Soc Choice Welf (2015) 45:707–722
DOI 10.1007/s00355-015-0876-5

Expected utility without full transitivity

Walter Bossert · Kotaro Suzumura

Received: 9 September 2014 / Accepted: 22 January 2015 / Published online: 31 January 2015

Abstract We generalize the classical expected-utility criterion by weakening transitivity to Suzumura consistency. In the absence of full transitivity, reflexivity and completeness no longer follow as a consequence of the system of axioms employed and a richer class of rankings of probability distributions results. This class is characterized by means of standard expected-utility axioms in addition to Suzumura consistency. An important feature of some members of our new class is that they allow us to soften the negative impact of so-called paradoxes that involve preference reversals without abandoning the expected-utility framework altogether.

JEL Classification D81

1 Introduction

The formal treatment of the expected-utility criterion has a long-standing tradition in the theory of individual choice under uncertainty, going back as far as von Neumann and Morgenstern's (1944, 1947) seminal contribution. While numerous criticisms have been leveled at the descriptive suitability of expected-utility theory (often in the context of experimental studies), the criterion has proven to be rather robust in that it remains on a sound normative foundation. Nevertheless, perceived inconsistencies

W. Bossert (✉)
Department of Economics and CIREQ, University of Montreal, Station Downtown,
P.O. Box 6128, Montreal, QC H3C 3J7, Canada
e-mail: walter.bossert@videotron.ca

K. Suzumura
School of Political Science and Economics, Waseda University, 1-6-1 Nishi-Waseda,
Shinjuku-ku, Tokyo 169-8050, Japan
e-mail: ktr.suzumura@gmail.com

that involve various forms of preference reversals [Tversky and Thaler (1990)] such as those illustrated by Allais (1953), Kahneman and Tversky (1979) and Machina (1983) constitute serious challenges that need to be responded to if (at least some form of) expected-utility theory is to continue to be an attractive option in descriptive contexts as well.

In an attempt to address paradoxes of this nature, several alternative theories have been developed over the years. These include Kahneman and Tversky's (1979) prospect theory and regret theory as investigated by Bell (1982) and Loomes and Sugden (1982). Moreover, there is a rapidly growing literature on behavioral approaches to economic decision-making; see Simon (1955), Camerer (1995) and Rabin (1998), for instance.

The above-mentioned alternative approaches represent clear-cut departures from the expected-utility criterion. In contrast, the objective of this paper is to attempt to soften the negative impact of a large class of preference-reversal examples without abandoning the expected-utility framework altogether. This is achieved by retaining most of the traditional expected-utility axioms but weakening transitivity to Suzumura consistency. As is well-known, transitivity (along with other standard expected-utility axioms) implies that the resulting decision rule is reflexive and complete—any two probability distributions can be ranked. In the absence of full transitivity, this implication is no longer valid and, thus, new decision rules emerge as additional possibilities.

Many of the behavioral approaches alluded to above explicitly start out with the hypothesis that economic agents are not necessarily fully rational but that they make choices under what is often referred to as bounded rationality. An advantage of the theories developed in this context is their ability to explain specific observable patterns of behavior in a coherent manner. On the other hand, many of these models are restricted to rather specific situations and, thus, are difficult to justify as general methods to describe observed choices.

The notion of bounded rationality frequently appears in cases where the decision problem under consideration is deemed to be too complex to allow for full rationality in the traditional sense. This reasoning appears to apply analogously to situations in which the inherent complexity leads us to the assumption that economic agents may not be able to rank all possible probability distributions under consideration. Thus, this complexity argument can be invoked in support of our approach which allows for non-comparability as well. That completeness may be a rather strong assumption in the context of expected-utility theory has been argued in many earlier contributions—von Neumann and Morgenstern (1944, 1947) themselves make this point; other authors who question the completeness axiom include Thrall (1954), Luce and Raiffa (1957), Aumann (1962, 1964, 2000), Fishburn (1971, 1972) and Dubra et al. (2004). After all, there are several instances where the imposition of completeness might create artificial puzzles and even impossibilities; the earlier contributions just cited are merely examples of such problems that may be triggered by the completeness assumption. In particular, a detailed discussion of the plausibility of incompleteness can be found in Aumann (1962) and, clearly, his arguments continue to be as compelling as they were at the time. By the same token, Fishburn (1971, 1972) "one-way expected utility" approach, to be discussed in more detail shortly, is of lasting importance in the field.

The class of decision rules that we characterize generalizes the expected-utility criterion in that some pairs of probability distributions may be considered to be non-comparable. Thus, this class is considerably richer than the traditional expected-utility criterion because it ranges from the classical (fully transitive) case itself to a minimal criterion that only imposes rankings for a small subset of the pairs of probability distributions. An important consequence of making these more general (possibly incomplete) rules available is that paradoxes that involve preference reversals can be avoided by a suitable choice of a member of our class—namely, a generalized expected-utility criterion that may be silent regarding the relative desirability of two probability distributions that appear in the paradox under consideration. Although our results primarily contribute to theoretical aspects of choice under uncertainty, the class of rules that we characterize may also be of use in addressing issues that involve preference reversals.

There are two conceptually distinct aspects to replacing transitivity with the weaker property of Suzumura consistency. First, we do not impose full transitivity because this, together with our other axioms, would prevent us from considering possibly incomplete decision rules. Thus, this part of our approach is an absolute necessity, keeping in mind that our objective is to obtain a richer class of rules in the first place. Second, if we were to eliminate transitivity from our list of axioms without replacing it by some other coherence property, this would result in a set of rules that are much too permissive. Without any further restrictions, cyclical rules would be members of the corresponding class which clearly is not desirable. Among the possible weakenings of transitivity, Suzumura consistency appears to be the most natural and attractive option, as we argue below.

