School of Economics

Working Paper Series

Representation of Binary Choice Probabilities. Part I: Scalability

Matthew Ryan

2016/04

Representation of Binary Choice Probabilities. Part I: Scalability

Matthew Ryan
School of Economics
Auckland University of Technology

September 25, 2016

Abstract

Scalability refers to the existence of a utility scale on alternatives, with respect to which binary choice probabilities are suitably monotone. This is a fundamental concept in psychophysical theory (Falmagne, 1985). We introduce a new notion of scalability which we call *strict scalability*, and establish axiomatic foundations for this concept. Strict scalability lies between the classical notion of simple scalability, which was axiomatised by Tversky and Russo (1969), and the weaker notion of monotone scalability, which was axiomatised by Fishburn (1973). When the set of alternatives is countable, a binary choice probability is strictly scalable if and only if it satisfies the familiar condition of weak substitutability.

1 Introduction

These notes explore the axiomatic foundations of classical representations of binary choice probabilities. In Part I, we examine scalability – the existence of a utility scale on alternatives, with respect to which binary choice probabilities are suitably monotone. Our primary purpose is to fill an unnoticed gap between the characterisations of monotone scalability by Fishburn (1973) and simple scalability by Tversky and Russo (1969). Between these two notions lies what we call *strict scalability.* A binary choice probability is strictly scalable if it is monotone scalable via a utility function that represents (in the usual sense) the associated stochastic preference relation.[1] When alternatives are drawn from a countable set, we show that binary choice probabilities are strictly scalable if and only if they satisfy the well-known strong stochastic transitivity condition – equivalently, the weak substitutability condition (Davidson and Marschak, 1959). Given the importance of these equivalent axioms in the literature on probabilistic choice, it is useful to understand their implications for the representation of binary choice probabilities.

In Part II, we study the axiomatic foundations of Fechnerian representations.

2 Binary choice probabilities

A *complete binary choice specification (CBCS)* is a pair (A, P), where A is a non-empty set of alternatives and P is a *binary choice probability (BCP)*. The latter is a mapping

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

that satisfies

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

for all $a, b \in A$. Hence, by definition,

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

for any $a \in A$. If $a \neq b$, then $P(a, b)$ is the probability (frequency) with which the decision-maker selects a when (given repeated opportunities of) choosing between a or b, abstention not being an option.

If A is finite, we can let $A = \{a_1, a_2, ..., a_n\}$ so that P corresponds to the matrix

$$\boldsymbol{P} = \begin{bmatrix} P(a_1, a_2) & P(a_1, a_2) & \cdots & P(a_1, a_n) \\ P(a_2, a_1) & P(a_2, a_2) & \cdots & P(a_2, a_n) \\ \vdots & \vdots & \ddots & \vdots \\ P(a_n, a_1) & P(a_n, a_2) & \cdots & P(a_n, a_n) \end{bmatrix} \tag{2}$$

This matrix satisfies $\boldsymbol{P} + \boldsymbol{P}^T = \mathbf{1}$, where $\mathbf{1}$ is a matrix with 1 in every cell.

The "completeness" qualifier refers to the fact that the domain of P is the entire Cartesian product $A \times A$, so the domain of the CBCS can be characterised by the set of alternatives, A. For a CBCS it is conventional to consider representations in terms of a utility function, $u : A \to \mathbb{R}$, defined

[1] We say that one alternative is weakly stochastically preferred to another if the former is chosen over the latter with probability at least $\frac{1}{2}$.

on alternatives – see Section 3. The best known such representation is the classical Fechner model, in which $P(a,b)$ is a non-decreasing function of the utility difference $u(a) - u(b)$ for some suitable choice of utility scale.

More generally, one might imagine a BCP defined on some subset $B \subseteq A \times A$. Consider, for example, the set

$$\Delta(A,P) = \{(a,b) \in A \times A \mid 0 < P(a,b) < 1\}.$$

If $(a,b) \in \Delta(A,P)$ then neither alternative is *absolutely preferred* to the other (Davidson and Marschak, 1959, Definition 1). Some analyses restrict attention to CBCS's defined on $\Delta(A,P)$, or assume that $\Delta(A,P) = A \times A$ (see, for example, Davidson and Marschak, 1959; Tversky and Russo, 1969).

