Economics Bulletin

Volume 35, Issue 4

A Strict Stochastic Utility Theorem

Matthew Ryan
Auckland University of Technology

Abstract

This note strengthens the sufficiency part of Blavatskyy's (2008) {em Stochastic Utility Theorem} and corrects an error in the necessity part. To do so, we introduce the distinction between a {em stochastic utility representation} and a {em strict stochastic utility representation} for binary choice probabilities.

Citation: Matthew Ryan, (2015) "A Strict Stochastic Utility Theorem", *Economics Bulletin*, Volume 35, Issue 4, pages 2664-2672
Contact: Matthew Ryan - mryan@aut.ac.nz.
Submitted: May 25, 2015. **Published:** December 13, 2015.

1 Introduction

Since the mid-1990s there has been renewed vigour in the debate about the descriptive merits of expected utility (EU). John Hey sounded the call to battle with his conjecture that "one can explain experimental anlayses of decision making under risk better (and simpler) as EU plus noise – rather than through some higher level functional – as long as one specifies the noise appropriately." (Hey, 1995, p.640). A number of other contributors have since weighed in on this empirical question, including (*inter alia*) Ballinger and Wilcox (1997), Loomes and Sugden (1998), Buschena and Zilberman (2000), Butler, Isoni and Loomes (2012), and Loomes and Pogrebna (2014).

Hey's conjecture also signals a shift in the field of battle, from deterministic to stochastic models of choice behaviour. Theorists have rallied to the call by providing axiomatic foundations for models that may be characterised as "EU plus noise". The leading examples are Gul and Pesendorfer (2006), Blavatskyy (2008) and Dagsvik (2008).

The Gul and Pesendorfer (2006) model is of the random utility variety. A decision-maker is characterised by a probability distribution over von Neumann-Morgenstern utility functions. When presented with a choice amongst risky prospects, a utility function is selected at random according to the given distribution and an expected-utility-maximising choice is made.

The models of Blavatskyy (2008) and Dagsvik (2008), by contrast, are "single utility" models in the Fechnerian tradition of psychophysics (Falmagne, 2002). For the case of *binary* choice problems – the sort typically encountered in the experimental literature – a decision-maker is characterised by a single von Neumann-Morgenstern utility function u plus an auxiliary function f that converts expected utility *differences* into choice probabilities. When presented with the binary choice $\{a, b\}$, where a and b are risky prospects, the probability with which the decision-maker chooses alternative a, denoted $P(a, b)$, is determined as follows:

$$P(a,b) = f(u(a) - u(b)).$$

She chooses b with the complementary probability $P(b, a) = 1 - P(a, b)$. The function f is required to be non-decreasing and to satisfy $f(x) + f(-x) = 1$. It follows that the decision-maker chooses the alternative with the higher expected utility at least half of the time, and this likelihood is weakly increasing in the difference between the higher and the lower expected utility.

The essential difference between Blavatskyy (2008) and Dagsvik (2008) is that the latter requires f to be *strictly* increasing, in the spirit of Debreu (1958), while Blavatskyy (2008) does not. The two sets of axioms are also very different.[1]

The purpose of this note clarify and strengthen the representation theorem of Blavatskyy

[1] Dagsvik (2008) also considers stochastic models of multinomial choice, while Blavatskyy (2008) confines attention to the binary case.

(2008). First, we observe that one of Blavatskyy's axioms is redundant (i.e., it is implied by the rest). Second, we show that the necessity of Blavatskyy's axioms only obtains if f is required to satisfy an additional property – it should not be constant on any neighbourhood of 0 unless u is constant. Finally, we introduce the distinction between a *stochastic utility representation* and a *strict stochastic utility representation* to emphasise the behavioural implications of this new restriction on f.

2 A Strict Stochastic Utility Theorem

Let A denote the unit simplex in $\mathbb{R}^n$, interpreted as the set of lotteries over the outcomes in $X = \{x_1, ..., x_n\}$. Following convention, we write $a\lambda b$ for $\lambda a + (1-\lambda)\, b$ whenever $a, b \in A$ and $\lambda \in [0,1]$.

Consider binary choice problems in which the pairs of alternatives are drawn from the set A. Since choice behaviour may exhibit randomness, each decision-maker will be characterised by a collection of binary choice probabilities rather than a preference relation. A *binary choice probability function* is a mapping

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

If $a \neq b$, the quantity $P(a,b)$ is the probability (or, in behavioural terms, the expected frequency) with which the decision-maker selects a when given the choice between a or b. No behavioural interpretation is given to $P(a,b)$ when $a = b$, but it is conventional to define binary choice probability functions on the entire Cartesian product $A \times A$ for convenience.

