

Behavioral assumptions for a class of utility theories: A program of experiments

R. Duncan Luce

Published online: 16 June 2010

© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract An outline is given of the behavioral properties (axioms) that have been proposed, and to some extent empirically evaluated, concerning uncertain (often risky) alternatives, the joint receipt of alternatives, and possible linking properties. Recent theoretical work has established the existence of three inherently distinct risk types of people—risk seeking, risk neutral, and risk averse—and so evaluations of theories must take respondent type into account. A program of experiments is sketched making clear exactly which empirical studies need to be repeated with respondents partitioned by risk type.

Keywords Gambles · Joint receipt · p-Additive representation · Risk types · Uncertain alternatives

JEL Classification C91 · D46 · D81

The term “behavioral economics” or “experimental economics” seems to have been mostly used for experimental studies of various models of economic interactions. But in my view it should also include experimental evaluation of the behavioral axioms underlying theories of individual behavior which have

This work was supported in part by the University of California, Irvine. I also thank a number of people who have commented on the manuscript: Drs. Young-Hee Cho, Enrico Diecidue, George Karabatsos, A. A. J. Marley, Arthur Merin, Ragnar Steingrímsson, Clinton Davis-Stober, and Michael H. Birnbaum who, as a referee, made many helpful suggestions.

R. D. Luce (✉)
Institute for Mathematical Behavioral Sciences, University of California,
Irvine, CA 92697-5100, USA
e-mail: rdluce@uci.edu

been evolving for more or less realistic economic situations and leading to forms of utility representations.

During the past 20 years, I have been fairly deeply involved in trying to understand how two distinct economic structures inter-relate and the sorts of numerical representations that can arise as a result of formalized assumptions concerning these inter-relations. The one structure has to do with consequences attached either to uncertain events resulting in what are often called “gambles” or to risky events (i.e., with known probability of events occurring) resulting in what are often called “lotteries”. This certainly has been the mainstay for well over half a century of anything that purports to be a theory of utility for uncertain or risky situations. The second structure has to do with the concatenation of valued items which is called their joint receipt (see Section 1.2 for more detailed discussion of this concept). This riskless aspect of utility seems functionally to have been declared to be outside the scope of utility theory even though behavioral axiomatizations of uncertain alternatives result in a utility function over riskless alternatives which are now taken for granted. Nonetheless, a few of us think that joint receipt is inherent to economic situations and that by including it we gain considerable richness of structure that can be effectively exploited.

My purpose here is to formulate in one place the key behavioral assumptions (axioms) that theorists have proposed and that might be evaluated empirically in, first, the (binary) case of $n = 2$ branches, and then in the general case of n branches which is approached by starting with the binary results and studying recursive relations known as branching and upper gamble decomposition (see Section 3). The proposed experimental program is quite large, and I am unable to undertake it because I no longer teach or supervise graduate students and do not have a suitable laboratory. My hope is that this article may stimulate others to do so.

In many ways, the type of behavioral science this represents is intellectually far closer to the types of physics that arose during the 16th through 19th centuries—the study of macroscopic phenomena and the discovery of the laws of mechanics, motion, thermodynamics, hydro-dynamics, electromagnetism, relativity, etc.—than it is “to the opening of black boxes”—atomic structure, quantum physics, much of planetary theory, plate tectonics, geology, etc. typical of much physics in the 20th and first decade of the 21st centuries. What we currently seem able to do is attempt to discover the behavioral axioms, some of which are invariances, that form compact summaries of behavioral regularities. And sometimes these axioms formulate enough structure to be able to derive numerical representations of them. A well known example concerns the axioms giving rise to the subjective expected utility (SEU) representation (originated by Savage 1954). Of course, I fully realize that huge, very expensive efforts are being made, often involving computational brain models that are loosely tied to imaging data, that are intended to open the black box that is the human brain and/or mind, but those “internal networks” never seem to become firmly agreed upon. Important though it may be, such reverse engineering is inherently very, very difficult, as modern physicists

and engineers are perfectly aware, especially when one is not certain what the appropriate components are. Imagine trying to infer what gives rise to observable behavior of some computer when you don't know of what electrical components went into its construction. So the focus here is the "laws" that lead to representations and that need to be checked experimentally.