One rationale for such a restriction is to avoid a well-known problem with classical Fechner models. These models cannot explain the absolute preference for a dominating option over a transparently dominated alternative (Luce and Suppes, 1965, p.334). Dominance induces absolute preference irrespective of utility differences.

Nevertheless, we study complete BCPs without assuming that $\Delta(A,P) = A \times A$, since we do not wish to exclude scenarios in which absolute preference arises for reasons which *can* be adequately captured by utility differences. At a formal level, if there exist pairs $(a,b) \in A \times A$ with transparent dominance relationships, then we assume that (A,P) describes a hypothetical extension of $P : B \to [0,1]$ for some $B \subset A \times A$ which excludes such pairs (as is often done in experimental settings). For example, if there is a Fechner model for $P : B \to [0,1]$, then this model can be used to obtain the desired extension. Note, however, that many of the properties of P used in our analysis implicitly impose a minimal degree of completeness on B. One should be alert to steps in proofs which rely upon the existence of particular elements from $A \times A$ in B.

Given P, it is natural to impute the following *stochastic preference relation* on A: for any $a, b \in A$,

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

where the second equivalence follows from (1). In other words, a is "weakly stochastically preferred" to b iff the decision-maker chooses a over b at least half of the time. The asymmetric and symmetric parts of $\succsim^P$ are denoted $\succ^P$ and $\sim^P$ respectively, and satisfy

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

and

$$a \sim^P b \quad \Leftrightarrow \quad P(a,b) = \frac{1}{2}.$$

Note that $\succsim^P$ is complete by construction but need not be transitive. It will be transitive iff (A, P) satisfies "weak stochastic transitivity":

Definition 1 *A CBCS satisfies* ***weak stochastic transitivity (WST)*** *if*

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

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

Fishburn (1973) introduced another useful binary relation that can be constructed from P:

$$a \succsim_0^P b \quad \Leftrightarrow \quad P(a,c) \geq P(b,c) \quad \text{for any } c \in A \tag{4}$$

It is clear that $\succsim_0^P$ is transitive, though it need not be complete. Fishburn (1973, Theorem 1) proves that it is complete iff P satisfies a condition he calls *weak independence* (Definition 11 below), which is logically independent of WST – see Fishburn (1973, Figure 1). Even if P satisfies WST *and* weak independence, so that $\succsim^P$ and $\succsim_0^P$ are both weak orders, it need not be the case that $\succsim^P = \succsim_0^P$ (Corollary 22). A necessary and sufficient condition for these two binary relations to coincide will be provided later (Corollary 20).

3 Representations

"Scalability" refers (loosely speaking) to the possibility of representing binary choice probabilities using a "utility scale" defined on the set of alternatives. The following definitions describe three increasingly demanding notions of scalability.

Definition 2 *A CBCS satisfies* ***monotone scalability (MS)*** *if there exist functions $u : A \to \mathbb{R}$ and $F : u(A) \times u(A) \to [0,1]$ such that F is non-decreasing (respectively, non-increasing) in its first (respectively, second) argument, and*

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

for all $a, b \in A$. In this case, we say that (A, P) is ***monotone scalable*** *through (u, F).*

Definition 3 *A CBCS satisfies* ***strict scalability (StS)*** *if there exists a function* $u : A \rightarrow \mathbb{R}$ *and a function* $F : u(A) \times u(A) \rightarrow [0,1]$ *that is non-decreasing (respectively, non-increasing) in its first (respectively, second) argument, such that* u *represents* $\succsim^P$ *(i.e.,* $a \succsim^P b$ *iff* $u(a) \geq u(b)$ *for any* $a, b \in A$*) and*

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

for all $a, b \in A$*. In this case, we say that* (A, P) *is* ***strictly scalable*** *through* (u, F)*.*

Definition 4 *A CBCS satisfies* ***simple scalability (SS)*** *if there exist functions* $u : A \rightarrow \mathbb{R}$ *and* $F : u(A) \times u(A) \rightarrow [0,1]$ *such that* F *is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument, and*

$$P(a,b) = F(u(a), u(b)) \tag{5}$$

for all $a, b \in A$*. In this case, we say that* (A, P) *is* ***simply scalable*** *through* (u, F)*.*

The SS concept is classical in the literature on BCPs. The MS concept is less well-known but of long standing – it was introduced by Fishburn (1973). The StS concept is defined for the first time here (so far as we are aware). It rules out the possibility that there exist $a, b \in A$ with $u(a) > u(b)$ but $P(a,b) = \frac{1}{2}$. In other words, it rules out imperfect utility discrimination – the possibility that the utility difference between a and b cannot be detected from the frequency of choices from $\{a, b\}$.