Given a binary choice probability function P, it is natural to induce the following binary relation on A: for any $a, b \in A$:

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

We will call $\succsim^P$ the decision-maker's *stochastic preference relation*. Thus, a is weakly stochastically preferred to b iff the decision-maker is expected to choose a over b at least as frequently as she is expected to choose b over a. We define $\succ^P$ and $\sim^P$ from $\succsim^P$ in the usual way.

Blavatskyy (2008) says that $P : A \times A \to [0,1]$ has a *stochastic utility representation* if there exists a linear utility function $u : A \to \mathbb{R}$ and a non-decreasing function $f : \mathbb{R} \to \mathbb{R}$ satisfying $f(x) + f(-x) = 1$ for all $x \in \mathbb{R}$ such that

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

for any $a, b \in A$.

Note that $f(x) + f(-x) = 1$ implies $f(0) = \frac{1}{2}$, so $P(a,b) \geq \frac{1}{2}$ when $u(a) \geq u(b)$ and $P(a,b) \leq \frac{1}{2}$ when $u(a) \leq u(b)$. In other words, if P has a stochastic utility representation then the decision-maker maximises expected utility with error: she chooses the alternative with the higher expected utility with probability at least $\frac{1}{2}$ though not necessarily with certainty. Moreover, the probability with which the expected-utility-maximising choice is made is weakly increasing in the difference between the higher and the lower expected utility.

Since f need not be *strictly* increasing, it is possible that $P(a,b) = \frac{1}{2}$ even if $u(a) > u(b)$. This has the unfortunate consequence that u need not represent (in the usual sense) the stochastic preference relation $\succsim^P$. To see why, observe that $f(x) + f(-x) = 1$ implies

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

and hence

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

We may therefore have $a \sim^P b$ (i.e., $P(a,b) = \frac{1}{2}$) when $u(a) > u(b)$. This happens if f is constant in some neighbourhood of 0 but u is non-constant.

Let us say that P has a *strict stochastic utility representation* if it has a stochastic utility representation (2) such that u represents $\succsim^P$. This is equivalent to saying that P has a stochastic utility representation such that either f is non-constant on any open neighbourhood of 0 or u is constant.

Note that the function f in a strict stochastic utility representation need not be continuous, even at 0. Indeed, there are important discontinuous examples, such as the Harless and Camerer (1994) model of constant errors.

One half of Blavatskyy's *Stochastic Utility Theorem* (Blavatskyy, 2008, Theorem 1) establishes that the following five conditions are sufficient for $P : A \times A \to [0,1]$ to possess a stochastic utility representation:

Axiom 1 (Completeness) *For any $a, b \in A$, we have $P(a,b) = 1 - P(b,a)$.*

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

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

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

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) *For any $a, b, c, d \in A$ and any $\lambda \in [0, 1]$*

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

Axiom 5 (Interchangeability) *For any $a, b \in A$,*

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

In fact, a careful reading of Blavatskyy's proof reveals a slightly stronger result: he shows that Axioms 1-5 suffice for a *strict* stochastic utility representation. We may strengthen it further by the following observation:

Lemma 2.1 *Axiom 5 is implied by Axioms 1 and 2.*

Proof. Davidson and Marschak (1959, p.240) prove that, given Axiom 1, strong stochastic transitivity (Axiom 2) is equivalent to the following *weak substitutability* property:[2] for any $a, b, c \in A$,

$$P(a, b) \geq \frac{1}{2} \quad \Rightarrow \quad P(a, c) \geq P(b, c) \tag{4}$$

Axiom 5 is a direct implication of weak substitutability, and hence of Axioms 1 and 2. □

As a Corollary we have the following strengethened version of the "sufficiency" part of Blavatskyy (2008, Theorem 1).

Corollary 2.1 *If P is satisfies Axioms 1-4 then it has a strict stochastic utility representation.*

Although Corollary 2.1 follows directly from Blavatskyy (2008) and Lemma 2.1, we give a proof in the Appendix.

The other half of Blavatskyy's *Stochastic Utility Theorem* asserts that P has a stochastic utility representation *only if* it satisfies Axioms 1-5. This is not quite correct. A binary choice probability function with a *non-strict* stochastic utility representation may violate Axioms 2, 3 and 5.

