



**AUT**

**Economics Working Paper Series**

*Faculty of Business, Economics and Law, AUT*

**Reconciling Dominance and Stochastic Transitivity  
in Random Binary Choice**

Matthew Ryan

2020/05

# Reconciling Dominance and Stochastic Transitivity in Random Binary Choice

Matthew Ryan\*

School of Economics

Auckland University of Technology

February 24, 2020

## Abstract

Ryan (2017) introduced a condition on binary stochastic choice between lotteries which we call *Weak Transparent Dominance (WTD)*. Consider a binary choice set containing two different mixtures over a “best” and a “worst” possible prize, so that one option transparently dominates the other. The WTD axiom says that the probability of choosing the dominant alternative depends only on the difference in the chance of winning the “best” prize across the two lotteries. A person whose choices always respect first-order stochastic dominance (FOSD) will satisfy this condition, but WTD is a weaker requirement. We show that WTD and strong stochastic transitivity (SST), together with a mild technical condition, ensure the existence of a Fechner model for choice probabilities. This implies, in particular, that for choice probabilities satisfying WTD and our technical condition, there is no observable difference between scalability (Krantz, 1964; Tversky and Russo, 1969) and compatibility with a Fechner model.

**JEL Codes:** D01, D81

**Keywords:** stochastic choice, risk, stochastic dominance, scalability, Fechner

---

\*Email: mryan@aut.ac.nz

# 1 Introduction

Fechnerian models of random binary choice, such as the binary logit model, express choice probabilities as functions of the *utility difference* between the two alternatives. Leonard Savage (*inter alia*) identified a weakness of such models (see Luce and Suppes, 1965, pp.334-337): if utility difference is the sole driver of choice probability, 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 utility difference, even if no dominance relationship exists in the latter case. However, for most standard utility models this is empirically untenable.

When choosing between *lotteries*, the presence of first-order stochastic dominance (FOSD) relationships can therefore be problematic for Fechnerian model fitting. These and other empirical problems motivate the use of “context-dependent” generalisations of the Fechner model (Wilcox, 2008, 2011). For example, Blavatsky (2011) provides axiomatic foundations for a model in which choice probabilities are functions of expected utility differences, but only after conditioning on a context defined by the upper (dominating) and lower (dominated) envelope for the choice pair.<sup>1</sup>

The root of the problem is neatly captured by Tversky (1972, p.284): “Choice probabilities [...] reflect not only the utilities of the alternatives, but also the difficulty of comparing them”. Fechnerian models miss the comparability dimension. As noted by Tversky (1972),<sup>2</sup> the same deficiency afflicts the broader class of *scalable* choice probabilities (Krantz, 1964; Tversky and Russo, 1969). In these models, the probability of choosing one option over another is non-decreasing in the utility of the former and non-increasing in the utility of the latter, irrespective of the relative comparability of the alternatives. The Fechnerian special case arises when the function mapping utilities to choice probabilities is linear – choice probabilities depend only on utility differences.

We explore this tension between scalability and dominance in the context of lottery choice. Our main result shows the following: choice probabilities that respect a weak form of stochastic dominance monotonicity (Axiom 3)<sup>3</sup> and satisfy a mild technical condition (Assumption 1) are scalable if and only if they possess a Fechnerian model. In particular, given our technical assumption, we show that any choice probabilities satisfying our Axiom 3 and strong stochastic transitivity (SST)<sup>4</sup> must have a Fechnerian representation.

---

<sup>1</sup>See Blavatsky (2011) for formal definitions.

<sup>2</sup>A similar observation had earlier been made by Krantz (1967, pp.235-6).

<sup>3</sup>This reproduces Axiom 7 from Ryan (2017).

<sup>4</sup>See Definition 1 below.

In short, if the tension is between stochastic transitivity and our weak form of dominance monotonicity can be resolved, then the distinction between scalability and Fechnerian structure vanishes; if not, then both structures are equally untenable. Stochastic dominance monotonicity is therefore equally problematic for scalability and Fechnerian structure in a rather precise sense: given a weak form of such monotonicity (and our technical condition), *both* models are *characterised* by SST.

## 2 Preliminaries

Let  $A$  be the unit simplex in  $\mathbb{R}^n$ . Points in  $A$  will be interpreted as *lotteries* over a given set  $X = \{x_1, \dots, x_n\}$  of 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  $i \in \{1, 2, \dots, n\}$  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  $\delta_j^i = 0$  for any  $j \neq i$ . Following standard convention, if  $a, b \in A$  and  $\lambda \in [0, 1]$  then  $a\lambda b$  denotes the convex combination  $\lambda a + (1 - \lambda)b$ .

