notion of *monotone scalability* (Fishburn 1973).

Axiom 7 (Solvability) *For any* $a, b, c \in A$ *and any* $\rho \in [0, 1]$ *if*

$$P(a, b) \geq \rho \geq P(a, c)$$

then $P(a, d) = \rho$ *for some* $d \in A$.

Theorem 4 *Let* P *be a BCP. Then* P *satisfies Axioms 1 –3 and 5–7 iff there exists a mixture-linear weak utility* $u : A \to \mathbb{R}$ *and a continuous, non-decreasing function* $G : \Gamma_u \to [0, 1]$ *satisfying* (6) *for any* $z \in \Gamma_u$, *such that* (3) *holds for all* $a, b \in A$.

We could also add solvability to Theorem 2 to ensure that F is continuous in each argument.[11] The reader will easily be able to adapt the proof of Theorem 4 to this purpose so we omit the details.

3.3 Reconstructing Dagsvik (2008)

Using Axioms 5–6 we can provide a new, modular axiomatic foundation for simple scalability with respect to a mixture-linear scale, plus the Fechner and the continuous Fechner (Dagsvik 2008, Theorem 4) representations with mixture-linear utility. To ensure the strict monotonicity of F and G in these respective representations we must strengthen the SST condition:

Axiom 8 (SSST) *For any* $a, b, c \in A$,

$$\min\{P(a, b),\ P(b, c)\} \geq [>]\ \frac{1}{2} \ \Rightarrow\ P(a, c) \geq [>] \max\{P(a, b),\ P(b, c)\}.$$

This axiom plays a decisive role in Russo and Tversky's (1969) characterisation of strict scalability. It goes by various names in the literature. Tversky and Russo themselves refer to it as "strong stochastic transitivity" but this term is now firmly affixed to the weaker concept defined by Axiom 2. Fishburn (1973) calls it "strict stochastic transitivity", abbreviated "SSST". Roberts (1971) calls it the "strong version of strong stochastic transitivity", also using the acronym SSST. We therefore refer to Axiom 8 simply by the SSST acronym, consistent with both Fishburn and Roberts.

Replacing SST with SSST in Theorems 2–4 we have:

Theorem 5 *Let* P *be a BCP. Then* P *satisfies Axioms Axioms 1, 3, 5 and 8 iff it satisfies balanced simple scalability with a mixture-linear scale.*

Theorem 6 *Let* P *be a BCP. Then* P *satisfies Axioms 1, 3, 5–6 and 8 iff* P *has a balanced Fechner representation* (u, G) *with* u *mixture-linear.*

Theorem 7 *Let* P *be a BCP. Then* P *satisfies Axioms 1, 3, 5–8 iff* P *has a balanced Fechner representation* (u, G) *with* u *mixture-linear and* G *continuous.*

[11] Using the other properties of F it follows that F is (jointly) continuous, as observed by Debreu (1958).

Once again, we could add solvability to Theorem 5 to ensure that F is continuous in each argument.

Figure 1 summarises the results of this section. The results of Sect. 3.2 follow exactly the same pattern, but with the strict monotonicity of F and G relaxed to weak monotonicity, and Axiom 2 (SST) replacing Axiom 8 (SSST).

Theorems 3 and 7 characterise the models in Blavatskyy (2008) and Dagsvik (2008, Theorem 4) respectively. Dagsvik's own axiomatisation of his model is quite different to the one given here. He uses a single axiom that he calls *strong independence* in place of Axioms 5–6.[12] It is known that CCI does not imply strong independence (Dagsvik 2015, Example 1), however it is an open question whether strong independence implies CCI. The precise logical relationship between these two axioms remains unclear. Dagsvik also uses an Archimedean alternative to Axiom 3 and the *quadruple condition* (Debreu 1958) in place of the weaker SSST axiom:[13]

Axiom 9 (Quadruple condition (QC)) *For any $a, b, c, d \in A$,*

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

The axiomatic toolkit developed here allows us to set the results of Blavatskyy and Dagsvik upon a common foundation, consisting of Axioms 1, 3, 5 and 6. For Blavatskyy's model we add SST, while for Dagsvik's we add the stronger SSST condition (to obtain the strict monotonicity of G) and also add solvability (to ensure continuity of G).

Just as Axioms 5 and 6 decompose CCI into pieces that attend to linearity of u and F separately, continuity (Axiom 3) and solvability (Axiom 7) attend respectively to the continuity of u and F (or G). While the boundary between scalability and Fechnerian structure is important to understand, there is no obvious merit in separating these two aspects of continuity. In the Appendix (see Theorem 10) we show that it is possible to

[12] Strong independence is the following condition:

$$P(a, b) \geq P\left(\hat{a}, \hat{b}\right) \quad \Rightarrow \quad P(a\lambda c, b\lambda c) \geq P\left(\hat{a}\lambda c, \hat{b}\lambda c\right)$$

for all $a, b, \hat{a}, \hat{b}, c \in A$ and all $\lambda \in [0, 1]$.

[13] More precisely, for a *balanced* binary choice probability, QC implies SSST but not conversely. I have not been able to find a proof of this fact, though one is not difficult to construct. (Details available on request.) One can also piece together a proof from extant material, and doing so reveals an element of confusion on this issue in the literature. Debreu (1958) showed that any balanced BCP which satisfies QC also satisfies the following *monotonicity* condition (Suppes et al. 1989, Chapter 17, Definition 5): for any $a, b, c, a', b', c' \in A$

$$P(a, b) \geq P\left(a', b'\right) \text{ and } P(b, c) \geq P\left(b', c'\right) \quad \Rightarrow \quad P(a, c) \geq P\left(a', c'\right)$$

with strict inequality in the consequent whenever either antecedent inequality is strict. It is straightforward to see that monotonicity implies SSST (given balancedness): take $a' = b' = c'$. Hence, QC implies SSST. Despite the claim in Suppes et al. (1989, Chapter 17 following Theorem 1), monotonicity does not imply QC in general (or even for mixture sets, as any finite set is a mixture set for a suitably defined mixture operation – see Ryan 2010, Corollary 1): a counter-example can be found in the proof of Block and Marschak (1960, Theorem 4.1). It follows that SSST does not imply QC.

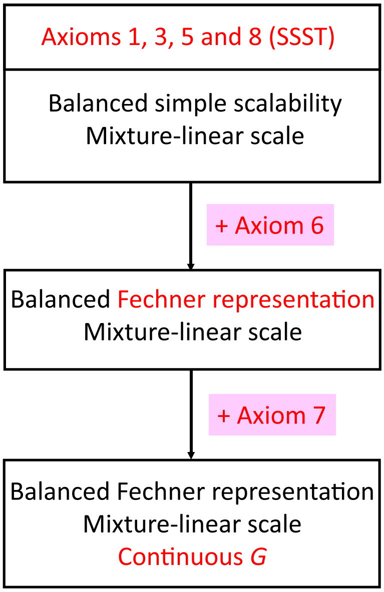

Fig. 1 Summarising theorems 5–7

consolidate Axioms 3 and 7 into a single condition, though at the expense of a mild strengthening of weak independence.