Of course, StS is still compatible with the possibility that $u(a) > u(b) > u(c)$ but there is no difference between the frequency with which a is chosen from $\{a, c\}$ and the frequency with which b is chosen from $\{b, c\}$. In other words, the decision maker may not discriminate between a and b *in comparison with* c. The SS concept imposes this "indirect discrimination" as well. In particular, the following result verifies that StS is intermediate between MS and SS.

Lemma 5 *If* (A, P) *is simply scalable through* (u, F) *then* u *represents* $\succsim^P$*.*

Proof: Since F is strictly increasing (respectively, strictly decreasing) in its first (respectively, second) argument, we have

$$\begin{aligned} P(a,b) \geq P(b,a) \quad &\Leftrightarrow \quad F(u(a), u(b)) \geq F(u(b), u(a)) \\ &\Leftrightarrow \quad u(a) \geq u(b) \end{aligned}$$

for all $a, b \in A$. It follows from (3) that u represents $\succsim^P$. □

Note that (1) imposes restrictions on F in each of these representations. If (A,P) satisfies MS (or StS or SS), then

$$F(x,y)+F(y,x)=1$$

for all $x,y\in u(A)$. Hence, $F(x,x)=\frac{1}{2}$ for any $x\in u(A)$. The requirement that u represents $\succsim^P$ places an additional implicit restriction on F. For StS to be satisfied we must have

$$F(x,y)>[<]\frac{1}{2}\quad\Leftrightarrow\quad x>[<]y$$

for any $x,y\in u(A)$ with $x\neq y$.

The rest of this paper examines the axiomatic foundations of these three notions of scalability. We review known results in the case of MS and SS, but our axiomatisation of StS is new.

Before doing so, we note that a *Fechnerian* representation is obtained if a BCP is scalable and choice probabilities can be expressed in terms of utility *differences*. That is, there is a utility scale $u:A\to\mathbb{R}$ and a non-decreasing function $F:u(A)\to[0,1]$ such that

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

for all $a,b\in A$. For each of our three notions of scalability there is a Fechnerian analogue. (The definition just given is the analogue of MS.) The axiomatisation of these Fechnerian models, and the axiomatic gap between each notion of scalability and its Fechnerian cousin, are issues that we explore in Part II of these notes.

Before proceeding to the axioms, we make a couple of observations which will be useful in the sequel. The first is the following fact:

Lemma 6 *Let (A,P) be simply (strictly) [monotone] scalable through (u,F). If $h:u(A)\to\mathbb{R}$ is strictly increasing and $\hat{u}=h\circ u$, then there exists an $\hat{F}$ such that (A,P) is simply (strictly) [monotone] scalable through $\left(\hat{u},\hat{F}\right)$.*

Proof: The condition

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

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

Second, if A is finite and (A, P) is monotone scalable, then we may enumerate

$$A = \{a_1, a_2, ..., a_n\}$$

such that $u(a_1) \geq u(a_2) \geq \cdots \geq u(a_n)$. The matrix (2) will then be non-decreasing along each row (from left to right) and non-increasing down each column (from top to bottom). This property will prove convenient, so we assume that A has been enumerated in this fashion whenever (A, P) is MS and A is finite.

4 Transitivity, substitution and independence

Tversky and Russo (1969) prove that simple scalability may be characterised by any one of three equivalent properties:[2]

Definition 7 *A CBCS satisfies* ***substitutability*** *if*

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

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

Definition 8 *A CBCS satisfies* ***strict stochastic transitivity (StST)*** *if*

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

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

Definition 9 *A CBCS satisfies* ***independence*** *if the following holds for all* $a, b, c, d \in A$*:*

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

It is obvious that substitutability and independence are identical restrictions on a BCP.[3] The equivalence of these two with StST is less obvious, and their mutual equivalence with simple scalability even less so.

[2] The "strict stochastic transitivity" terminology is from Fishburn (1973). In Tversky and Russo (1969), "strong stochastic transitivity" is used to refer to what we have labelled StST. Most authors use "strong stochastic transitivity" for the property that we introduce below under the label SST. We adopt the more conventional usage.