All but one of the axioms that we employ are standard in the literature on choice under uncertainty and in many other branches of the literature. In particular, we employ notions of solvability, monotonicity and independence. Loosely speaking, solvability is related to continuity properties, monotonicity rules out counter-intuitive rankings in relatively straightforward comparisons and independence is a separability condition.

The only axiom that is less familiar in the context of choice under uncertainty is Suzumura consistency and, for that reason, we discuss it in some detail. This coherence property of a binary relation was introduced by Suzumura (1976). It rules out all preference cycles that involve at least one instance of strict preference. Thus, Suzumura consistency is stronger than acyclicity which merely prohibits cycles such that all preferences involved are strict. Furthermore, Suzumura consistency is implied by transitivity. If a relation is reflexive and complete, Suzumura consistency also implies transitivity but this implication does not hold if reflexivity or completeness is violated. Sen's (1969) axiom of quasi-transitivity, which demands that the asymmetric part of a relation be transitive, and Suzumura consistency are independent. Because Suzumura consistency is equivalent to transitivity in the presence of reflexivity and completeness, it can be considered a very natural weakening; note that quasi-transitivity fails to imply transitivity even if a relation is reflexive and complete and, of course, the same observation applies for acyclicity (which is weaker than quasi-transitivity).

Further forceful arguments in support of Suzumura consistency can be made. A well-known theorem due to Szpilrajn (1930) establishes that transitivity is sufficient for the existence of an ordering extension of a relation. This is a fundamental result that has been applied in numerous settings. A remarkable strengthening of Szpilrajn's

(1930) extension theorem is proven by Suzumura (1976) who shows that Suzumura consistency is necessary and sufficient for the existence of an ordering extension, thus providing a very clear demarcation line between the set of relations that can be extended to an ordering and those that cannot. This is another attractive feature of Suzumura consistency as compared to quasi-transitivity and acyclicity: neither of these properties can be used to obtain such an equivalence result. In addition, Bossert et al. (2005) show that there exists a well-defined Suzumura-consistent closure of any relation, just as is the case for transitivity. No such closure operations exist in the cases of quasi-transitivity and acyclicity. These observations reinforce our statement that Suzumura consistency is indeed a natural weakening of transitivity, and a detailed analysis of the axiom is carried out in Bossert and Suzumura (2010) where we demonstrate the usefulness of this coherence property in numerous individual and collective decision problems including, among others, topics in revealed preference theory. See also Suzumura (1978) for the crucial service rendered by Suzumura consistency in the social-choice theoretic analysis of individual rights. Possible violations of transitivity and some of their consequences are examined by authors such as Fishburn (1982, 1991) and Fishburn and LaValle (1988) but none of these studies employ the notion of Suzumura consistency, which is an essential novel feature of our approach.

There are several studies that analyze expected-utility theory in the context of reflexive and transitive but not necessarily complete relations over probability distributions, such as those carried out by Aumann (1962, 1964, 2000), Fishburn (1971, Theorem 1; 1972, Theorem A) and Dubra et al. (2004). However, these contributions retain full transitivity as an assumption and, as a consequence, obtain results that are quite different in nature from ours.

Aumann (1962) considers rankings of probability distributions that are reflexive and transitive but not necessarily complete, in conjunction with a continuity property and a variant of the independence axiom. He shows that, under his assumptions, there exists an additive function such that if two distributions are indifferent, then they must generate the same expectation according to this function and, likewise, if a distribution is strictly preferred to another, the former is associated with a greater expectation than the latter. Clearly, the ranking generated by such a function is not necessarily complete and, as pointed out by Dubra et al. (2004, footnote 2), "…[this] approach falls short of yielding a representation theorem, for it does not *characterize* the preference relations under consideration." An analogous remark applies to Theorems A, B and C of Fishburn (1972)—and to Theorem 1 of Fishburn (1971), which is identical to his Theorem A. In these results, sufficient (but not necessary) conditions for one of the implications in our theorems are provided. This implication demands that if a probability distribution p is strictly preferred to a probability distribution q, then the expected utility of p must exceed that of q. Because we work within a richer setting, our results are independent of Fishburn's. Fishburn (1971) also establishes necessary and sufficient conditions for the above implication but, again, these results do not imply—and are not implied by—our observations. We reiterate at this point that none of Fishburn (1971, 1972) results—or any others in the context of choice under uncertainty that we are aware of—employs the axiom of Suzumura consistency.

Dubra et al. (2004) establish an expected multi-utility theorem that characterizes reflexive and transitive but possibly incomplete preferences on probability distributions by means of a continuity property and a version of the independence axiom. The idea underlying the multi-utility approach is that a reflexive and transitive dominance criterion can be established by means of a set of possible utility functions such that a distribution is considered at least as good as another if and only if the expectation according to the former is greater than or equal to that of the latter for all utility functions in this set.

Aumann (1962) and Dubra et al. (2004) do not impose a property akin to the monotonicity condition alluded to earlier, which is a major reason why completeness does not follow from their axioms even in the presence of full transitivity. This, in addition to the absence of Suzumura consistency in their work, is yet another feature that distinguishes these approaches from ours.