A *binary choice probability (BCP)* is a mapping  $P : A \times A \rightarrow [0, 1]$  that satisfies

$$P(a, b) + P(b, a) = 1 \quad (1)$$

This is called the *balance* or *completeness* condition.<sup>5</sup> If  $a \neq b$ , then  $P(a, b)$  is the probability of choosing  $a$  from the binary choice set  $\{a, b\}$ . No behavioural interpretation is given to  $P(a, b)$  when  $a = b$ , but (1) implies that  $P(a, a) = \frac{1}{2}$  for all  $a \in A$ .

Let  $\mathcal{P}$  denote the set of all BCPs. Associated with any  $P \in \mathcal{P}$  is its *base relation*,  $\succsim^P \subseteq A \times A$ , defined as follows:

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

The binary relations  $\succ^P$  and  $\sim^P$  are determined from  $\succsim^P$  in the usual way.

**Definition 1.** A *weak utility* for  $P$  is a function  $u : A \rightarrow \mathbb{R}$  that represents the base relation: that is,  $u(a) \geq u(b)$  iff  $a \succsim^P b$ .

We will focus on the subset of BCPs that satisfy an auxiliary condition.

**Assumption 1.** There exist  $h, \ell \in \{1, 2, \dots, n\}$  and a continuous weak utility,  $u : A \rightarrow \mathbb{R}$ , for  $P$  such that (I)  $u(\delta^h) \geq u(a) \geq u(\delta^\ell)$  for all  $a \in A$ , and (II)  $u(\delta^h \lambda \delta^\ell)$  is strictly increasing in  $\lambda$ .

---

<sup>5</sup>It is sometimes imposed as an axiomatic restriction on  $P$  but we incorporate it directly into the definition of a BCP for convenience.

Let  $\mathcal{R} \subseteq \mathcal{P}$  be the set of BCPs that satisfy Assumption 1. Note that Assumption 1 restricts only the base relation,  $\succsim^P$ . There are well-known axiomatic conditions on  $\succsim^P$  that are equivalent to Assumption 1. These may be translated into conditions on  $P \in \mathcal{P}$  via (2). The role of Assumption 1 is to ensure that any  $a \in A$  has a unique  $\lambda$  such that  $a \sim^P \delta^h \lambda \delta^\ell$  and that these “probability equivalents” are weak utilities. In particular,  $P \in \mathcal{R}$  if  $P$  possesses a weak utility from almost any standard class of utility functions for lotteries, including EU, Implicit Expected Utility (Dekel, 1986) and Rank-Dependent Expected Utility (Quiggin, 1982).

Finally, we introduce four classes of models for BCPs.<sup>6</sup>

**Definition 2** (Ryan, 2018). *We say that  $P \in \mathcal{P}$  is **strictly scalable** if there exists  $(u, F)$  such that*

$$P(a, b) = F(u(a), u(b))$$

for all  $a, b \in A$ , where  $u$  is a weak utility for  $P$  and  $F : u(A) \times u(A) \rightarrow \mathbb{R}$  is weakly increasing (respectively, weakly decreasing) in its first (respectively, second) argument. In this case, we say that  $P \in \mathcal{P}$  is strictly scalable by  $(u, F)$ .

**Definition 3** (Tversky and Russo, 1969). *We say that  $P \in \mathcal{P}$  is **simply scalable** if it is strictly scalable by some  $(u, F)$  such that  $F$  is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument. In this case, we say that  $P \in \mathcal{P}$  is simply scalable by  $(u, F)$ .*

**Definition 4.** *We say that  $P \in \mathcal{P}$  has a **Fechner model** if it is strictly scalable by some  $(u, F)$  such that  $F$  depends only on utility differences: that is,  $F(x, y) = F(x', y')$  whenever  $x - y = x' - y'$ .*

**Definition 5.** *We say that  $P \in \mathcal{P}$  has a **strong Fechner model** if it is simply scalable by some  $(u, F)$  such that  $F$  depends only on utility differences.*

It is useful to note that if  $P \in \mathcal{P}$  is strictly scalable, and if  $u$  is a weak utility for  $P$ , then  $P$  is strictly scalable by  $(u, F)$  for some  $F : u(A) \times u(A) \rightarrow [0, 1]$ . The same is true if “strictly” is replaced by “simply”.