[3] Set $b = c$ in Definition 9 to observe that independence implies substitutability. For the converse, if substitutability holds and $P(a, c) \geq P(b, c)$, then

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

and hence $P(a, d) \geq P(b, d)$.

Tversky and Russo (1969) assume that $\Delta(A, P) = A \times A$. A careful reading of Tversky and Russo's argument shows this restriction plays no role, but we prove the following to keep the present analysis self-contained.

Theorem 10 *Given a CBCS (A, P), the following are equivalent:*

(i) (A, P) *satisfies SS;*

(ii) (A, P) *satisfies substitutability;*

(iii) (A, P) *satisfies StST.*

(iv) (A, P) *satisfies independence.*

Proof: We first show that (i) implies (iii). Using the properties of F implied by SS and (1), we have

$$\begin{aligned} & \min\{F(u(a), u(b)),\ F(u(b), u(c))\} \ \geq [>]\ \frac{1}{2} \\ \Rightarrow\quad & u(a) \geq [>] u(b) \geq [>] u(c) \\ \Rightarrow\quad & F(u(a), u(c)) \geq [>] \max\{F(u(a), u(b)),\ F(u(b), u(c))\} \end{aligned}$$

from which StST follows.

Next, we show that (iii) implies (ii). By Lemma 16, (A, P) satisfies weak substitutability, so we need only show the converse to the weak substitutability implication. Suppose $P(a, c) \geq P(b, c)$ and $P(a, b) < \frac{1}{2}$. We show that a contradiction necessarily follows. Suppose $P(a, c) > \frac{1}{2}$. Since $P(b, a) > \frac{1}{2}$, StST gives $P(b, c) > P(a, c)$, which is a contradiction. Suppose, instead, that $P(a, c) \leq \frac{1}{2}$. Then

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

which implies

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

From $P(c, b) \geq \frac{1}{2}$, $P(b, a) > \frac{1}{2}$ and StST we deduce

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

Thus,

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

Applying StST we get $P(c, a) > P(c, b)$ which contradicts $P(a, c) \geq P(b, c)$.

We now establish the equivalence of (i), (ii) and (iii) by showing that (ii) implies (i). Since we already know that (ii) and (iv) are equivalent, this will complete the proof. Fix some $\overline{a} \in A$ and define $u : A \rightarrow [0,1]$ by $u(a) = P(a, \overline{a})$. Now define $F : u(A) \times u(A) \rightarrow [0,1]$ by $F(u(a), u(b)) = P(a, b)$.[4] Substitutability ensures that F is well-defined:

$$\begin{aligned}
(u(a), u(b)) = (u(a'), u(b')) \quad &\Leftrightarrow \quad P(a, a') = P(b, b') = \frac{1}{2} \\
&\Leftrightarrow \quad P(a, b) = P(a', b) \text{ and } P(b, a') = P(b', a') \\
&\Rightarrow \quad P(a, b) = P(a', b').
\end{aligned}$$

Substitutability also ensures that F has the required monotonicity properties. To see that F is strictly increasing in its first argument, note that for any $a, a', b \in A$ we have:

$$\begin{aligned}
u(a) > u(a') \quad &\Leftrightarrow \quad P(a, a') > \frac{1}{2} \\
&\Leftrightarrow \quad P(a, b) > P(a', b) \\
&\Leftrightarrow \quad F(u(a), u(b)) > F(u(a'), u(b)).
\end{aligned}$$

Similar logic shows that F is strictly decreasing in its second argument. □

Fishburn (1973) obtains a characterisation of monotone scalability by weakening independence as follows:

Definition 11 *A CBCS satisfies* ***weak independence*** *if the following holds for all $a, b, c, d \in A$:*

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

Theorem 12 (Fishburn, 1973) *A CBCS satisfies MS iff it satisfies weak independence and there exists a utility representation for $\succsim_0^P$.*

Fishburn (1973, Theorem 1) shows that weak independence is equivalent to $\succsim_0^P$ being complete (i.e., a weak order), so the additional requirement that there exist a utility representation for $\succsim_0^P$ is equivalent to the requirement that A contain a countable $\succsim_0^P$-dense subset (*ibid.*, pp.350-351). When A itself is countable, this latter condition is trivially satisfied.

Corollary 13 *If (A, P) is a CBCS and A is countable, then the following are equivalent:*

[4] Note that Debreu (1958) uses essentially the same construction.