In Sect. 2, we introduce our notation and basic definitions. Section 3 contains a preliminary characterization of decision rules on the basis of our axioms without independence, followed by our main result. Section 4 illustrates that, in addition to the theoretical contribution of our work, some members of our new class of rules may serve to address perceived paradoxes that involve preference reversals. Section 5 concludes.

2 Definitions

Suppose there is a fixed finite set of alternatives $X = \{x_1, \ldots, x_n\}$, where $n \in \mathbb{N} \setminus \{1, 2\}$. We exclude the one-alternative and two-alternative cases because they are trivial: clearly, the case $n = 1$ is degenerate and in the case $n = 2$ we are immediately back to the classical expected-utility criterion once our monotonicity property (see below) is imposed. The set $\Delta = \{p \in \mathbb{R}^n_+ \mid \sum_{i=1}^n p_i = 1\}$ is the unit simplex in $\mathbb{R}^n_+$, interpreted as the set of all probability distributions on X. For all $i \in \{1, \ldots, n\}$, the ith unit vector in $\mathbb{R}^n$ is denoted by e^i.

A (binary) relation on Δ is a set $\succsim \subseteq \Delta^2$. As usual, the symmetric and asymmetric parts $\sim$ and $\succ$ of $\succsim$ are defined by letting, for all $p, q \in \Delta$,

$$p \sim q \;\Leftrightarrow\; p \succsim q \text{ and } q \succsim p$$

and

$$p \succ q \;\Leftrightarrow\; p \succsim q \text{ and } \neg(q \succsim p).$$

We interpret the relation $\succsim$ as the preference relation (or the decision rule) used by an agent to rank probability distributions. The relations $\sim$ and $\succ$ are the corresponding indifference relation and strict preference relation.

The transitive closure $\overline{\succsim}$ of $\succsim$ is defined by letting, for all $p, q \in \Delta$,

$$p \,\overline{\succsim}\, q \Leftrightarrow \text{there exist } K \in \mathbb{N} \text{ and } r^0, \ldots, r^K \in \Delta \text{ such that}$$
$$p = r^0 \text{ and } r^{k-1} \succsim r^k \text{ for all } k \in \{1, \ldots, K\} \text{ and } r^K = q.$$

It is assumed that there exist two distinct alternatives x_j and x_k in X such that the probability distribution that assigns a probability of one to x_j is strictly preferred to the distribution that yields x_k with certainty. Without loss of generality, we assume that these alternatives are given by $x_j = x_1$ and $x_k = x_n$. (There is no claim that x_1 is the best and x_n is the worst among the alternatives in $X = \{x_1, \ldots, x_n\}$; see also the remarks following the statement of our Theorem 2 to the effect that the elements of X need not be fully ordered). To reflect this assumption, we explicitly incorporate the statement $(1, 0, \ldots, 0) \succ (0, \ldots, 0, 1)$ into our formal results. We reiterate that the choice of x_1 and x_n does not involve any loss or generality and allows us to simplify our exposition considerably.

Clearly, the above-described assumption rules out the universal-indifference relation. This feature of our approach is intended. Recall that our goal is to identify a rich class of decision rules that allow us to declare pairs of distributions non-comparable even though they are ranked by the classical expected-utility criterion. Note that universal indifference is explicitly ruled out by our monotonicity axiom.

We now introduce some properties of the decision rule $\succsim$. The first three of these are simply the properties that define an ordering. Note that reflexivity and completeness are usually not imposed when formulating some of the versions of the classical expected-utility theorem, such as that of Kreps (1988); they are implied by the set of axioms employed in the requisite result when transitivity is one of these axioms. We do not require reflexivity and completeness in our generalized expected-utility theorem either. Because the full force of transitivity is not imposed, a more general class of (not necessarily reflexive and complete) decision rules can be obtained. As mentioned in the introduction, relaxing transitivity is needed in order to open up the possibility of incompleteness.

Reflexivity For all $p \in \Delta,\ p \succsim p$.

Completeness For all $p, q \in \Delta$ such that $p \neq q$,
$p \succsim q$ or $q \succsim p$.

Transitivity For all $p, q, r \in \Delta,\ [p \succsim q \text{ and } q \succsim r] \Rightarrow p \succsim r$.

The first axiom used in our results is Suzumura consistency.

Suzumura consistency For all $p, q \in \Delta,\ p \overline{\succsim} q \Rightarrow \neg(q \succ p)$.

Recall that, in the presence of reflexivity and completeness, Suzumura consistency and transitivity are equivalent but, because our results do not involve reflexivity and completeness, transitivity is not implied. Again, see Suzumura (1976) and Bossert and Suzumura (2010) for more detailed discussions.

The remaining three properties are standard in decision theory as well as, suitably reformulated, in numerous other areas within microeconomic theory. The first of these amounts to a continuity condition, the second ensures that the direction of preference is in accord with the interpretation of the relation $\succsim$ as a decision rule for choice under uncertainty, and the third is a separability property. As discussed earlier, the seemingly special role played by the alternatives labeled x_1 and x_n in the axioms does not involve any loss of generality.

Solvability For all $p \in \Delta$, there exists $\alpha \in [0, 1]$ such that $p \sim (\alpha, 0, \ldots, 0, 1 - \alpha)$.