It is important to recognize that utility representations, e.g., SEU, neither reside in the mind of the decision maker (DM) nor are used directly by the DM to make choices any more than the partial differential equations of classical physics and their solutions lurk within the objects whose behavior they characterize. Both representations are creations of scientific attempts to summarize compactly the relevant behavior, thereby making it more convenient to derive predictions from the collection of behavioral "laws."

That is exactly why we bother to work out the representations: to discover behavioral implications of the underlying assumptions (ultimately called "laws," should they withstand empirical testing). We will see a vivid example of this below in Sections 1.2.3 and 1.2.4. Further, a great deal of the empirical research that the psychologist Michael H. Birnbaum has used to attack the class of rank-dependent models (including the famed cumulative prospect theory, CPT, of Tversky and Kahneman 1992) is of that character. Of course, aficionados of CPT simply dismiss or ignore Birnbaum's findings. For a summary of his results until 2004 with detailed references to his relevant publications, see Marley and Luce (2005) and also Birnbaum (2008).

My aim here is to propose an empirical program having two distinct parts. The first is designed to evaluate several interrelated putative *binary* "laws" in a fashion that should pinpoint which, if any, assumptions appear to be wrong. Once the binary case is clarified, we turn to some possible generalizations to gambles with $n > 2$ branches. These are based on inductive axioms that simply have not yet been evaluated empirically. One quite novel feature of the program is that the binary results lead to a qualitative classification of people into 3 risk types (Section 1.2.3), which for good reasons I call "risk seeking," "risk neutral," and "risk averse."¹ This partition must be taken into account in empirically evaluating properties of the model. This, of course, complicates the experimental program considerably.

Of course behavioral properties for which the type distinction does not matter need not be repeated; these are listed in Section 4.1 and include transitivity and monotonicity of preference, commutativity and associativity of joint receipt, and several less familiar concepts. The focus of the rest of Section 4 is on how to decide for individual respondents which type they are and how the behavioral laws linking gambles and joint receipt vary with type.

Formal theorems and their proofs are not stated here.

¹These terms are commonly used referring to shapes of utility functions, and I relate my classification to them, in particular to the Arrow-Pratt risk classification (Section 2.3).

1 Underlying empirical structures

1.1 Primitives

1.1.1 Consequences and preference order

Suppose that X is the (rich) set of valued, pure consequences under consideration. By “pure” I mean viewed as certain, i.e., without perceived risk. A special, but important, example is money. But X can include far more substantive consequences than that, e.g., goods at stores when they are viewed as riskless, etc.

We postulate that the DM exhibits preferences over consequences that can be summarized as a *weak (preference) order* \succsim over X . The key property of \succsim is transitivity:

$$\text{If } x \succsim y \text{ and } y \succsim z, \text{ then } x \succsim z. \quad (1)$$

As usual, $x \sim y$ means that both $x \succsim y$ and $y \succsim x$ hold. Because the indifference relation \sim is an equivalence relation, we are in reality working with equivalence classes.

In principle transitivity can be studied empirically, but in practice some fairly troublesome statistical issues have been encountered in making this evaluation. Some data analyses proved to be in error, but the current consensus is that we need not reject transitivity. The details until 1999 are found in Luce (2000, pp. 37–39). Crucially important were analyses of Iverson and Falmagne (1985), Regenwetter and Davis-Stober (2008), and Regenwetter et al. (2010). A somewhat different explanation of the data is offered by Birnbaum and Gutierrez (2007).

1.2 Joint receipts (JR)

Suppose that $x, y \in X$, then let $x \oplus y \in X$ mean having both x and y . Thus \oplus is a binary operation (of *joint receipt*, often abbreviated JR) on X . Examples of joint receipt are ubiquitous—anytime one shops for two (or more) things, that purchase results in a joint receipt of goods. Buying a portfolio of financial assets is another example. There are subtle issues of definition of a good. In buying a pair of shoes, is that the joint receipt of the left shoe and the right shoe or is it simply that the pair of shoes is a unitary good? The complications of the former interpretation far exceed its usefulness to the consumer.