(i) (A, P) *satisfies MS;*

(ii) (A, P) *satisfies weak independence.*

There is no known analogue of weak independence in the form of a substitutability or stochastic transitivity condition.

The results of Tversky and Russo (1969) and Fishburn (1973) raise the obvious question of whether there is a condition intermediate in strength between independence and weak independence that characterises strict scalability? In the next section, we obtain such a condition, which turns out to be surprisingly familiar.

5 Strict scalability

The following two concepts have a long and distinguished heritage in the literature on BCPs.

Definition 14 *A CBCS satisfies* ***weak substitutability*** *if*

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

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

Definition 15 *A CBCS satisfies* ***strong stochastic transitivity (SST)*** *if*

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

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

It is well known that these two concepts are in fact equivalent.[5]

Lemma 16 (Davidson and Marschak, 1959) *A CBCS satisfies SST iff it satisfies weak substitutability.*

Proof: Let (A, P) be a CBCS that satisfies SST and let $P(a,b) \geq \frac{1}{2}$. If $P(b,c) \geq \frac{1}{2}$ then $P(a,c) \geq P(b,c)$ by SST. Suppose $P(b,c) < \frac{1}{2}$, and hence $P(c,b) > \frac{1}{2}$. If $P(a,c) \geq \frac{1}{2}$ then $P(a,c) \geq P(b,c)$ by SST. If $P(a,c) < \frac{1}{2}$, then $P(c,a) > \frac{1}{2}$ so $P(c,b) \geq P(c,a)$ by SST and hence $P(a,c) \geq P(b,c)$.

[5] Davidson and Marschak, like Tversky and Russo, restrict attention to CBCS's with $\Delta(A,P) = A \times A$, so we include a version of their proof to verify that this restriction plays no role.

Conversely, suppose (A, P) satisfies weak substitutability and

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

Then $P(a,c) \geq P(b,c)$ and $P(b,a) \geq P(c,a)$. The latter implies $P(a,c) \geq P(a,b)$. □

The next example shows that weak substitutability is *strictly* weaker than substitutability.

Example 17 *We adapt an example from Luce and Suppes (1965, p.346). Let* $A = \{a_1, a_2, a_3\}$ *and specify* P *as follows:*

$$\frac{1}{6}\begin{bmatrix} 3 & 4 & 4 \\ 2 & 3 & 4 \\ 2 & 2 & 3 \end{bmatrix}.$$

Note that $a_1 \succ^P a_2 \succ^P a_3$ *in this example. Thus, weak substitutability requires that*

$$P(a_i, a_k) \geq P(a_j, a_k) \quad \textit{for all } k$$

whenever $i \leq j$*. That is, if* $\boldsymbol{P}$ *is the matrix representation for* P*, and* $\boldsymbol{P}_{k,\cdot}$ *denotes the* k^{th} *row of* $\boldsymbol{P}$*, then weak substitutability requires* $\boldsymbol{P}_{i,\cdot} \geq \boldsymbol{P}_{j,\cdot}$ *whenever* $i \leq j$*. It is clear that this is the case. However, substitutability is violated:* $P(a_2, a_1) = P(a_3, a_1)$ *but* $P(a_2, a_3) > \frac{1}{2}$*.*

It is also obvious that weak substitutability implies weak independence.[6] Example 21 below shows the converse to be false.

From Theorem 10 and Lemma 16 we therefore have the following (well known) relationships, with neither of the omitted converses being valid:

$$\begin{array}{lcl} \text{weak independence} & \Rightarrow & \text{weak substitutability} \ \Leftrightarrow\ \text{SST} \\ & \Rightarrow & \text{substitutability} \ \Leftrightarrow\ \text{StST} \ \Leftrightarrow\ \text{independence.} \end{array}$$

It will be convenient to introduce one further substitution property.

Definition 18 *A CBCS satisfies* ***quasi-substitutability*** *if the following holds for all* $a, b \in A$*:*

$$P(a,b) \geq \frac{1}{2} \quad \Leftrightarrow \quad [P(a,c) \geq P(b,c) \ \textit{ for all } c \in A]$$

[6] The former rules out the possibility that there exist $a, b, c, d \in A$ with

$$P(a,c) > P(b,c) \quad \text{and} \quad P(a,d) < P(b,d).$$

Thus, (A, P) satisfies quasi-substitutability iff $\succsim^P = \succsim_0^P$.