To conclude this section, we briefly discuss Dagsvik's (2008) Theorem 2. This characterises balanced Fechnerian representations of the following form:

$$P(a, b) = G(h(u(a)) - h(u(b))) \tag{9}$$

with G continuous and strictly increasing, $h : u(A) \to \mathbb{R}$ strictly increasing, and u a mixture-linear weak utility. Note that this is a continuous Fechner representation for the weak utility function $h \circ u$. It is still a model of "EU maximisation with noise" but the noise is not Fechnerian unless h is linear. For example, if u is mixture-linear with $u(A) \subseteq (0, \infty)$, $h(z) = \ln(z)$ and

$$G(x) = \frac{1}{1 + \exp(-x)}$$

then (9) is the binary Luce (1959) model:

$$P(a, b) = \frac{u(a)}{u(a) + u(b)}.$$

The following example verifies that P may have a balanced Fechner representation of the form (9) with u mixture-linear, yet not possess a Fechner representation $\left(\hat{u}, \hat{G}\right)$ with $\hat{u}$ mixture-linear.

Example 2 Suppose $u : A \to [0, 1]$ is mixture-linear and surjective, $G(z) = \frac{1}{2}(1 + z)$ and $P(a, b) = G\left(u(a)^2 - u(b)^2\right)$ for any $a, b \in A$. Note that G satisfies (6) for any $z \in \Gamma_u = [-1, 1]$. Hence, P is balanced and any Fechner representation for P must therefore also be balanced. Straightforward calculation gives:

$$P(a\lambda c, b\lambda c) = \frac{1}{2}\left(1 + \lambda^2\left[u(a)^2 - u(b)^2\right]\right) + \lambda(1 - \lambda)u(c)[u(a) - u(b)].$$

It follows that P violates CCI, which is a necessary condition for the existence of a balanced Fechner representation $\left(\hat{u}, \hat{G}\right)$ with $\hat{u}$ mixture-linear (Theorem 1).

How can we bridge from our Theorem 5 to Dagsvik (2008, Theorem 2)? Suppose that P is a balanced BCP and that (u, F) is a simple scale representation for P with u mixture-linear. It follows that if v is a weak utility for P then $v = h \circ u$ for some strictly increasing $h : u(A) \to \mathbb{R}$ (recalling that u is a weak utility for P by Definition 2). Therefore, if P has a Fechner representation (which will necessarily be balanced) it must have a representation of the form (9) with u mixture-linear and G strictly increasing. Necessary and sufficient conditions for the latter representation are therefore obtained by adding to the axioms of Theorem 5 any set of conditions which are necessary and sufficient to ensure the existence of a (balanced) Fechner representation. Many such conditions are known—see Köbberling (2006) for an overview. Classically, Debreu (1958) showed that balance, solvability and QC suffice. They are slightly more than necessary: balance and QC are implied by the existence of a balanced Fechner representation but solvability is only necessary if G is continuous. Recalling that QC implies SSST (see footnote 13), we have therefore shown that strengthening SSST to QC in Theorem 5, and adding solvability, gets us to the representation in Dagsvik (2008, Theorem 2):

Theorem 8 *Let P be a BCP. Then P satisfies Axioms 1, 3, 5, 7 and 9 iff P has a balanced Fechner representation (v, G) with G continuous and $v = h \circ u$ for some mixture-linear $u : A \to \mathbb{R}$ and some strictly increasing $h : u(A) \to \mathbb{R}$.*

Theorem 8 is essentially identical to Dagsvik's (2008) Theorem 2, except that Dagsvik uses his *independence* axiom in place of our *weak independence*. Dagsvik shows that linearity of h can be obtained by strengthening his independence condition to *strong independence* (see footnote 12). It is evident from our analysis that we could, alternatively, add stochastic symmetry while simultaneously weakening QC to SSST—compare Theorems 7 and 8.

Note that bridging from Theorem 5 to a representation of the form (9) with u mixture-linear and G strictly increasing (but not necessarily continuous) is problematic since conditions that are necessary and sufficient for a balanced Fechner representation on general mixture set domains are, to the best of my knowledge, unknown. Debreu (1958, p. 441) lamented that solvability was the "least satisfactory" of his axioms, being excessive to his purpose, but this deficiency is yet to be resolved.[14]

[14] Though Köbberling (2006, Theorem 1) has made some progress in this direction by showing that a weaker form of solvability suffices.

3.4 Preference-bounded domains

In this section we restrict attention to BCPs for which A is bounded with respect to the base relation in the following sense.

Definition 4 The binary choice probability P has a preference-bounded domain if there exist $\underline{a}, \overline{a} \in A$ such that $\overline{a} \succsim^P a \succsim^P \underline{a}$ for all $a \in A$.

For example, if A is the set of lotteries over outcomes in $X = \{x_1, x_2, \ldots, x_n\}$, and if P has a mixture-linear weak utility, then P has a preference-bounded domain: take $\underline{a}$ and $\overline{a}$ to be worst and best outcomes in X respectively (or rather, as the degenerate lotteries with these respective outcomes).

For BCPs with preference-bounded domain we may replace stochastic symmetry with a stochastic dominance condition:

Axiom 10 (Weak transparent dominance [WTD]) *For any* $\beta, \gamma, \lambda, \mu \in [0, 1]$ *with* $\beta - \gamma = \lambda - \mu$,

$$P\left(\overline{a}\beta\underline{a}, \overline{a}\gamma\underline{a}\right) = P\left(\overline{a}\lambda\underline{a}, \overline{a}\mu\underline{a}\right) \tag{10}$$

The WTD axiom is a slight generalisation of Axiom 7 in Ryan (2017). All the lotteries appearing in (10) are mixtures of $\overline{a}$ and $\underline{a}$. The WTD axiom says that the probability of choosing one such mixture over another depends only on the difference between the respective weights on $\overline{a}$. It is intuitive that a mixture with a higher weight on $\overline{a}$ will "dominate" one with a lower weight. If A is the set of lotteries over outcomes in $X = \{x_1, x_2, \ldots, x_n\}$ and P has a mixture-linear weak utility, then this corresponds to a transparent form of first-order stochastic dominance—hence the name of the axiom. A sufficient condition for a balanced binary choice probability, P, to satisfy WTD is that $P\left(\overline{a}\beta\underline{a}, \overline{a}\gamma\underline{a}\right) = 1$ whenever $\beta > \gamma$. In other words, if choice always respects "transparent" dominance, then WTD will be satisfied. This is an intuitively plausible condition, but WTD is even weaker.

Theorem 9 *Let* P *be a BCP with preference-bounded domain that satisfies balanced simple scalability with a mixture-linear scale. Then* P *has a Fechner representation* (u, G) *with* u *mixture linear iff* P *satisfies WTD.*

In other words, given a preference-bounded domain, we can replace stochastic symmetry with WTD in Theorem 6. This provides an alternative axiomatic bridge from strict scalability to a Fechner representation.[15] In fact, this is true even more generally than Theorem 9 suggests: mixture-linearity of the utility scale is not essential. As the proof of Theorem 9 makes clear, it suffices that "probability equivalents" are well-defined and unique for every $a \in A$ (i.e., $a \sim^P \overline{a}\lambda\underline{a}$ for some unique $\lambda \in [0, 1]$) and that the mapping from lotteries to probability equivalents is a weak utility for P. If these conditions hold, and if P satisfies balanced simple scalability and WTD, then P has a Fechner representation.