We assume that $\langle X, e, \oplus, \sim \rangle$ satisfies the usual axioms of what is technically called an abelian (weakly commutative) group with identity e (see Cho and Luce 1995; Cho et al. 1994). In particular, in addition to \sim being an equivalence relation, the assumptions are that for all $x, y, z \in X$, there is an element $e \in X$, called “no change from the status quo,” that is an *identity* of the operation:

$$x \oplus e \sim x, \quad (2)$$

that *commutativity* holds:

$$x \oplus y \sim y \oplus x \tag{3}$$

and that *associativity* holds:

$$(x \oplus y) \oplus z \sim x \oplus (y \oplus z). \tag{4}$$

The binary definition of \oplus extends to any finite number of goods because we have assumed that \oplus is associative. In the psychophysical context and for some interpretations of \oplus , failures of (3) have been found. I know of nothing comparable in utility theory.

If $x \in X$ and if $x \succsim e$, then x is called a *gain*. If $x \precsim e$, then x is called a *loss*. Clearly joint receipt of two gains is a gain and of two losses is a loss. Joint receipt of a gain and a loss can be perceived as either a gain or a loss. As we shall see, the mixed case of both gains and losses is an ever present complication in the theory.

In principal, the above axioms can be experimentally evaluated although I am unaware of such a study. For associativity, presumably one would determine for (4) certainty equivalents, i.e., the pure consequence indifferent to the more complex object, for each side

$$u \sim (x \oplus y) \oplus z, \quad u' \sim x \oplus (y \oplus z), \tag{5}$$

and then ask whether (statistically²)

$$u \approx u'. \tag{6}$$

A similar test of commutativity is also possible.

Two psychophysical examples are auditory pure tones of different intensities to the two ears and a similar visual one with light patches of different intensities to the two eyes. Such tests have been conducted by Steingrimsson and Luce (2005) and by Steingrimsson (2009, 2010). If we denote the respective stimuli (x, y) and (y, x) and match them to (u, u) and (u', u') the question is whether or not (6) holds within the accuracy of the data. References in either of these articles gives the theoretical background.

Both here and in some later cases, such as the Thomsen condition, we run afoul of empirical experience that attempts to deal with an indifference \sim by having the respondent provide matches by pure alternatives has proved to be more problematic than are choices of order based upon \succsim . Dealing with this is experimentally important and it is easy to become confused.

1.2.1 Hölder’s axioms

Our operation \oplus plays a role analogous to the concatenation operations of elementary physics, e.g., two masses on a pan balance, two rods abutted, etc.

²This is quite a tricky issue that has received some attention in psychophysics, but hardly an accepted solution.

We assume that $\langle X, e, \oplus, \succsim \rangle$ on equivalence classes is a solvable, Archimedean ordered, abelian group with an isomorphism onto the additive real numbers. For the equivalence classes, the operation \oplus is closed, has an identity, is commutative and associative, and each element has an inverse satisfying the usual axioms of a solvable, Archimedean ordered, abelian group (Hölder 1901; Krantz et al. 1971, Chapters 2 and 3).

The key testable assumption beyond transitivity is *monotonicity*: For all $x, y, z \in X$,

$$x \succsim y \Leftrightarrow x \oplus z \succsim y \oplus z. \quad (7)$$

The literature up to 1999 is discussed in Luce (2000, pp. 137–139, 237–238) and it seems favorable toward monotonicity of joint receipt. For a later and very general discussion of testing many of the axioms I will mention, see the important article by Karabatsos (2005).

1.2.2 *p*-Additive representations

Classically, and certainly in Hölder's theorem as usually formulated, the only representations that are studied are mappings into the real numbers denoted \mathbb{R} (or sometimes into the non-negative real numbers denoted \mathbb{R}^+) under just addition, i.e., into the structure $\langle \mathbb{R}, \geq, + \rangle$. But recall that the typical theories for uncertain alternatives, such as SEU and CPT, have representations that involved both addition $+$ and multiplication \times . So, why not admit the possibility that the representations of \oplus are onto suitable (defined below) subintervals of $\langle \mathbb{R}, \geq, +, \times \rangle$ that are closed under both addition and multiplication?

Admitting that possibility, under the usual Hölder assumptions, the possible polynomial representations are of the form:

$$U(x \oplus y) = U(x) + U(y) + \delta U(x)U(y), \quad \delta = -1, 0, 1, \quad (8)$$

where U is order preserving and $U(e) = 0$. These are called *p*-additive because they are the only polynomial forms with $U(e) = 0$ that transform into addition. If we limit ourselves to mappings of the form

$$U(x \oplus y) = F(U(x), U(y)),$$

where $F(u, v)$ is a function of u, v that is expressible in terms of the field operations $+$ and \times , along with the induced subtraction and division operations, and assuming F is a rational function, then (8) constitute the only representations.³

³Personal communication from C. T. Ng, November 17, 2009.