Example 2.1 *Suppose that u is linear on A with $u(A) = [0, 1]$ and $f : \mathbb{R} \to \mathbb{R}$ is the function*

$$f(x) = \begin{cases} \frac{3}{4} & \text{if } x \geq \frac{1}{2} \\ \frac{1}{2} & \text{if } x \in \left(-\frac{1}{2}, \frac{1}{2}\right) \\ \frac{1}{4} & \text{if } x \leq -\frac{1}{2} \end{cases}$$

[2] If we replace "⇒" with "⇔" in (4) we obtain the *substitutability* condition (Tversky and Russo, 1969).

Note that f is non-decreasing and satisfies $f(x)+f(-x)=1$. Let $P(a,b)=f(u(a)-u(b))$ for all $a,b\in A$. Then P is a binary choice probability function with a (non-strict) stochastic utility representation. Let $a,b,c\in A$ with $u(a)=\frac{7}{8}$, $u(b)=\frac{5}{8}$ and $u(c)=\frac{1}{4}$. We have $P(a,b)=\frac{1}{2}$ and

$$P(a,c)=\frac{3}{4}>\frac{1}{2}=P(b,c)$$

which contradicts Axiom 5. We also have $P(c,b)=P(b,a)=\frac{1}{2}$ and $P(c,a)=\frac{1}{4}$ which contradicts Axiom 2. Finally:

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

This is a violation of Axiom 3.

To ensure that Axioms 1-5 (equivalently, Axioms 1-4) hold, the function f in a stochastic utility representation must not be constant in any open neighbourhood of 0, unless u is constant. In other words, the stochastic utility representation must be *strict*. The following is proved in the Appendix.

Theorem 2.1 *The binary choice probability function P has a strict stochastic utility representation iff P satisfies Axioms 1-4.*

Theorem 2.1 is the main result of the paper. It strengthens the sufficiency part and corrects the necessity part of Blavatskyy (2008, Theorem 1). In short, Blavatskyy's (2008) analysis applies to what we have called *strict stochastic utility representations*. Obtaining necessary and sufficient conditions for a stochastic utility representation remains an open question. However, it also appears to be a less interesting one, as the inconsistency between $\succsim^P$ and the utility function in a stochastic utility representation is conceptually unappealing.

3 Concluding remarks

To summarise, we have established three results. First, Axiom 5 (Interchangeability) is redundant to Theorem 1 in Blavatskyy (2008). Second, the existence of a stochastic utility representation does *not* imply Axioms 1-5. Hence, we introduce the notion of a *strict* stochastic utility representation, in which u represents $\succsim^P$. Finally, we establish that Axioms 1-4 are necessary and sufficient for the existence of a strict stochastic utility representation.

References

Ballinger, T.P. and N.T. Wilcox (1997) "Decisions, Error and Heterogeneity" *The Economic Journal* **107**, 1090-1105.

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

Buschena, D. and D. Zilberman (2000) "Generalized Expected Utility, Heteroscedastic Error, and Path Dependence in Risky Choice" *Journal of Risk and Uncertainty* **20(1)**, 67-88.

Butler, D., A. Isoni and G. Loomes (2012) "Testing the 'Standard' Model of Stochastic Choice under Risk" *Journal of Risk and Uncertainty* **45**, 191–213.

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

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

Debreu, G. (1958), "Stochastic Choice and Cardinal Utility," *Econometrica* **26(3)**, 440–444.

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

Fishburn, P.J. (1982) *The Foundations of Expected Utility*, D. Reidel Publishing: Dordrecht.

Gul, F. and W. Pesendorfer (2006) "Random Expected Utility" *Econometrica* **74(1)**, 121–146.

Harless, D. and C. Camerer (1994) "The Predictive Utility of Generalized Expected Utility Theories" *Econometrica* **62**, 1251–1289.

Hey, J.D. (1995) "Experimental Investigations of Errors in Decision Making under Risk" *European Economic Review* **39**, 633–640.

Loomes, G. and G. Pogrebna (2014) "Testing for Independence while Allowing for Probabilistic Choice" *Journal of Risk and Uncertainty* **49**, 189–211.

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

Tversky, A. and J.E. Russo (1969) "Substitutability and Similarity in Binary Choices" *Journal of Mathematical Psychology* **6**, 1–12.