15 We could likewise replace stochastic symmetry with WTD in Theorem 3 when P has a preference-bounded domain. We omit the details, but the reader will easily be able to supply them by following the proof of Theorem 9.

3.5 Multinomial choice

While Fecherian structures are explicitly limited to binary choice,[16] the notion of scalability was originally formulated for multinomial choice (Krantz 1964). Both of our results on scalability (Theorems 2 and 5) could easily be extended to a multinomial choice context. The only substantive change would be the replacement of SSST (Axiom 8) with Tversky's (1972b) *order independence* axiom, and SST (Axiom 2) with the *multinomial weak substitutability* condition of Ryan (2018b). Axiom 1 would also need to be replaced with the obvious multinomial generalisation, while Axiom 3 and weak independence require no modification other than translation into suitable notation for multinomial choice. See Ryan (2018b) for further details.

4 Concluding remarks

The Fechnerian representations of Blavatskyy (2008) and Dagsvik (2008, Theorem 4) are linear in two senses: utility has an EU form and choice probabilities are linear functions of lottery utilities. Both axiomatisations bundle these two types of linearity into a single axiom: CCI in the case of Blavatskyy (2008) and strong independence in Dagsvik (2008). This has the merit of elegance and efficiency, but obscures the separate foundations of these two very different aspects of the model. This paper provides a suitable decomposition. First, we show that a novel independence-type condition imposed on the base relation (Axiom 5) suffices to underwrite mixture-linearity of utility. Second, we give two additional conditions that each suffice to ensure Fechnerian noise: stochastic symmetry (Axiom 6) and, for BCPs with preference-bounded domains, weak transparent dominance (Axiom 10). An ancillary benefit of our modular axiomatic structure is that it places the results of Blavatskyy (2008) and Dagsvik (2008, Theorem 4) on a common axiomatic footing.

We conclude with some empirical implications of our analysis.

For preference-bounded domains, we have shown that WTD suffices for Fechnerian noise, given simple scalability with respect to a mixture-linear weak utility. This is somewhat surprising, and significant. Consider the case in which A is the set of lotteries over $X = \{x_1, x_2, \ldots, x_n\}$. If subjects' choices always respect "transparent" first-order stochastic dominance, then WTD will be satisfied. It is therefore hard to imagine that empirical tests will be powerful enough to detect violations of WTD, given the very low rates at which transparently dominated options are typically chosen (see Blavatskyy 2012, and references therein).[17] In other words, it would be difficult, empirically, to reject the existence of a strong mixture-linear utility function without also rejecting simple scalability with respect to an EU scale.

The fact that Fechnerian structure is barely more empirically restrictive than simple scalability, at least when the BCP domain is preference-bounded and weak utility is mixture-linear, may be interpreted positively or negatively. On the one hand, it gives

[16] However, see Blavatskyy (2018) for one possible multinomial extension.

[17] It is well-known that *some* lottery pairs can induce a *substantial* proportion of subjects to choose the stochastically dominated option (see, for example, Birnbaum and Navarrete 1998). However, these pairs do not exhibit the sort of transparent dominance entertained here.

a powerful warrant to Fechnerian representations under a maintained hypothesis of simple scalability. On the other hand, it crystallises a classical objection to the latter hypothesis. This objection is neatly captured by Tversky (1972a, p. 284):[18] "Choice probabilities [...] reflect not only the utilities of the alternatives, but also the difficulty of comparing them". If utility is the sole driver of choice, then the reliability with which a dominating option is chosen over a dominated alternative should carry over to any other choice pair with the same utilities (or utility difference in the Fechnerian case), even when no dominance relationship exists. This is intuitively implausible. Simple scalability misses this "context-dependence" of choice probabilities.[19]

For this reason, such models are usually appropriate only when restricted to sets of binary comparisons of similar "difficulty". Otherwise, Theorem 9 indicates the (limited) range of options for evading this tension between scalability and comparability context. If choices always respect transparent first-order stochastic dominance then WTD is satisfied and no problem arises. However, it is easy to see that this means there can be no randomness (non-decisiveness) in *any* binary choice between options with different weak utilities. Essentially, we are back to deterministic choice. To resolve the tension while retaining non-trivial randomness in choice behaviour, transparently dominated options must be chosen with positive probability, and these probabilities must conform with WTD. This is the essential line of demarcation between scalability and Fechnerian structure, when weak utility has the EU form.

Of course, our analysis is restricted to two specific models of "noisy" EU maximisation. The question of whether some form of stochastic EU is an adequate basis for describing, let alone predicting, behaviour, or whether some alternative model is "better", requires a lens with a much broader focus than ours.[20] We have concentrated on one particular boundary in this broader lansdscape, with the aim of drawing it more sharply.

Acknowledgements My thanks to Aurélien Baillon and two anonymous referees for numerous suggestions which have materially improved the paper. I have also benefitted from the comments of audiences at the University of Queensland, Queen Mary University of London, the 37th Australasian Economic Theory Workshop (University of Technology Sydney) and the DECIDE Workshop on Experimental Economics (University of Auckland).

[18] A similar observation had earlier been made by Krantz (1964, pp. 235–236).

[19] Blavatskyy (2011) provides axiomatic foundations for a context-dependent generalisation of the Fechner representation that accommodates dominance relationships.

[20] Stott (2006) and Blavatskyy and Pogrebna (2010) apply such a lens, testing various combinations of "core theory" (i.e., the form of weak utility) and stochastic specification for best fit to a given dataset. The key conclusion to emerge is that both aspects must be considered *jointly*: there is not a "best" core theory, only "best" combinations of core theory and stochastic specification; and estimated utility parameters depend on the noise structure within which the core theory is embedded (see also Bhatia and Loomes 2017, on the latter point).

Appendix

A Proofs for Section 3.2

The following will be needed for the proof of Proposition 1 (and may be of independent interest).

Lemma 1 *Let P be a balanced BCP and (v, F) a simple scale representation for P with $v(A)$ a non-degenerate interval and F continuous in each argument. If P has a representation of the form (3) with u a weak utility and G non-decreasing, then $u = h \circ v$ for some continuous and strictly increasing function $h : v(A) \to \mathbb{R}$.*

Proof Since u and v are both weak utilities for P, there is a strictly increasing function $h : v(A) \to \mathbb{R}$ such that $u = h \circ v$. Thus

$$F(x, y) = G(h(x) - h(y)) \tag{11}$$

for all $x, y \in v(A)$. It remains to show that h is continuous.

Since $v(A)$ is a non-degenerate interval and h is strictly increasing, u is non-constant. As noted in Ryan (2015), it follows that G cannot be constant on any open interval containing 0.[21] We will show that this implies the continuity of h.