Quasi-substitutability appears to be strictly intermediate between weak substitutability and substitutability, but is actually equivalent to the former.

Lemma 19 *A CBCS satisfies quasi-substitutability iff it satisfies weak substitutability.*

Proof: The "only if" part is obvious. For the "if" part, suppose (A, P) satisfies weak substitutability and $P(b, a) > \frac{1}{2}$. We must show that $P(a, c) < P(b, c)$ for at least one $c \in A$. Weak substitutability ensures that $P(b, c) \geq P(a, c)$ for all $c \in A$. If equality holds for all $c \in A$ then

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

which contradicts $P(a, b) < \frac{1}{2}$. □

Corollary 20 *A CBCS satisfies* $\succsim^P = \succsim_0^P$ *iff it satisfies weak substitutability.*

Since completeness of $\succsim_0^P$ is equivalent to weak independence, and transitivity of $\succsim^P$ is equivalent to WST, it follows that weak substitutability implies both weak independence (since $\succsim_0^P$ must be complete if $\succsim^P = \succsim_0^P$) and WST (since $\succsim^P$ must be transitive if $\succsim^P = \succsim_0^P$). Example 21 establishes that the converse is false. That is, even if $\succsim_0^P$ and $\succsim^P$ are both weak orders, it need not be the case that $\succsim^P = \succsim_0^P$.

Example 21 *Let* $A = \{a_1, a_2, a_3\}$ *and specify* P *as via the following matrix (denoted* $\boldsymbol{P}$*):*

$$\frac{1}{12}\begin{bmatrix} 6 & 9 & 8 \\ 3 & 6 & 6 \\ 4 & 6 & 6 \end{bmatrix}.$$

We have $a_1 \succ^P a_2$, $a_2 \sim^P a_3$ *and* $a_1 \succ^P a_3$ *so* P *satisfies WST. It is also clear that* $\boldsymbol{P}_{i,\cdot} \geq \boldsymbol{P}_{j,\cdot}$ *or* $\boldsymbol{P}_{j,\cdot} \geq \boldsymbol{P}_{i,\cdot}$ *for any* i, j*, so* P *satisfies weak independence. However, weak substitutability is violated:* $P(a_2, a_1) < P(a_3, a_1)$ *but* $P(a_2, a_3) = \frac{1}{2}$.

Combining the foregoing observations with Lemma 16 we have:

Corollary 22 *For any CBCS, SST implies both WST and weak independence, but not conversely (i.e., the conjunction of WST and weak independence does* ***not*** *imply SST).*

Fishburn (1973) proves that SST implies weak independence but not conversely – see his Figure 1 – but the stronger result established here has not, to the best of our knowledge, been observed previously.

We can now prove our main result.

Theorem 23 *A CBCS satisfies StS iff it satisfies weak substitutability and there exists a utility representation for* $\succsim^P$.

Proof: Let (A, P) be strictly scalable through (u, F). In particular, u represents $\succsim^P$. Then

$$\begin{aligned} P(a,b) \geq \frac{1}{2} \quad &\Rightarrow \quad u(a) \geq u(b) \\ &\Rightarrow \quad F(u(a), u(c)) \geq F(u(b), u(c)) \quad \text{for all } c \in A \\ &\Rightarrow \quad P(a,c) \geq P(b,c) \quad \text{for all } c \in A \end{aligned}$$

where the second implication uses the fact that F is non-decreasing in its first argument. Hence, (A, P) satisfies weak substitutability.

Next, suppose (A, P) satisfies quasi-substitutability (which is equivalent to weak substitutability – Lemma 19) and let $u : A \to \mathbb{R}$ be a representation for $\succsim^P$. Then:

$$\begin{aligned} u(c') \geq u(c) \quad &\Leftrightarrow \quad P(c', c) \geq \frac{1}{2} \\ &\Leftrightarrow \quad P(c', d) \geq P(c, d) \text{ for any } d \in A \\ &\Leftrightarrow \quad P(d, c) \geq P(d, c') \text{ for any } d \in A. \end{aligned}$$

and hence

$$\begin{aligned} u(c') = u(c) \quad &\Leftrightarrow \quad P(c', d) = P(c, d) \text{ for any } d \in A \\ &\Leftrightarrow \quad P(d, c) = P(d, c') \text{ for any } d \in A. \end{aligned}$$

It follows that the condition

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

determines a well-defined function $F : u(A) \times u(A) \to [0, 1]$ with the required monotonicity properties. □

Given weak substitutability, which implies WST, the existence of a utility representation is equivalent to A containing a countable $\succsim^P$-dense subset, which is trivially satisfied for countable A.