Monotonicity For all $\alpha, \beta \in [0, 1]$,

$$(\alpha, 0, \ldots, 0, 1 - \alpha) \succsim (\beta, 0, \ldots, 0, 1 - \beta) \Leftrightarrow \alpha \geq \beta.$$

Independence For all $p, q \in \Delta$ and for all $\alpha, \beta, \gamma \in [0, 1]$, if $p \sim (\alpha, 0, \ldots, 0, 1-\alpha)$ and $q \sim (\beta, 0, \ldots, 0, 1-\beta)$, then

$$\gamma p + (1-\gamma)q \sim \gamma(\alpha, 0, \ldots, 0, 1-\alpha) + (1-\gamma)(\beta, 0, \ldots, 0, 1-\beta).$$

Our formulation of the independence axiom differs from some of the traditional variants in some respects. One of the standard versions requires that an indifference between two distributions p and q implies that any convex combination of p and any distribution r with weights γ and $1-\gamma$ is indifferent to the convex combination of q and r with the same weights γ and $1-\gamma$. In the presence of transitivity (or merely transitivity of the indifference relation $\sim$), our axiom is implied by this alternative property because it restricts the requisite implication to a subset of pairs of distributions. The reason why we employ this alternative formulation is that we intend to arrive at a characterization result without having to impose transitivity of $\sim$. For instance, Luce's (1956) well-known coffee-sugar example provides a powerful argument against the use of transitive indifference: a decision maker may find it very difficult to perceive small differences and, thus, the indifference relation may very well fail to be transitive. If one is willing to require $\sim$ to be transitive, a more restrictive class of decision rules is characterized in the presence of our remaining axioms; in this case, all indifferences according to the expected-utility criterion have to be respected. Details on this alternative result are available from the authors on request.

Unlike Kreps (1988) and other authors, we treat $\succsim$ as the primitive concept rather than $\succ$. As is standard when $\succ$ is considered to be the primary relation, Kreps (1988) imposes asymmetry and negative transitivity on $\succ$. If the relation $\succsim$ is required to be an ordering (that is, reflexive, complete and transitive), it does not matter whether we start out with $\succsim$ or with the associated asymmetric part $\succ$ as the primitive notion because the conjunction of asymmetry and negative transitivity of $\succ$ implies that the corresponding relation $\succsim$ is an ordering; see Kreps (1988, pp. 9–10). Thus, adopting Kreps's (1988) setting would result in an immediate conflict with our main objective of examining the consequences of weakening the transitivity requirement in the context of expected-utility theory.

3 Suzumura-consistent expected utility

The main result of this paper establishes that if transitivity is weakened to Suzumura consistency, a class of generalized expected-utility criteria is characterized. These relations allow for violations of reflexivity or completeness in some situations. The preferences imposed by solvability and monotonicity continue to be required but, because full transitivity is no longer used, any additional pairs that belong to the expected-utility relation may or may not be included. Thus, the new class contains as special cases the standard (reflexive and complete) expected-utility criterion as the "maximal" relation and the one where the only preferences are those imposed by solvability and monotonicity as the "minimal" relation satisfying the axioms. In particular, this means that any other pair that is weakly (strictly) ranked by the expected-utility criterion may either be weakly (strictly) ranked or non-comparable according to our generalization. Thus, undesirable observations such as preference reversals can be ameliorated by replacing one of the two counter-intuitive preferences with non-comparability. We illustrate this feature in Sect. 4.

As a first step, we characterize the class of all decision rules that satisfy Suzumura consistency, solvability and monotonicity. This theorem is of some interest in its own right: even in the absence of independence (the quintessential condition that underlies the expected-utility criterion), weakening transitivity to Suzumura consistency yields a precisely defined class of decision rules. We also use parts of its proof in establishing our main result.

Theorem 1 *Suppose that X contains at least three alternatives and that $\succsim$ is a relation on Δ such that $(1, 0, \ldots, 0) \succ (0, \ldots, 0, 1)$. The relation $\succsim$ satisfies Suzumura consistency, solvability and monotonicity if and only if there exists a function $\varphi\colon \Delta \to [0, 1]$ such that the pair $(\succsim, \varphi)$ satisfies*

(i) $\big[(\alpha, 0, \ldots, 0, 1 - \alpha) \succsim (\beta, 0, \ldots, 0, 1 - \beta) \Leftrightarrow \alpha \geq \beta\big]$ *for all* $\alpha, \beta \in [0, 1]$;
(ii) $p \sim (\varphi(p), 0, \ldots, 0, 1 - \varphi(p))$ *for all* $p \in \Delta$;
(iii) $p \sim q \;\Rightarrow\; \varphi(p) = \varphi(q)$ *for all* $p, q \in \Delta$;
(iv) $p \succ q \;\Rightarrow\; \varphi(p) > \varphi(q)$ *for all* $p, q \in \Delta$.

Proof 'If.' Solvability and monotonicity follow immediately from (i) and (ii). Suppose, by way of contradiction, that Suzumura consistency is violated. Then there exist $p, q \in \Delta$ such that $p \overline{\succsim} q$ and $q \succ p$. Thus, by definition of the transitive closure, there exist $K \in \mathbb{N}$ and $r^0, \ldots, r^K \in \Delta$ such that $p = r^0$ and $r^{k-1} \succsim r^k$ for all $k \in \{1, \ldots, K\}$ and $r^K = q$. Consider any $k \in \{1, \ldots, K\}$. If $r^{k-1} \sim r^k$, it follows that

$$\varphi(r^{k-1}) = \varphi(r^k)$$

because of (iii). If $r^{k-1} \succ r^k$, (iv) implies

$$\varphi(r^{k-1}) > \varphi(r^k).$$

Thus, for all $k \in \{1, \ldots, K\}$, we have

$$\varphi(r^{k-1}) \geq \varphi(r^k).$$

Combining these inequalities for all $k \in \{1, \ldots, K\}$ and using $p = r^0$ and $r^K = q$, it follows that

$$\varphi(p) \geq \varphi(q). \tag{1}$$

By (iv), $q \succ p$ implies

$$\varphi(q) > \varphi(p),$$

contradicting (1). Thus, Suzumura consistency is satisfied.