Suppose, contrary to what we seek to show, that h is *not* continuous. Then there is some $\{x^n\}_{n=1}^{\infty} \subseteq v(A)$ with $x^n \to \hat{x} \in v(A)$ as $n \to \infty$ but $h(x^n)$ does not converge to $h(\hat{x})$. That is, there exists some $\varepsilon > 0$ such that either $h(x^n) - h(\hat{x}) \geq \varepsilon$ infinitely often, or $h(x^n) - h(\hat{x}) \leq -\varepsilon$ infinitely often. We only consider the former case, as the latter may be handled similarly. Since $x^n \to \hat{x} \in v(A)$ and F is continuous in each argument, we have

$$F(x^n, \hat{x}) \to F(\hat{x}, \hat{x}) = \frac{1}{2} \tag{12}$$

as $n \to \infty$ (where we have used the fact that $F(x, y) + F(y, x) = 1$). Since $h(x^n) - h(\hat{x}) \geq \varepsilon$ infinitely often and G is non-decreasing, it follows from (11) that:

$$F(x^n, \hat{x}) = G\left(h(x^n) - h(\hat{x})\right) \geq G(\varepsilon) \text{ for infinitely many } n \tag{13}$$

Moreover, since G is non-decreasing and satisfies $G(x) + G(-x) = 1$, we have:

$$G(\varepsilon) \geq G(0) = \frac{1}{2}.$$

[21] If it were, then there would exist $a, b \in A$ with $u(a) \neq u(b)$ and

$$P(a, b) = G(u(a) - u(b)) = \frac{1}{2}$$

which contradicts the fact that u is a weak utility for P.

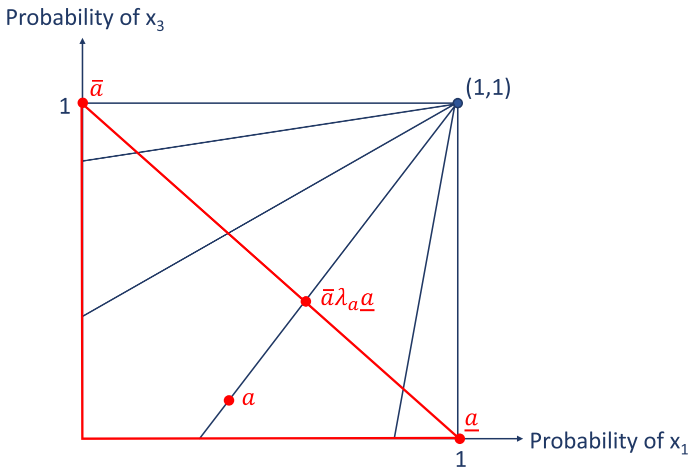

Fig. 2 MM triangle for the proof of Proposition 1

From (12) and (13) we therefore conclude that $G(\varepsilon) = G(0)$; hence

$$G(z) = \frac{1}{2}$$

for all $z \in [0, \varepsilon]$ and

$$G(z) = 1 - G(-z) = \frac{1}{2}$$

for all $z \in [0, -\varepsilon]$. Thus G is constant on $(-\varepsilon, \varepsilon)$, which is the desired contradiction. □

Proof (***Proposition*** 1) We construct a suitable counter-example. Let A be the unit simplex in $\mathbb{R}^3$ (endowed with the usual mixture operation). Think of this as the set of all lotteries over some fixed set $X = \{x_1, x_2, x_3\}$ of "prizes". Figure 2 depicts A in the form of a Marschak–Machina (MM) triangle, with the probability of x_1 measured on the horizontal axis and the probability of x_3 on the vertical. If $b = (b_1, b_2, b_3) \in A$ then we abuse notation and also use b to denote the corresponding point (b_1, b_3) in the triangle. Let $\underline{a} = (1, 0, 0)$ and $\overline{a} = (0, 0, 1)$. The points $\underline{a}$ and $\overline{a}$ are indicated in Fig. 2. (In the following, it will be useful to imagine that x_3 is the best prize and x_1 the worst.)

Note that the line joining any point in the triangle to the point $(1, 1)$ outside the triangle has a unique intersection with the hypotenuse. For each $a \in A$, define $\lambda_a \in [0, 1]$ by the requirement that the line joining the point $a \in A$ to the point $(1, 1)$ passes through $\overline{a}\lambda_a\underline{a}$. Figure 2 illustrates. Now define P as follows:

$$P(a, b) = \begin{cases} \frac{1}{2} + \frac{1}{2}(\lambda_a)^2(\lambda_a - \lambda_b) \text{ if } \lambda_a \geq \lambda_b \\ \\ \frac{1}{2} - \frac{1}{2}(\lambda_b)^2(\lambda_b - \lambda_a) \text{ if } \lambda_a < \lambda_b \end{cases}$$

for each $a, b \in A$. It is easy to see that the range of P is contained in [0, 1] so P is a binary choice probability.

We first verify that P satisfies Axioms 1–3. It is obvious that P is balanced (i.e., satisfies Axiom 1). Note that

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

for any $a, b \in A$. Defining $v : A \rightarrow \mathbb{R}$ by $v(a) = \lambda_a$ and $F : [0, 1]^2 \rightarrow [0, 1]$ by

$$F(x, y) = \begin{cases} \frac{1}{2} + \frac{1}{2}(x)^2(x - y) \text{ if } x \geq y \\ \\ \frac{1}{2} - \frac{1}{2}(y)^2(y - x) \text{ if } x < y \end{cases}$$

we see that (v, F) is a simple scale representation for P. In particular, v is a weak utility for P with $v(A) = [0, 1]$, and F is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument. It follows that P satisfies SST (Tversky and Russo 1969). Since λ_a varies continuously with $a \in A$, and $F(x, y)$ is continuous in x for any y, we deduce that $P(a\lambda b, c)$ varies continuously with λ. Hence, Axiom 3 is satisfied.

In summary: P is a BCP that satisfies Axioms 1–3, and any weak utility for P is a strictly increasing function of v. The contours (indifference curves) of any weak utility for P are therefore the contours of v, which are described by the lines emanating from the point (1, 1) in Fig. 2 (or rather, by the intersections of such lines with the triangle). It is obvious that no such utility function can be mixture-linear. Thus, P has no mixture-linear weak utility.

It remains to show that P has no representation of the form (3) with u a weak utility and G non-decreasing. Suppose it did. Then, by Lemma 1, there exists some strictly increasing and continuous function $h : [0, 1] \rightarrow \mathbb{R}$ such that

$$P(a, b) = F(\lambda_a, \lambda_b) = G(h(\lambda_a) - h(\lambda_b))$$

for all $a, b \in A$. Since F is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument, we claim that G must be strictly increasing on its domain. (We will prove this claim shortly.) Thus:

$$\begin{aligned} P(a, b) \geq P(c, d) \quad &\Leftrightarrow \quad F(\lambda_a, \lambda_b) \geq F(\lambda_c, \lambda_d) \\ &\Leftrightarrow \quad h(\lambda_a) - h(\lambda_b) \geq h(\lambda_c) - h(\lambda_d) \end{aligned}$$

for any $a, b, c, d \in A$. From the equivalence of the first and last of these inequalities, it follows that P must satisfy:

$$P(a, b) = P(c, d) \quad \Leftrightarrow \quad P(a, c) = P(b, d) \tag{14}$$

for any $a, b, c, d \in A$. Let $a, b, c, d \in A$ be chosen such that $\lambda_a = 1$, $\lambda_b = \frac{7}{8}$, $\lambda_c = \frac{1}{2}$ and $\lambda_d = 0$. Then

$$(\lambda_a)^2(\lambda_a - \lambda_b) = \frac{1}{8} = (\lambda_c)^2(\lambda_c - \lambda_d) \quad \Rightarrow \quad F(\lambda_a, \lambda_b) = F(\lambda_c, \lambda_d)$$
$$\Leftrightarrow \quad P(a, b) = P(c, d)$$

but

$$(\lambda_a)^2(\lambda_a - \lambda_c) = \frac{1}{2} < \left(\frac{7}{8}\right)^3 = (\lambda_b)^2(\lambda_b - \lambda_d) \quad \Rightarrow \quad F(\lambda_a, \lambda_c) < F(\lambda_b, \lambda_d)$$
$$\Leftrightarrow \quad P(a, c) < P(b, d)$$

which contradicts (14).