**Lemma 1.** *Let  $P \in \mathcal{P}$  be strictly (respectively, simply) scalable by  $(u, F)$ . If  $h : u(A) \rightarrow \mathbb{R}$  is strictly increasing and  $\hat{u} = h \circ u$ , then there exists an  $\hat{F} : \hat{u}(A) \times \hat{u}(A) \rightarrow [0, 1]$  such that  $P$  is strictly (respectively, simply) scalable through  $(\hat{u}, \hat{F})$ .*

---

<sup>6</sup>The reader is warned that the terminology for Fechnerian models is not standardised. The terminology in Definitions 4 and 5 is somewhat idiosyncratic to this paper.

**Proof:** The condition

$$\hat{F}(x, y) = F(h^{-1}(x), h^{-1}(y))$$

determines a well-defined function  $\hat{F} : \hat{u}(A) \times \hat{u}(A) \rightarrow [0, 1]$  which shares the same monotonicity properties as  $F$ . Moreover, if  $u$  represents  $\succsim^P$  then so does  $\hat{u}$ .  $\square$

### 3 Main Results

Our main results (Theorems 1 and 2) provide conditions under which  $P \in \mathcal{R}$  possesses a model within one of the four classes. These conditions involve the following axioms, the first two of which are well-known,<sup>7</sup> while the third was introduced (unnamed) in Ryan (2017).

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

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

**Axiom 2** (Strict Stochastic Transitivity [StST]). *For all  $a, b, c \in A$ ,*

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

**Axiom 3** (Weak Transparent Dominance [WTD]). *If  $\delta^h, \delta^\ell \in \{\delta^1, \dots, \delta^n\}$  are such that  $\delta^i \succsim^P \delta^\ell$  and  $\delta^h \succsim^P \delta^i$  for all  $i$ , then*

$$P(\delta^h \alpha \delta^\ell, \delta^h \beta \delta^\ell) = P(\delta^h \lambda \delta^\ell, \delta^h \mu \delta^\ell) \tag{3}$$

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

Only the WTD axiom requires further explanation. Each lottery appearing in (3) is a mixture of a “best” ( $x_h$ ) and a “worst” ( $x_\ell$ ) possible outcome in  $X$ . Thus, the lottery pair on each side of (3) is ordered by “transparent dominance” (Birnbbaum and Navarrete, 1998, p.52): the lotteries in the pair differ (if at all) only in the probability assigned to the “best” outcome. (This motivates the name of Axiom 3.) The WTD axiom says that the probability of choosing one such lottery over another depends only on the difference in the chance of securing the “best” outcome. A sufficient condition for  $P$  to satisfy WTD is that  $P(\delta^h \alpha \delta^\ell, \delta^h \beta \delta^\ell) = 1$  whenever  $\alpha > \beta$ . In other words, if the decision-maker is certain to respect FOSD when choosing between pairs from the set  $\{\delta^h \lambda \delta^\ell \mid \lambda \in [0, 1]\}$  then s/he

---

<sup>7</sup>Our nomenclature follows Fishburn (1973).

satisfies WTD. However, WTD is a much weaker restriction on choice probabilities than this.

We may now prove:

**Theorem 1.** *Suppose  $P \in \mathcal{R}$  satisfies WTD. Then the following are equivalent:*

- (i)  $P$  satisfies SST.
- (ii)  $P$  is strictly scalable.
- (iii)  $P$  has a Fechner model.

**Proof:** That (i) implies (ii) and (iii) implies (i) follows from Ryan (2018, Theorem 14). It remains to show that (ii) implies (iii). Assume, then, that  $P \in \mathcal{R}$  is strictly scalable. Let  $u$  be the weak utility whose existence is guaranteed by Assumption 1. Since  $u$  is continuous and satisfies conditions (I) and (II) of Assumption 1, for every  $a$  there exists a unique  $\langle a \rangle \in [0, 1]$  such that  $u(a) = u(\delta^h \langle a \rangle \delta^\ell)$ , and  $u(a) \geq u(b)$  iff  $\langle a \rangle \geq \langle b \rangle$ . It follows that the function  $v : A \rightarrow [0, 1]$  defined by  $v(a) = \langle a \rangle$  is a(nother) weak utility for  $P$ . By Lemma 1 there is some  $F : v(A) \times v(A) \rightarrow \mathbb{R}$  such that  $P$  is strictly scalable by  $(v, F)$ . We claim that  $F(x, y)$  depends only on  $x - y$ . Suppose  $x - y = \hat{x} - \hat{y}$ . Let

$$a = \delta^h x \delta^\ell$$

$$b = \delta^h y \delta^\ell$$

$$\hat{a} = \delta^h \hat{x} \delta^\ell$$

and

$$\hat{b} = \delta^h \hat{y} \delta^\ell$$

so that  $F(x, y) = P(a, b)$  and  $F(\hat{x}, \hat{y}) = P(\hat{a}, \hat{b})$ . Axiom 3 implies  $P(a, b) = P(\hat{a}, \hat{b})$  as required. Thus,  $P$  has a Fechner model.  $\square$

Ryan (2017, Theorem 2) is a special case of Theorem 1 that imposes additional conditions to ensure the existence of a weak utility of the Implicit Expected Utility form; these conditions imply Assumption 1. Theorem 1 elucidates a more general logic that underpins this specialised result.