'Only if.' Part (i) follows immediately from monotonicity.

To prove (ii), let $p \in \Delta$ and $\alpha \in [0, 1]$ be such that

$$p \sim (\alpha, 0, \ldots, 0, 1 - \alpha). \tag{2}$$

The existence of α is guaranteed by solvability. Furthermore, α is unique. To see this, suppose, by way of contradiction, that there exists $\beta \in [0, 1] \setminus \{\alpha\}$ such that

$$p \sim (\beta, 0, \ldots, 0, 1 - \beta). \tag{3}$$

If $\beta > \alpha$, we have

$$(\beta, 0, \ldots, 0, 1 - \beta) \succ (\alpha, 0, \ldots, 0, 1 - \alpha) \tag{4}$$

by monotonicity and, by (2) and (3),

$$(\alpha, 0, \ldots, 0, 1 - \alpha) \succsim (\beta, 0, \ldots, 0, 1 - \beta)$$

which, together, with (4), leads to a contradiction to Suzumura consistency. The same argument applies if $\beta < \alpha$. Thus, α must be unique for p and we can write it as a function $\varphi \colon \Delta \to [0, 1]$, that is,

$$p \sim (\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) \quad \text{for all } p \in \Delta. \tag{5}$$

This establishes the existence of a function φ such that (ii) is satisfied.

Now we prove (iii). Suppose, by way of contradiction, that there exist $p, q \in \Delta$ such that $p \sim q$ and $\varphi(p) \neq \varphi(q)$. Without loss of generality, suppose that $\varphi(q) > \varphi(p)$. By monotonicity,

$$(\varphi(q), 0, \ldots, 0, 1 - \varphi(q)) \succ (\varphi(p), 0, \ldots, 0, 1 - \varphi(p)). \tag{6}$$

Furthermore, because

$$(\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) \sim p$$

and

$$p \sim q$$

and

$$q \sim (\varphi(q), 0, \ldots, 0, 1 - \varphi(q)),$$

it follows that

$$(\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) \succsim (\varphi(q), 0, \ldots, 0, 1 - \varphi(q))$$

which, together with (6), contradicts Suzumura consistency.

To prove (iv), suppose, by way of contradiction, that there exist $p, q \in \Delta$ such that $p \succ q$ and $\varphi(p) \leq \varphi(q)$. By monotonicity,

$$(\varphi(q), 0, \ldots, 0, 1 - \varphi(q)) \succsim (\varphi(p), 0, \ldots, 0, 1 - \varphi(p)).$$

Furthermore, because we also have

$$q \sim (\varphi(q), 0, \ldots, 0, 1 - \varphi(q))$$

and

$$(\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) \sim p,$$

it follows that $q \overline{\succsim} p$ which, together with $p \succ q$, leads to a contradiction of Suzumura consistency. □

Now we can state and prove our main result.

Theorem 2 *Suppose that X contains at least three alternatives and that $\succsim$ is a relation on Δ such that $(1, 0, \ldots, 0) \succ (0, \ldots, 0, 1)$. The relation $\succsim$ satisfies Suzumura consistency, solvability, monotonicity and independence if and only if there exists a function $U\colon X \to \mathbb{R}$ such that the pair $(\succsim, U)$ satisfies*

(0) $U(x_1) = 1$ *and* $U(x_n) = 0$;
(i) $\left[(\alpha, 0, \ldots, 0, 1 - \alpha) \succsim (\beta, 0, \ldots, 0, 1 - \beta) \Leftrightarrow \alpha \geq \beta\right]$ *for all* $\alpha, \beta \in [0, 1]$;
(ii) $p \sim \left(\sum_{i=1}^{n} p_i U(x_i), 0, \ldots, 0, 1 - \sum_{i=1}^{n} p_i U(x_i)\right)$ *for all* $p \in \Delta$;
(iii) $p \sim q \Rightarrow \sum_{i=1}^{n} p_i U(x_i) = \sum_{i=1}^{n} q_i U(x_i)$ *for all* $p, q \in \Delta$;
(iv) $p \succ q \Rightarrow \sum_{i=1}^{n} p_i U(x_i) > \sum_{i=1}^{n} q_i U(x_i)$ *for all* $p, q \in \Delta$.

Proof ‘If.’ Solvability and monotonicity follow from (i) and (ii). That Suzumura consistency is satisfied is an immediate consequence of substituting $\varphi(p) = \sum_{i=1}^{n} p_i U(x_i)$ for all $p \in \Delta$ and applying the requisite result of the ‘if’ part of Theorem 1.