Corollary 24 *If (A, P) is a CBCS and A is countable, then the following are equivalent:*

(i) (A, P) *satisfies StS;*

(ii) (A, P) *satisfies weak substitutability;*

(iii) (A, P) *satisfies SST.*

Interestingly, if A is not countable, then (A, P) may satisfy weak substitutability but fail to satisfy even *monotone* scalability. Fishburn (1973, p.351) constructs a CBCS (A, P) that is not monotone scalable. In particular, the binary relation $\succsim_0^P$ does not possess a utility representation. It is straightforward to observe that $\succsim^P = \succsim_0^P$ for Fishburn's example, from which we deduce that (A, P) satisfies weak substitutability (Corollary 20).

6 Discussion

Figure 1 summarises the relationships described in this paper. It uses the same conventions as Fishburn (1973, Figure 1). Conditions within a single box are equivalent. If an arrow points in only one direction it means that the converse implication is false. If there is no path from one box to another via a series of arrows then the two conditions are logically independent. The implied relationships within and between dashed boxes were already established by previous authors (with the exception of the equivalence between weak substitutability and $\succsim^P = \succsim_0^P$).

Our main contribution can be viewed from two perspectives. From one perspective, our contribution is to establish axiomatic foundations for strict scalability. The notion of strict scalability occupies a natural intermediate position between monotone and simple scalability. In Ryan (2015) we introduced a similar variant on standard Fechnerian representations. Alternatively, when A is countable, our contribution may be viewed as establishing a convenient representation for any BCP that satisfies the familiar weak substitutability condition.

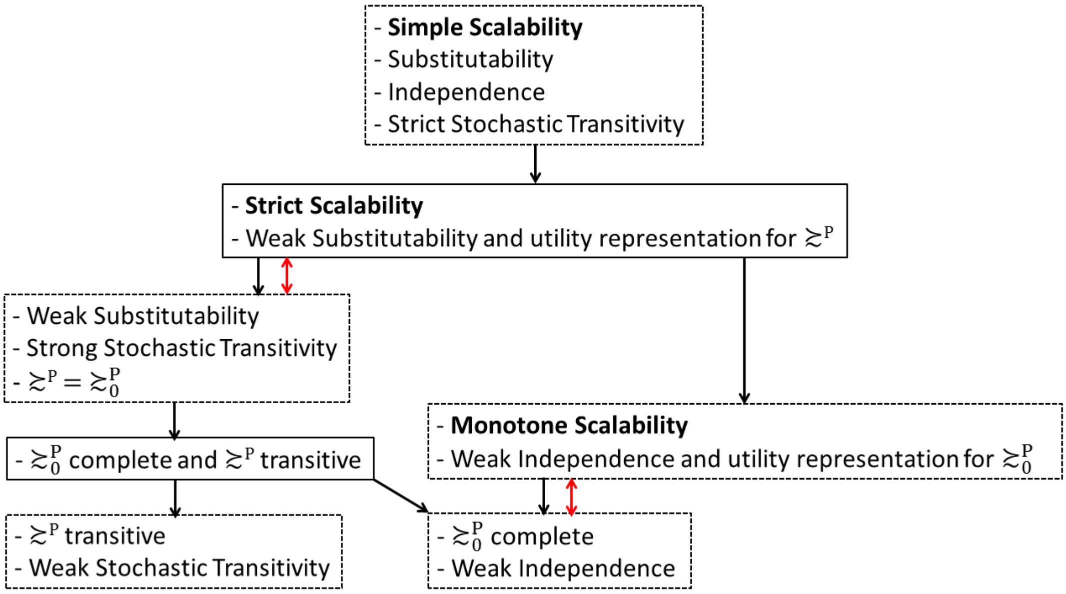

FIGURE 1: Red arrows apply only if A is countable

References

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

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

FALMAGNE, J.C. (1985), *Elements of Psychophysical Theory.* Oxford University Press.

FISHBURN, P.J. (1973), "Binary Choice Probabilities: On the Varieties of Stochastic Transitivity," *Journal of Mathematical Psychology* **10**, 327–352.

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.

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

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