To complete the proof we verify the claim made above. Let $\Sigma = h([0, 1])$ and consider a graph with $h(\lambda_a)$ measured along the horizontal axis and $h(\lambda_b)$ along the vertical. Since h is continuous, Σ is an interval so $\Sigma \times \Sigma$ is a square bisected diagonally by the 45 degree line. See Fig. 3. Let $\hat{G}(x, y) = G(x - y)$ for any $(x, y) \in \Sigma \times \Sigma$. The function $\hat{G}$ is therefore constant along any line parallel to the 45 degree line (or rather, along the portion of such a line that intersects $\Sigma \times \Sigma$). Since P is balanced, it suffices to consider the portion of $\Sigma \times \Sigma$ that lies on or below the 45 degree line and to show that $\hat{G}$ is strictly increasing in $x - y$ on this part of its domain. Let $z \in \Sigma \times \Sigma$ and $z' \in \Sigma \times \Sigma$ be such that $z_1 \geq z_2$ and $z'_1 \geq z'_2$, with $z_1 - z_2 > z'_1 - z'_2$ (so z' lies closer to the 45 degree line than z). If $z_1 > z'_1$ and $z_2 < z'_2$ (i.e., z is strictly southeast of z') then

$$F\left(h^{-1}(z_1), h^{-1}(z_2)\right) > F\left(h^{-1}\left(z'_1\right), h^{-1}\left(z'_2\right)\right)$$

and therefore $\hat{G}(z) > \hat{G}\left(z'\right)$. Otherwise, z must lie to the northeast or southwest of z'. In either case, we can move along the line through z' that is parallel to the 45 degree line to reach some point z'' that is northwest of z and conclude that $\hat{G}(z) > \hat{G}\left(z''\right) = \hat{G}\left(z'\right)$. See Fig. 3. □

Proof (***Proposition*** 2) Suppose, to the contrary, that $a \sim^P b$ but

$$a\frac{1}{2}c \succ^P b\frac{1}{2}c.$$

Then

$$P\left(a\frac{1}{2}c, b\frac{1}{2}c\right) > \frac{1}{2}.$$

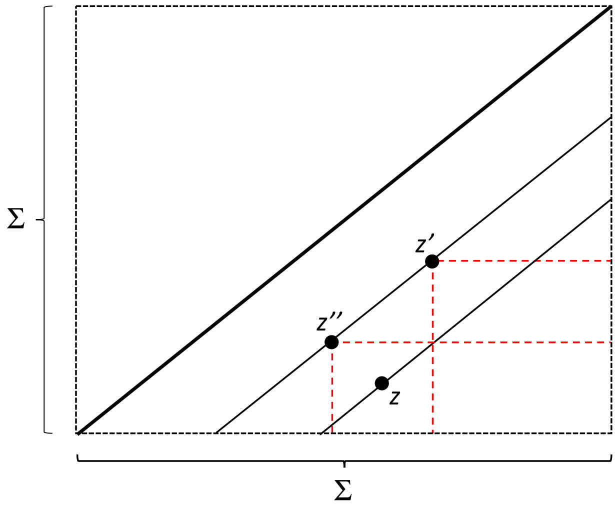

Fig. 3 Verifying strict monotonicity of G

Using Axiom 5 we have

$$\min \left\{ P\left(a, a\frac{1}{2}b\right), P\left(a\frac{1}{2}b, b\right) \right\} > \frac{1}{2} \tag{15}$$

so $P(a, b) \geq \frac{1}{2}$ by WST. If $P(a, b) = \frac{1}{2}$ then balance gives $P(b, a) = \frac{1}{2}$ and we obtain

$$P\left(b, a\frac{1}{2}b\right) \geq \frac{1}{2}$$

from WST and (15). Using balance again we have

$$P\left(a\frac{1}{2}b, b\right) \leq \frac{1}{2}$$

which contradicts (15). We can therefore rule out $P(a, b) = \frac{1}{2}$, so $P(a, b) > \frac{1}{2}$. But this contradicts $a \sim^P b$. □

Proof (***Proposition*** 3) It was shown in text that CCI implies Axioms 5 and 6 . We now develop a counter-example to the converse.

Let A be the unit simplex in $\mathbb{R}^3$ endowed with the usual mixture operation. We interpret A as the set of all lotteries over some fixed set $X = \{x_1, x_2, x_3\}$ of "prizes". We use $\|\cdot\|$ to denote the usual Euclidean norm on $\mathbb{R}^3$:

$$\|z\| = \sqrt{z_1^2 + z_2^2 + z_3^2}.$$

In preparation for defining a suitable BCP, we need two further pieces of notation. First, given any $a, b \in A$ with $a \neq b$, let $D(a, b)$ denote the Euclidean length of the longest line segment that passes through a and b and remains entirely within the simplex. (Think of $D(a, b)$ as the "width" of the simplex along the line through a and b.) Second, let $\geq^*$ be the following lexicographic binary relation on the simplex:

$$a \geq^* b \quad \Leftrightarrow \quad [a_3 > b_3 \text{ or } (a_3 = b_3 \text{ and } a_1 \leq b_1)] .$$

(Once again, imagine that x_3 is the best prize and x_1 the worst.) Let $>^*$ denote the asymmetric part of $\geq^*$, so

$$a >^* b \quad \Leftrightarrow \quad [a_3 > b_3 \text{ or } (a_3 = b_3 \text{ and } a_1 < b_1)] .$$

Note that $\geq^*$ is a linear order (i.e., complete, antisymmetric and transitive). In particular, for any $a, b \in A$ with $a \neq b$, we have $a >^* b$ or $b >^* a$ (but not both).[22] Furthermore, $>^*$ satisfies the following mixture-independence-type condition:

$$a >^* b \quad \Leftrightarrow \quad a\lambda c >^* b\lambda c \tag{16}$$

for any $a, b, c \in A$ and any $\lambda \in (0, 1]$.

Now define P as follows:

$$P(a, b) = \begin{cases} \frac{1}{2} & \text{if } a = b \\ \frac{1}{2} + \frac{1}{2}\left(\frac{\|a-b\|}{D(a,b)}\right) & \text{if } a >^* b \\ \frac{1}{2} - \frac{1}{2}\left(\frac{\|a-b\|}{D(a,b)}\right) & \text{if } b >^* a \end{cases}$$

We will show that P satisfies Axioms 5 and 6 but not CCI.