To prove that the members of the class of decision rules identified in the statement of Theorem 2 satisfy independence, suppose $p, q \in \Delta$ and $\alpha, \beta \in [0, 1]$ are such that $p \sim (\alpha, 0, \ldots, 0, 1 - \alpha)$ and $q \sim (\beta, 0, \ldots, 0, 1 - \beta)$. By (i) and (ii) and the uniqueness of α and β, this means that $\alpha = \sum_{i=1}^{n} p_i U(x_i)$ and $\beta = \sum_{i=1}^{n} q_i U(x_i)$. We have to show that, for all $\gamma \in [0, 1]$,

$$\gamma p + (1-\gamma)q \sim \gamma\left(\sum_{i=1}^{n} p_i U(x_i), 0, \ldots, 0, 1 - \sum_{i=1}^{n} p_i U(x_i)\right)$$
$$+ (1-\gamma)\left(\sum_{i=1}^{n} q_i U(x_i), 0, \ldots, 0, 1 - \sum_{i=1}^{n} q_i U(x_i))\right)$$

or, equivalently,

$$\gamma p + (1-\gamma)q \sim \left(\sum_{i=1}^{n}(\gamma p_i + (1-\gamma)q_i)U(x_i), 0, \ldots, 0,\right.$$
$$\left. 1 - \sum_{i=1}^{n}(\gamma p_i + (1-\gamma)q_i)U(x_i)\right). \tag{7}$$

Letting $\gamma p + (1-\gamma)q$ play the role of p in (ii), (7) follows.

'Only if.' The function φ can be constructed as in the proof of Theorem 1. Now define $U: X \to \mathbb{R}$ by letting

$$U(x_i) = \varphi(e^i) \quad \text{for all } i \in \{1, \ldots, n\}. \tag{8}$$

Substituting $p = e^1$ in (5) and using (8), it follows that

$$e^1 = (1, 0, \ldots, 0) \sim (U(x_1), 0, \ldots, 0, 1 - U(x_1))$$

and $U(x_1) = 1$ follows immediately from uniqueness which, in turn, follows from Suzumura consistency and monotonicity. That $U(x_n) = 0$ follows analogously and, thus, the proof of (0) is complete.

Part (i) is an immediate consequence of monotonicity.

In view of parts (ii), (iii) and (iv) of Theorem 1, the proof is complete once we show that

$$\varphi(p) = \sum_{i=1}^{n} p_i U(x_i) \quad \text{for all } p \in \Delta. \tag{9}$$

Let $p, q \in \Delta$ and $\gamma \in [0, 1]$. By definition of φ, we have

$$p \sim (\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) \;\text{ and }\; q \sim (\varphi(q), 0, \ldots, 0, 1 - \varphi(q)).$$

By independence,

$$\begin{aligned}\gamma p + (1-\gamma)q &\sim \gamma(\varphi(p), 0, \ldots, 0, 1 - \varphi(p)) + (1-\gamma)(\varphi(q), 0, \ldots, 0, 1 - \varphi(q))\\ &= (\gamma\varphi(p)+(1-\gamma)\varphi(q), 0, \ldots, 0, \gamma(1-\varphi(p))+(1-\gamma)(1-\varphi(q)))\\ &= (\gamma\varphi(p) + (1-\gamma)\varphi(q), 0, \ldots, 0, 1 - [\gamma\varphi(p) + (1-\gamma)\varphi(q)])\,.\end{aligned}$$

By definition of φ, this means that

$$\varphi(\gamma p + (1-\gamma)q) = \gamma\varphi(p) + (1-\gamma)\varphi(q). \tag{10}$$

We now use the definition of U in (8) and the functional equation (10) to prove (9). The proof proceeds by induction on the number of positive components of p, that is, on the cardinality of the set $\{i \in \{1, \ldots, n\} \mid p_i > 0\}$. This step in our proof is borrowed from Kreps's (1988) proof of a version of the classical expected-utility theorem.

If $\{i \in \{1, \ldots, n\} \mid p_i > 0\}$ contains a single element j, it follows that $p = e^j$. Clearly, $\varphi(e^j) = \sum_{i=1}^n p_i U(x_i) = U(x_j)$ in this case and (9) is satisfied.

Now let $1 < m \leq n$ and suppose (9) is satisfied for all probability distributions in Δ with $m - 1$ positive components. Let p be such that p has m positive components and let $i \in \{1, \ldots, n\}$ be such that $p_i > 0$. Define a distribution $q \in \Delta$ by

$$q_j = \begin{cases} 0 & \text{if } j = i, \\ \frac{p_j}{1-p_i} & \text{if } j \neq i. \end{cases}$$

By definition, q has $m - 1$ positive components and we can express p as

$$p = p_i e^i + (1 - p_i)q.$$

By (10),

$$\varphi(p) = p_i \varphi(e^i) + (1 - p_i)\varphi(q)$$

and, using (8) and applying the induction hypothesis to q, it follows that

$$\varphi(p) = p_i U(x_i) + (1 - p_i) \sum_{\substack{j=1 \\ j \neq i}}^{n} \frac{p_j}{1 - p_i} U(x_i) = \sum_{i=1}^{n} p_i U(x_i),$$

as was to be shown. □

The function U in the theorem statement is often referred to as a von-Neumann-Morgenstern function.

As mentioned in the introduction, Fishburn's (1971, 1972) one-way expected-utility results are concerned with part (iv) of the above theorem only and, moreover, he does not invoke Suzumura consistency in any of his observations.

In most of the formulations of the classical expected-utility theorem, reflexivity and completeness are implied due to the presence of transitivity in the set of axioms employed. Clearly, this is not the case if transitivity is weakened to Suzumura consistency. Moreover, not even the restriction of $\succsim$ to pairs of unit vectors needs to be an ordering. That is, Suzumura consistency in conjunction with the remaining axioms of Theorem 2 is not sufficient to guarantee that certain alternatives (the elements of X) are fully ordered.