By replacing SST with StST, we obtain an analogous result for the strictly monotone models:

**Theorem 2.** *Suppose  $P \in \mathcal{R}$  satisfies WTD. Then the following are equivalent:*

(i)  $P$  satisfies *StST*.

(ii)  $P$  is simply scalable.

(iii)  $P$  has a strong Fechner model.

**Proof:** The proof follows that of Theorem 1, *mutatis mutandis*. That (i) implies (ii) and (iii) implies (i) now follows from Tversky and Russo (1969). The rest of the argument is the same, except that “simply scalable” replaces “strictly scalable” and “strong Fechner” replaces “Fechner”.  $\square$

## 4 Discussion

Theorems 1 and 2 show that the tension between dominance and SST afflicts scalability and Fechnerian models alike, and when  $P \in \mathcal{R}$  the distinction between these models vanishes provided choice probabilities respect transparent dominance in the sense of Axiom 3.

Our results have implications for the prospects of Fechnerian models that embed any standard utility function over lotteries, since almost any such model generates a BCP in  $\mathcal{R}$ . If WTD holds, then validation of such models is equivalent to verifying SST plus the existence of a weak utility of the required form. Moreover, it is hard to imagine that empirical tests will be powerful enough to reject WTD, given the very low rates at which transparently dominated options are typically observed to be chosen (Loomes and Sugden, 1998, p.593).<sup>8</sup> In short, the fate of Fechnerian models of lottery choice rests almost entirely on the empirical validity of SST.

## References

- [1] Birnbaum, M. H., and Navarrete, J. B. (1998). Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty*, 17(1), 49-79.

---

<sup>8</sup>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, I am not aware of any published paper that finds significant violations of FOSD monotonicity in the context of *transparent* dominance (though there is one unpublished exception: Pan, 2019).

- [2] Blavatsky, P.R. (2008). Stochastic utility theorem. *Journal of Mathematical Economics*, 44, 1049–1056.
- [3] Blavatsky, P. R. (2011). A model of probabilistic choice satisfying first-order stochastic dominance. *Management Science*, 57(3), 542-548.
- [4] Debreu, G. (1960). Review of RD Luce, Individual choice behavior: A theoretical analysis. *American Economic Review*, 50(1), 186-188.
- [5] Dekel, E. (1986). An axiomatic characterization of preferences under uncertainty: Weakening the independence axiom. *Journal of Economic Theory*, 40(2), 304-318.
- [6] Fishburn, P.J. (1973). Binary choice probabilities: on the varieties of stochastic transitivity. *Journal of Mathematical Psychology*, 10, 327–352.
- [7] Krantz, D.H. (1964). The scaling of small and large color differences. PhD thesis, University of Pennsylvania.
- [8] Krantz, D. H. (1967). Rational distance functions for multidimensional scaling. *Journal of Mathematical Psychology*, 4(2), 226-245.
- [9] Loomes, G. and Sugden, R. (1998). Testing different stochastic specifications of risky choice. *Economica*, 65(260), 581-598.
- [10] 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.
- [11] Pan, A. (2019). Empirical tests of stochastic binary choice models. Unpublished working paper, University of Auckland.
- [12] Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior & Organization*, 3(4), 323-343.
- [13] Ryan (2017). Binary choices that satisfy stochastic betweenness. *Journal of Mathematical Economics*, 70, 176–184.
- [14] Ryan (2018). Strict scalability of choice probabilities. *Journal of Mathematical Psychology*, 84, 89–99.
- [15] Tversky, A. (1972). Elimination by aspects: A theory of choice. *Psychological Review*, 79(4), 281-299.

- [16] Tversky, A. and Russo, J.E. (1969). Substitutability and similarity in binary choices. *Journal of Mathematical Psychology*, 6, 1–12.
- [17] Wilcox, N.T. (2008). Stochastic models for binary discrete choice under risk: a critical primer and econometric comparison. In J.C. Cox and G.W. Harrison (eds) *Research in Experimental Economics: Risk Aversion in Experiments*, Volume 2, pp.197-292. Emerald, Bingley, U.K.
- [18] Wilcox, N.T. (2011). ‘Stochastically more risk averse’: a contextual theory of stochastic discrete choice under risk. *Journal of Econometrics*, 162(1), 89–104.