It is easy to check that P is a balanced BCP and that

$$P(a, b) > \frac{1}{2} \quad \Leftrightarrow \quad a >^* b \tag{17}$$

To see that P satisfies Axiom 5, use (17) and the fact that (16) implies

$$a\lambda c >^* b\lambda c \quad \Leftrightarrow \quad a >^* b \quad \Leftrightarrow \quad a >^* a\mu b \quad \Leftrightarrow \quad a\eta b >^* b \tag{18}$$

for any $a, b, c \in A$ and any $\lambda, \mu, \eta \in [0, 1]$. That P also satisfies Axiom 6 follows from (18) and two further observations: first, that

$$\|a - a\lambda b\| = (1 - \lambda)\|a - b\| = \|b\lambda a - b\| ;$$

[22] Since A is the unit simplex, if $a \neq b$ then $a \in A$ and $b \in A$ differ in at least *two* components.

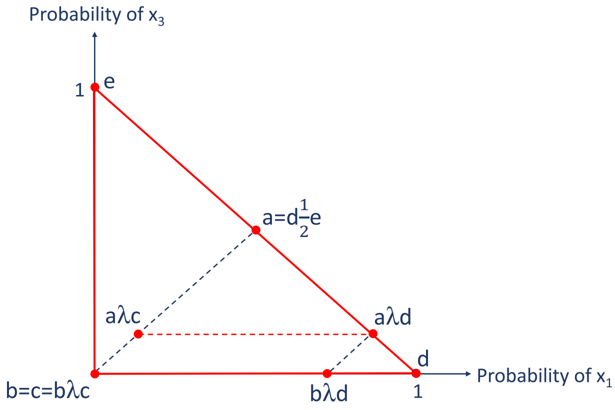

Fig. 4 MM triangle for the proof of Proposition 3. Note that $a >^* b$

and second, that all the vectors in $\{a, b, a\lambda b, b\lambda a\}$ are collinear when $a \neq b$, so $D(a, a\lambda b) = D(b\lambda a, b)$. However, P violates CCI: if $a = \left(\frac{1}{2}, 0, \frac{1}{2}\right)$, $b = c = (0, 1, 0)$ and $d = (1, 0, 0)$, then $a >^* b$ and for any $\lambda \in (0, 1)$ we have

$$P(a\lambda c, b\lambda c) < 1 = P(a\lambda d, b\lambda d).$$

The MM triangle in Fig. 4 illustrates: the probability of x_1 is measured on the horizontal axis and the probability of x_3 on the vertical. □

Proof (***Theorem*** 2) We first show that (i) implies (ii). We start by establishing that $\succsim^P$ has a mixture-linear representation. The argument closely follows Step 1 in the proof of Corollary 2.1 in Ryan (2015). The base relation is complete by Axiom 1 and transitive by strong stochastic transitivity (Axiom 2). Using Axiom 3 we deduce that the sets

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

and

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

are closed for any $a, b, c \in A$. Using Lemma 2 and the fact that SST implies WST we deduce that $\succsim^P$ satisfies the following condition (Fishburn 1982, Axiom B2): for any $a, b, c \in A$

$$a \sim^P b \quad \Rightarrow \quad a\frac{1}{2}c \sim^P b\frac{1}{2}c.$$

Theorem 1 in Fishburn (1982, Chapter 2) now guarantees that P has a mixture-linear weak utility.

Let u be a mixture-linear representation for $\succsim^P$. Using Ryan (2018b, Lemma 11 and Theorem 14) there is a weak utility $\hat{u}$ for P and a function $\hat{F} : \hat{u}(A) \times \hat{u}(A) \to [0, 1]$ that is non-decreasing (respectively, non-increasing) in its first (respectively, second) argument, such that

$$P(a, b) = \hat{F}\left(\hat{u}(a), \hat{u}(b)\right)$$

for all $a, b \in A$. Since u and $\hat{u}$ are both weak utilities for P, we have $\hat{u} = h \circ u$ for some strictly increasing $h : u(A) \to \mathbb{R}$. Defining $F : u(A) \times u(A) \to [0, 1]$ by

$$F(x, y) \; = \; \hat{F}(h(x), h(y))$$

we see that F is non-decreasing (respectively, non-increasing) in its first (respectively, second) argument and (2) holds for all $a, b \in A$. The fact that P is balanced (Axiom 1) ensures that F satisfies (5) for any $x, y \in u(A)$.

We next prove that (ii) implies (i). Axiom 1 follows from the fact that $F(x, y) + F(y, x) = 1$. Axiom 2 (SST) follows from the facts that u represents $\succsim^P$ (Definition 1) and the monotonicity properties of F: if $P(a, b) \geq \frac{1}{2}$ and $P(b, c) \geq \frac{1}{2}$ then $u(a) \geq u(b) \geq u(c)$ so

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

To verify continuity (Axiom 3) we use the mixture-linearity of u and the fact that u represents $\succsim^P$ to deduce

$$P(a\lambda b, c) \geq \frac{1}{2} \;\; \Leftrightarrow \;\; u(a\lambda b) \geq u(c) \;\; \Leftrightarrow \;\; \lambda[u(a) - u(b)] \geq [u(c) - u(b)]$$

and

$$P(a\lambda b, c) \leq \frac{1}{2} \;\; \Leftrightarrow \;\; u(a\lambda b) \leq u(c) \;\; \Leftrightarrow \;\; \lambda[u(a) - u(b)] \leq [u(c) - u(b)].$$

Weak independence (Axiom 5) also follows from the mixture-linearity of u and the fact that u represents $\succsim^P$: if $a\frac{1}{2}c \succ^P b\frac{1}{2}c$ then

$$\frac{1}{2}u(a) + \frac{1}{2}u(c) = u\left(a\frac{1}{2}c\right) > u\left(b\frac{1}{2}c\right) = \frac{1}{2}u(b) + \frac{1}{2}u(c)$$

and hence $u(a) > u(b)$. Therefore

$$u(a) > \frac{1}{2}u(a) + \frac{1}{2}u(b) > u(b)$$

so $a \succ^P a\frac{1}{2}b \succ^P b$. $\square$

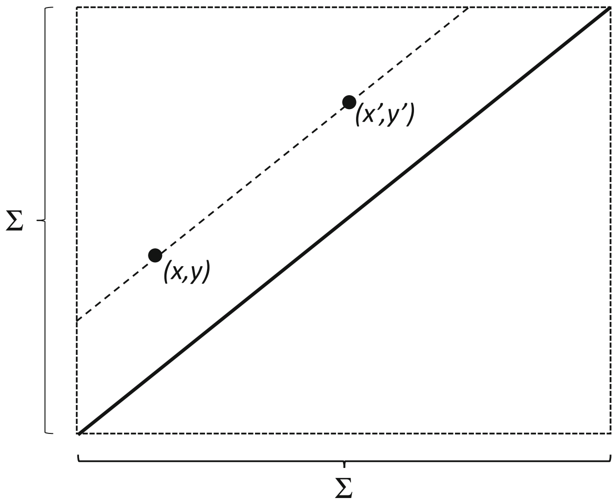

Fig. 5 Domain of F

Proof (***Theorem*** 3) Suppose P satisfies Axioms 1–3 and 5–6. Let (u, F) denote the representation for P from (ii) of the previous theorem. It remains to show that $F(x, y) = F(x', y')$ whenever $x, y, x', y' \in u(A)$ with $x - y = x' - y'$. Let $\Sigma = u(A)$ and recall that Σ is an interval. Figure 5 illustrates the domain of F.[23] Since $F(x, y) + F(y, x) = 1$ for any $x, y \in \Sigma$, the contours of F are symmetric about the 45 degree line (diagonal) in Fig. 5. Moreover, $F(x, x) = \frac{1}{2}$ so any two points on the 45 degree line are contained within the same contour of F. It therefore suffices to show that any two distinct points on a line parallel to, and above, the 45 degree line occupy the same contour of F. Let (x, y) and (x', y') be two such points, so $y - x = y' - x' > 0$. Without loss of generality (WLOG), we assume that $x' > x$ (as in Fig. 5). It follows that $\{y, x'\} \subseteq (x, y')$ with $y = \lambda x + (1 - \lambda) y'$ and $x' = (1 - \lambda) x + \lambda y'$ for some $\lambda \in (0, 1)$.[24] Let $a, b, a', b' \in A$ be such that $x = u(a)$, $y = u(b)$, $x' = u(a')$ and $y' = u(b')$. Then stochastic symmetry and the mixture-linearity of u imply