The classical expected-utility criterion is a special ("maximal") case of the class characterized in Theorem 2. Another special ("minimal") case is obtained if no preferences are added to those imposed by (i) and (ii), that is, the case in which $\succsim$ is defined by letting

$$\left[(\alpha, 0, \ldots, 0, 1-\alpha) \succsim (\beta, 0, \ldots, 0, 1-\beta) \Leftrightarrow \alpha \geq \beta\right] \quad \text{for all } \alpha, \beta \in [0, 1]$$

and

$$p \sim \left(\sum_{i=1}^{n} p_i U(x_i), 0, \ldots, 0, 1 - \sum_{i=1}^{n} p_i U(x_i)\right) \quad \text{for all } p \in \Delta,$$

where $U(x_1) = 1$ and $U(x_n) = 0$.

4 Preference reversals

Preference reversals manifest themselves when agents (usually experimental subjects) exhibit a pattern of rankings that cannot be reconciled with the theory under examination—in our case, the classical expected-utility criterion. Specifically, consider two pairs of probability distributions p versus q and p' versus q'. A standard preference reversal occurs when an agent prefers p to q and q' to p' but even if each of the two rankings can be generated by some von-Neumann-Morgenstern function, the two functions do not correspond to the same expected-utility criterion. That is, there may exist von-Neumann-Morgenstern functions U and V such that the ranking $p \succ q$ is consistent with the criterion associated with U and the ranking $q' \succ p'$ can be generated by means of V but U and V cannot possibly generate the same decision criterion. Typical examples of such preference reversals are the common consequence effect and the common ratio effect discussed in Machina (1983).

The use of members of the class characterized in Theorem 2 allows us to circumvent essentially all preference reversals. As long as at least one of the two rankings $p \succ q$ and $q' \succ p'$ is compatible with some von-Neumann-Morgenstern function, a suitable choice of a generalized expected-utility criterion allows us to avoid a preference reversal. To see that this is indeed the case, suppose that the ranking $p \succ q$ can be obtained by means of the expected-utility criterion associated with a function U but the ranking $q' \succ p'$ is incompatible with the use of U as a von-Neumann-Morgenstern function. It is clear that, in such a case, a generalized expected-utility criterion involving U can be defined such that we have $p \succ q$ according to this criterion but p' and q' are non-comparable. Thus, our class is remarkably rich and flexible, an observation that we now illustrate by means of two prominent examples.

Consider first a special case of the common consequence effect [Machina (1983)]—namely, the well-known Allais (1953) paradox. Suppose that we have a set of alternatives $X = \{5, 1, 0\}$, where $x_1 = 5$ stands for receiving five million dollars, $x_2 = 1$ means receiving one million dollars and $x_3 = 0$ is an alternative in which the agent receives zero. Experimental evidence [see Kahneman and Tversky (1979), for instance] suggests that many subjects express something resembling the following rankings of specific probability distributions. Consider the pair of distributions $p = (0, 1, 0)$ versus $q = (0.1, 0.89, 0.01)$, on the one hand, and the pair of distributions $p' = (0, 0.11, 0.89)$ versus $q' = (0.1, 0, 0.9)$, on the other. Many agents appear to rank p as being better than q and q' as being better than p'. However, the combination of these two rankings is inconsistent with classical expected utility theory.

Indeed, for any von-Neumann-Morgenstern function $U: X \to \mathbb{R}$ such that $U(5) = 1$ and $U(0) = 0$, $p \succ q$ entails $U(1) > 0.1\,U(5) + 0.89\,U(1) + 0.01\,U(0)$ and, thus, $U(1) > 1/11$, whereas $q' \succ p'$ implies $0.1\,U(5) + 0.9\,U(0) > 0.11\,U(1) + 0.89\,U(0)$ and, thus, $U(1) < 1/11$, a contradiction. But these rankings can be reconciled with a generalized expected-utility criterion: pick a von-Neumann-Morgenstern function U such that $U(1) > 1/11$ and choose a generalized criterion from our class such that $p \succ q$ and p' and q' are non-comparable. Alternatively, of course, it is possible to pick V such that $V(1) < 1/11$, $q' \succ p'$ and p and q are non-comparable. Either option eliminates the offending preference reversal.

As a second example, we discuss the certainty effect [Kahneman and Tversky (1979)] which is a specific instance of the common ratio effect [again, see Machina (1983)]. Now suppose that $X = \{6, 3, 0\}$, where $x_1 = 6$ represents receiving \$6,000, $x_2 = 3$ means that the agent receives \$3,000, and $x_3 = 0$ is an alternative such that the agent receives zero. Furthermore, consider the pair of distributions $p = (0, 0.9, 0.1)$ versus $q = (0.45, 0, 0.55)$, on the one hand, and the pair of distributions $p' = (0, 0.002, 0.998)$ versus $q' = (0.001, 0, 0.999)$, on the other. Now a common pattern that emerges in experimental studies is that p is ranked as being better than q and q' is ranked as being better than p'. As is the case for the Allais paradox, the combination of these two rankings cannot be reconciled with classical expected utility theory. For any von-Neumann-Morgenstern function $U: X \to \mathbb{R}$ such that $U(6) = 1$ and $U(0) = 0$, $p \succ q$ means that $0.9\,U(3) + 0.1\,U(0) > 0.45\,U(6) + 0.55\,U(0)$ and, thus, $U(3) > 1/2$. But $q' \succ p'$ implies $0.001\,U(6) + 0.999\,U(0) > 0.002\,U(3) + 0.998\,U(0)$ and, thus, $U(3) < 1/2$, which establishes the desired contradiction. This preference-reversal problem can be resolved by means of a generalized expected-utility criterion in analogy with the previous example.