Appendix

Proof of Corollary 2.1. We break the proof into three steps.

Step 1. *In which Axioms 1-4 are used to show that the binary relation $\succsim^P$ has a linear representation.*

Let $\succsim^P$ be the binary relation on A derived from P using (3). This 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 $a, b, c \in A$. Finally, one may show that $\succsim^P$ satisfies the following independence condition: for any $a, b, c \in A$

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

Suppose, to the contrary, that $a \sim^P b$ but

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

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

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

Using common consequence independence (Axiom 4) twice we have

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

and

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

Applying strong stochastic transitivity we deduce the required contradiction:

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

Theorem 1 in Fishburn (1982, Chapter 2) therefore implies that $\succsim^P$ has a linear representation $u : A \to \mathbb{R}$, unique up to positive affine transformations. If u is constant, a strict stochastic utility representation follows trivially: define $f(x) = \frac{1}{2}$ for all x. Otherwise, we may assume that $u(A) = [0,1]$ and we do so for the remainder of the proof. We also define δ_i to be the i^{th} unit vector in $\mathbb{R}^n$ (i.e., the degenerate lottery that gives x_i with probability 1) and we assume that the elements of X are ordered such that $\delta_1 \succsim^P \delta_2 \succsim^P \cdots \succsim^P \delta_n$. It follows that $u(\delta_1) = 1$ and $u(\delta_n) = 0$.

Step 2. *In which Axioms 1 and 2 are used to show that P can be re-expressed as a function of the utilities of alternatives.*

The weak substitutability condition (4), together with Axiom 1, give

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

for any $a, b, c \in A$. It follows that there exists a function $\pi : [0,1]^2 \to [0,1]$ such that $P(a,b) = \pi(u(a), u(b))$ for any $a, b \in A$. Weak substitutability and Axiom 1 further imply that π is non-decreasing in its first argument, non-increasing in its second and satisfies $\pi(x,y) = 1 - \pi(y,x)$.

Step 3. *In which Axiom 4 is used to show that $\pi(x,y)$ depends only on $x - y$.*

Given any $k \in [0,1]$ note that

$$\begin{aligned} \left\{ (x,y) \in [0,1]^2 \;\middle|\; x - y = k \right\} &= \{ (y+k, y) \mid y \in [0, 1-k] \} \\ &= \{ (k + (1-k) z,\ (1-k) z) \mid z \in [0,1] \} \\ &= \{ k(1,0) + (1-k)(z,z) \mid z \in [0,1] \} \end{aligned}$$

From the linearity of u and the fact that $u(A) = [0,1]$ we have

$$P(k(\delta_1, \delta_n) + (1-k)(c,c)) = \pi(k(1,0) + (1-k)(u(c), u(c)))$$

for any $c \in A$. Common consequence independence implies that this quantity is independent of c. Thus, for any $k \in [0,1]$, the function π is constant on the set

$$\left\{ (x,y) \in [0,1]^2 \;\middle|\; x - y = k \right\} \tag{5}$$

Since $\pi(x,y) = 1 - \pi(y,x)$, it follows that π is also constant on (5) for any $k \in [-1, 0]$. From Step 2, we know that the value of π on (5) is non-decreasing in k. The stochastic utility representation now follows directly: define $f : [-1,1] \to [0,1]$ by $f(k) = \pi(x,y)$ for any $(x,y) \in [0,1]^2$ with $x - y = k$ and define f on $(-\infty, -1) \cup (1, \infty)$ in any fashion that ensures f is non-decreasing and satisfies $f(x) + f(-x) = 1$. Since u represents $\succsim^P$, this is a strict stochastic utility representation.

This completes the proof of Corollary 2.1. $\square$

Proof of Theorem 2.1. The "if" part is Corollary 2.1. We prove the "only if" part here.

Completeness (Axiom 1) is implied by $f(x) + f(-x) = 1$.

Axiom 2 (strong stochastic transitivity) follows from the facts that u represents $\succsim^P$ and f is non-decreasing: the former ensures $u(a) \geq u(b)$ whenever $P(a, b) \geq \frac{1}{2}$, and the latter implies

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

for all $x \geq 0$ and $y \geq 0$.

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

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

and

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

Finally, common consequence independence (Axiom 4) follows from the linearity of u, since

$$u(a\lambda c) - u(b\lambda c) = \lambda[u(a) - u(b)]$$

for any $c \in A$. $\square$