$$\begin{aligned} F(x, y) &= F\left(x, \lambda x + (1 - \lambda) y'\right) = P\left(a, a\lambda b'\right) \\ &= P\left(b'\lambda a, b'\right) = F\left(\lambda y' + (1 - \lambda) x, y'\right) = F\left(x', y'\right) \end{aligned}$$

23 Figure 5 depicts Σ as if it is bounded but this need not be the case.

24 Let $y = \lambda x + (1 - \lambda) y'$ and $x' = \mu x + (1 - \mu) y'$. Then

$$y - x = (1 - \lambda)\left(y' - x\right)$$

and

$$y' - x' = \mu\left(y' - x\right)$$

so $y - x = y' - x' > 0$ implies $\mu = 1 - \lambda$.

as required.

Conversely, suppose P has a mixture-linear strong utility, u. It suffices, given what was established in Theorem 2, to verify Axiom 6 (stochastic symmetry). This follows straightforwardly from the mixture-linearity of u:

$$
\begin{aligned}
P(a, a\lambda b) = P(b\lambda a, b) &\Leftrightarrow u(a) - u(a\lambda b) = u(b\lambda a) - u(b) \\
&\Leftrightarrow (1-\lambda)[u(a) - u(b)] = (1-\lambda)[u(a) - u(b)]
\end{aligned}
$$

□

Proof (***Theorem*** 4) Suppose P satisfies Axioms 1–3 and 5–7. Let (u, G) be the representation guaranteed by Theorem 3. We must show that G is continuous. Since $u(A)$ is an interval so is the domain of G (i.e., Γ_u). Given that G is non-decreasing, if G were not continuous there would be a gap in the range of G. This would imply a violation of solvability.

Conversely, suppose that (u, G) provides a representation for P of the sort indicated in the theorem. Given Theorem 3 we need only verify solvability. Suppose

$$G(u(a) - u(b)) \geq \rho \geq G(u(a) - u(c)).$$

Now define $h : [0, 1] \to [0, 1]$ by $h(\lambda) = G(u(a) - u(b\lambda c))$ and note that h is continuous (since G is continuous and u mixture-linear) with $h(0) \leq \rho \leq h(1)$. It follows that $h(\lambda) = \rho$ for some $\lambda \in [0, 1]$ by the Intermediate Value Theorem. Hence, $G(u(a) - u(b\lambda c)) = \rho$ for some $\lambda \in [0, 1]$. □

B Proofs for Section 3.3

Proof (***Theorem*** 5) Suppose P satisfies Axioms 1, 3, 5 and 8. The proof of Theorem 2 establishes that P has a mixture-linear weak utility, u. Since P satisfies balance and SSST, P is balanced and possesses a simple scale representation, $\left(\hat{u}, \hat{F}\right)$ (Tversky and Russo 1969). Following the same line of argument as in Theorem 2 it follows that $\hat{u} = h \circ u$ for some strictly increasing $h : u(A) \to \mathbb{R}$ and (u, F) is a simple scale representation for P, where $F : u(A) \times u(A) \to [0, 1]$ is defined as follows:

$$F(x, y) = \hat{F}(h(x), h(y)).$$

Conversely, suppose P is balanced and (u, F) is a simple scale representation for P with u mixture-linear. Theorem 2 ensures that P satisfies Axioms 1, 3 and 5, as well as SST. The monotonicity properties of F ensure that if $u(a) > u(b) > u(c)$, then

$$F(u(a), u(c)) > \max\{F(u(a), u(b)),\ F(u(b), u(c))\}.$$

Since u represents $\succsim^P$, SSST follows (given SST). □

Proof (*Theorem* 6) The result follows by the same argument as for Theorem 3, *mutatis mutandis*, with Theorem 5 used in place of Theorem 2. □

Proof (*Theorem* 7) This follows by the argument used to prove Theorem 4, *mutatis mutandis*, with Theorem 6 used in place of Theorem 3. □

C Theorem 7: Consolidating continuity conditions

We here consider whether the two varieties of continuity axiom in Theorem 7–Axioms 3 and 7—might be consolidated into a single axiom that ensures continuity of u and G simultaneously. This is possible if we strengthen weak independence slightly.
Axiom 5′ *For any* $a, b, c \in A$ *and any* $\lambda \in (0, 1]$,

$$P(a\lambda c, b\lambda c) > \frac{1}{2} \quad \Rightarrow \quad \min\left\{P\left(a, a\frac{1}{2}b\right),\ P\left(a\frac{1}{2}b, b\right)\right\} > \frac{1}{2}.$$

It is straightforward to see Axiom 5′ is an implication of CCI. Moreover, each of our results remains valid if Axiom 5 is replaced with Axiom 5′, since the latter is clearly necessary for the existence of a mixture-linear weak utility.

Given Axiom 5′ we may replace Axioms 3 and 7 in Theorem 7 with the following strengthened form of Axiom 7:
Axiom 7′ (Mixture Solvability) *For any* $a, b, c \in A$ *and any* $\rho \in [0, 1]$ *if*

$$P(a, b) \geq \rho \geq P(a, c)$$

then $P(a, b\lambda c) = \rho$ *for some* $\lambda \in [0, 1]$.

This mixture solvability condition is taken from Ryan (2018a). We may now state our final representation result:

Theorem 10 *Let P be a BCP. Then P satisfies Axioms 1, 5′, 6, 7′ and 8 iff P has a balanced Fechner representation* (u, G) *with u mixture-linear and G continuous.*

Proof The necessity of Axiom 5′ follows by a similar argument to that for the necessity of Axiom 5 in the proof of Theorem 2, and the necessity of mixture solvability by a similar argument to that for the necessity of solvability in the proof of Theorem 4. It remains to verify the sufficiency of the axioms. For this, we need only show that the assumed axioms imply Axiom 3 (continuity), since Axiom 5′ implies Axiom 5 and mixture solvability implies solvability.

We first modify the argument in the proof of Theorem 2 to show that $\succsim^P$ satisfies the following mixture-independence condition: for any $a, b, c \in A$ and any $\lambda \in (0, 1)$:

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

Suppose, to the contrary, that $a \succsim^P b$ and

$$b\lambda c \succ^P a\lambda c$$

for some $\lambda \in (0,1)$. Then Axiom $5'$ implies

$$\min\left\{P\left(b, a\frac{1}{2}b\right), P\left(a\frac{1}{2}b, a\right)\right\} > \frac{1}{2}.$$

Applying SSST we have $b \succ^P a$, which is the desired contradiction.