5 Concluding remarks

We view our contribution as being primarily theoretical in nature. Nevertheless, it seems to us that our new class of generalized expected-utility criteria has the potential to be useful in some experimental settings as well. There are several possible stances that one may take in the face of observed preference reversals. A common response consists of exploring theories that fundamentally diverge from the classical expected-utility criterion in the sense that the traditional expected-utility axioms are weakened or removed altogether, a path that has been followed in numerous earlier contributions; see the introduction for examples of this literature. Alternatively, one may consider the insistence on the completeness requirement to be at the root of the supposedly paradoxical outcome. Allowing for incomplete rankings is an approach that has been followed in some of the existing literature but if the remaining standard expected-utility axioms are to be retained, this can only be done by means of a weakening of transitivity; see, for instance, Kreps (1988).

Clearly, if transitivity is removed from the list of properties and no coherence condition is imposed at all, the resulting class of decision rules becomes much too rich and contains many rather undesirable rules, such as those that produce cyclical preferences. This is where our crucial property of Suzumura consistency comes into play: we employ this natural weakening of transitivity, thereby imposing some discipline on the set of admissible decision rules. The restriction thereby obtained clearly is non-trivial and lends itself to empirical testing. If the agents are given the option of treating two distributions as non-comparable instead of being forced to express a relative ranking, it may very well be the case that, according to the decisions of some of them, p is ranked as better than q and p' and q' are non-comparable (or p and q are non-comparable and q' is preferred to p'). Of course, the preference-reversal problem persists if the original rankings are retained even in the presence of a non-comparability option. But it seems to us that the use of an incomplete generalized expected-utility criterion may considerably reduce the instances of dramatically conflicting pairwise rankings without abandoning the core principles of expected-utility theory altogether. The common consequence effect and the common ratio effect defined by Machina (1983) are important examples to illustrate the considerably richer class of decision rules identified in our results but the potential use of our class is much broader—as established in the previous section, its reach extends to essentially all instances of preference reversals.

Acknowledgments We thank Clemens Puppe, an associate editor and a referee for their valuable comments on an earlier version of the paper. Financial support from a Grant-in-Aid for Specially Promoted Research from the Ministry of Education, Culture, Sports, Science and Technology of Japan for the Project on Economic Analysis of Intergenerational Issues (grant number 22000001), the Fonds de Recherche sur la Société et la Culture of Québec, and the Social Sciences and Humanities Research Council of Canada is gratefully acknowledged.

References

Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Americaine. Econometrica 21:503–546
Aumann RJ (1962) Utility theory without the completeness axiom. Econometrica 30:445–462
Aumann RJ (1964) Utility theory without the completeness axiom: a correction. Econometrica 32:210–212
Aumann RJ (2000) Collected papers, vol 1. MIT Press, Cambridge, MA
Bell DE (1982) Regret in decision making under uncertainty. Oper Res 30:961–981
Bossert W, Sprumont Y, Suzumura K (2005) Consistent rationalizability. Economica 72:185–200
Bossert W, Suzumura K (2010) Consistency, choice, and rationality. Harvard University Press, Cambridge, MA
Camerer C (1995) Individual decision making. In: Kagel JH, Roth AE (eds) Handbook of experimental economics. Princeton University Press, Princeton, NJ, pp 587–704
Dubra J, Maccheroni F, Ok EA (2004) Expected utility theory without the completeness axiom. J Econ Theory 115:118–133
Fishburn PC (1971) One-way expected utility with finite consequence spaces. Ann Math Stat 42:572–577
Fishburn PC (1972) Alternative axiomatizations of one-way expected utility. Ann Math Stat 43:1648–1651
Fishburn PC (1982) Nontransitive measurable utility. J Math Psychol 26:3l–67
Fishburn PC (1991) Nontransitive preferences in decision theory. J Risk Uncertain 4:113–134
Fishburn PC, LaValle IH (1988) Context-dependent choice with nonlinear and nontransitive preferences. Econometrica 56:1221–1239
Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. Econometrica 47:263–291
Kreps DM (1988) Notes on the theory of choice. Westview Press, Boulder, CO
Loomes G, Sugden R (1982) Regret theory: an alternative theory of rational choice under uncertainty. Econ J 92:805–824
Luce RD (1956) Semiorders and the theory of utility discrimination. Econometrica 24:178–191
Luce RD, Raiffa H (1957) Games and decisions. Wiley, New York, NY

Machina MJ (1983) Generalized expected utility analysis and the nature of observed violations of the independence axiom. In: Stigum BP, Wenstøp F (eds) Foundations of utility and risk theory with applications. Reidel, Dordrecht, pp 263–293

Rabin M (1998) Psychology and economics. J Econ Lit 36:11–46

Sen AK (1969) Quasi-transitivity, rational choice and collective decisions. Rev Econ Stud 36:381–393

Simon HA (1955) A behavioral model of rational choice. Q J Econ 69:99–118

Suzumura K (1976) Remarks on the theory of collective choice. Economica 43:381–390

Suzumura K (1978) On the consistency of libertarian claims. Rev Econ Stud 45:329–342

Szpilrajn E (1930) Sur l'extension de l'ordre partiel. Fundam Math 16:386–389

Thrall RM (1954) Applications of multidimensional utility theory. In: Thrall RM, Coombs CH, Davis RL (eds) Decision processes. Wiley, New York, NY, pp 181–186

Tversky A, Thaler RH (1990) Anomalies: preference reversals. J Econ Perspect 4:201–211

von Neumann J, Morgenstern O, (1944, second ed 1947) Theory of games and economic behavior. Princeton University Press, Princeton, NJ