Finally, given property (19) and mixture solvability (Axiom $7'$), the argument on p. 653 of Ryan (2018a) shows that P satisfies Axiom 3 (Continuity). □

D Proofs for Section 3.4

Proof (***Theorem*** 9) The "only if" part follows by straightforward calculation. For the "if" part let P be balanced and let $\left(\hat{u}, \hat{F}\right)$ be a simple scale representation for P with $\hat{u}$ mixture-linear. If $\overline{a} \sim^P \underline{a}$ the result is trivial, so assume $\overline{a} \succ^P \underline{a}$. It follows that for every $a \in A$, there is a unique $u(a) \in [0,1]$ such that $a \sim^P \overline{a} u(a) \underline{a}$. In particular:

$$u(a) = \left[\frac{1}{v(\overline{a}) - v\left(\underline{a}\right)}\right] v(a) - \left[\frac{v\left(\underline{a}\right)}{v(\overline{a}) - v\left(\underline{a}\right)}\right].$$

Hence $u : A \to [0,1]$ is a positive affine transformation of $\hat{u}$ and therefore a mixture-linear weak utility for P. By a now familiar argument we deduce that there exists some $F : [0,1]^2 \to [0,1]$ such that (u, F) is a simple scale representation for P. It now suffices to show that $F(x, y)$ depends only on $x - y$. Suppose $x, y, \hat{x}, \hat{y} \in [0,1]$ with $x - y = \hat{x} - \hat{y}$. Let

$$a = \overline{a} x \underline{a}$$
$$b = \overline{a} y \underline{a}$$
$$\hat{a} = \overline{a} \hat{x} \underline{a}$$

and

$$\hat{b} = \overline{a} \hat{y} \underline{a}$$

so that $F(x, y) = P(a, b)$ and $F\left(\hat{x}, \hat{y}\right) = P\left(\hat{a}, \hat{b}\right)$. Axiom 10 implies $P(a, b) = P\left(\hat{a}, \hat{b}\right)$, so $F(x, y) = F\left(\hat{x}, \hat{y}\right)$ as required. □

References

Bhatia, S., Loomes, G.: Noisy preferences in risky choice: a cautionary note. Psychol. Rev. **124**(5), 678–687 (2017)

Birnbaum, M.H., Navarrete, J.B.: Testing descriptive utility theories: violations of stochastic dominance and cumulative independence. J. Risk Uncertain. **17**(1), 49–79 (1998)

Blavatskyy, P.R.: Stochastic utility theorem. J. Math. Econ. **44**, 1049–1056 (2008)
Blavatskyy, P.R.: A model of probabilistic choice satisfying first-order stochastic dominance. Manag. Sci. **57**(3), 542–548 (2011)
Blavatskyy, P.R.: Probabilistic choice and stochastic dominance. Econ. Theor. **50**(1), 59–83 (2012)
Blavatskyy, P.: Fechner's strong utility model for choice among $n > 2$ alternatives: risky lotteries, Savage acts, and intertemporal payoffs. J. Math. Econ. **79**, 75–82 (2018)
Blavatskyy, P.R., Pogrebna, G.: Models of stochastic choice and decision theories: why both are important for analyzing decisions. J. Appl. Econom. **25**(6), 963–986 (2010)
Block, H.D., Marschak, J.: Random orderings and stochastic theories of response. In: Olkin, I., Ghurye, S., Hoeffding, W., Madow, W., Mann, H. (eds.) Contributions to Probability and Statistics I. Stanford University Press, Stanford (1960)
Dagsvik, J.K.: Axiomatization of stochastic models for choice under uncertainty. Math. Soc. Sci. **55**, 341–370 (2008)
Dagsvik, J.K.: Stochastic models for risky choices: a comparison of different axiomatizations. J. Math. Econ. **60**, 81–88 (2015)
Davidson, D., Marschak, J.: Experimental tests of a stochastic decision theory. In: Churchman, C.W., Ratoosh, P. (eds.) Measurement: Definitions and Theories. Wiley, New York (1959)
Debreu, G.: Stochastic choice and cardinal utility. Econometrica **26**(3), 440–444 (1958)
Fishburn, P.C.: Binary choice probabilities: on the varieties of stochastic transitivity. J. Math. Psychol. **10**(4), 327–352 (1973)
Fishburn, P.C.: The Foundations of Expected Utility. D. Reidel Publishing, Dordrecht
Fishburn, P.C.: SSB utility theory: an economic perspective. Math. Soc. Sci. **8**(1), 63–94 (1984)
Gul, F., Pesendorfer, W.: Random expected utility. Econometrica **74**(1), 121–146 (2006)
Herstein, I.N., Milnor, J.: An axiomatic approach to measurable utility. Econometrica **21**(2), 291–297 (1953)
Köbberling, V.: Strength of preference and cardinal utility. Econ. Theor. **27**(2), 375–391 (2006)
Krantz, D.H.: The Scaling of Small and Large Color Differences. Ph.D. thesis, University of Pennsylvania (1964)
Luce, R.D.: Individual Choice Behavior: A Theoretical Analysis. Wiley, New York (1959)
Luce, R.D., Suppes, P.: Preference, utility and subjective probability. In: Luce, R.D., Bush, R.B., Galanter, E. (eds.) Handbook of Mathematical Psychology, vol. III. Wiley, New York (1965)
Marschak, J.: Binary choice constraints and random utility indicators. In: Arrow, K.J., Karlin, S., Suppes, P. (eds.) Mathematical Methods in the Social Sciences. Stanford University Press, Stanford (1960)
Roberts, F.S.: Homogeneous families of semiorders and the theory of probabilistic consistency. J. Math. Psychol. **8**(2), 248–263 (1971)
Ryan, M.J.: Mixture sets on finite domains. Decisions Econ. Finan. **33**(2), 139–147 (2010)
Ryan, M.J.: A strict stochastic utility theorem. Econ. Bull. **35**(4), 2664–2672 (2015)
Ryan, M.J.: Binary choices that satisfy stochastic betweenness. J. Math. Econ. **70**, 176–184 (2017)
Ryan, M.J.: Uncertainty and binary stochastic choice. Econ. Theor. **65**(3), 629–662 (2018a)
Ryan, M.J.: Strict scalability of choice probabilities. J. Math. Psychol. **18**, 89–99 (2018b)
Stott, H.P.: Cumulative prospect theory's functional menagerie. J. Risk Uncertain. **32**(2), 101–130 (2006)
Suppes, P., Krantz, D.H., Luce, R.D., Tversky, A.: Foundations of Measurement, vol. II. Academic Press, San Diego (1989)
Tversky, A.: Elimination by aspects: a theory of choice. Psychol. Rev. **79**(4), 281–299 (1972a)
Tversky, A.: Choice by elimination. J. Math. Psychol. **9**, 341–367 (1972b)
Tversky, A., Russo, J.E.: Substitutability and similarity in binary choices. J. Math. Psychol. **6**, 1–12 (1969)