1.2.3 Three types of people

Corresponding to the value of δ , there are 3 classes or types of people. As we shall see, these types are strikingly different and so I believe that our proposed experiments must be partitioned accordingly for the data to be meaningful.

When $\delta = 0$, the representation is

$$U(x \oplus y) = U(x) + U(y), \tag{9}$$

which is a purely additive ratio scale (unique to choice of unit) which certainly is the type that has been mostly studied during the past 60 years. As Karabatsos (2005) showed, additivity is not well sustained in general.

When $\delta \neq 0$, we may rewrite (8) in terms of U as

$$1 + \delta U(x \oplus y) = [1 + \delta U(x)][1 + \delta U(y)], \tag{10}$$

which means there is a representation in terms of the transformation

$$V(x) := 1 + \delta U(x), \tag{11}$$

which satisfies the multiplicative property

$$V(x \oplus y) = V(x)V(y). \tag{12}$$

Note that no scale factor $\alpha \neq 1$ maintains this multiplicative representation, (12). In general, however, such a multiplicative representation is unique only up to an arbitrary power, $V \rightarrow V^\beta$, $\beta > 1$, but as we shall see even that degree of freedom is lost. Of course, from (12) it is immediate that $\ln V$ is an additive ratio scale.

1.2.4 Two scale types: ratio and absolute

As mentioned earlier, for those people satisfying (9), the utility function is a ratio scale, i.e., it is unique up to its unit. But for the other two types, $\delta \neq 0$, U must be an absolute scale. This follows from the fact that in (11) U is either added to or subtracted from 1. Of course, V itself is unique up to positive powers and so $\ln V$ is an additive ratio scale. The V scale maps onto $]0, \infty[$ in both non-zero cases. The difference being that V is order preserving for $\delta = 1$ and order reversing for $\delta = -1$. Put in words, U in the case $\delta = 0$ has a free unit, as say with mass, whereas for the cases with $\delta \neq 0$ there is no freedom in the choice of unit, as with probability. These differences are very significant as we see below.

1.2.5 An experimental-procedural implication

The fact that there are 3 types of people corresponding to $\delta = -1, 0, 1$, which as we shall see are qualitatively quite different, should be significant for the experimenter. It means that each participant in the study must be evaluated for type before any further data are examined. A criterion for determining type is provided in Section 2 below. Moreover, that fact certainly means that group

averages, including medians and comparisons of distributions, over an un-screened population are *totally* meaningless. Although averaging is a bit more justified over people of the same type, the fact that utility functions have the same (non-linear) form but with different parameters (see Section 2.4) means that here too the averaging of raw data is not really justified. This has received some attention in other areas, such as psychophysics, where the upshot is that only averages of linear functions are meaningful. Moreover, the average slope parameter equals the average of the individual parameters. Thus, in the present case, one must transform the data to linear form as outlined in Section 2.4.

Although the Hölder axioms clearly have interesting consequences, there is troubling empirical evidence (e.g., Luce 2000; Sneddon and Luce 2001) that some properties seem to hold for gains ($x \succsim e$) and, separately, for losses ($x \precsim e$) but not very well for gambles involving mixed gains and losses. Schneider and Lopes (1986) clearly ran into the same issues with gambling behavior. Whether or not this reflects the fact that a substantial number of respondents are of types $\delta \neq 0$ has not yet been carefully explored; it seems to me to be very important to do so.

1.3 Uncertain alternatives—gambles

1.3.1 Uncertain alternatives

We assume there are a great many families of chance “experiments,” i.e., sources of uncertainty. We denote by C, D , etc. typical disjoint chance events arising within a family of chance “experiments.” A general uncertain alternative—often, although somewhat misleadingly, called a *gamble*, which nonetheless is the term that I use—begins with a partition of the underlying chance event into n subevents, and to each is assigned a consequence, either an element of X or a first-order gamble. See Luce (2003).

A *binary* uncertain alternative is a gamble with $n = 2$ branches. We may think of it as an experiment whose “universal set” Ω may be partitioned into C and $\bar{C} := \Omega \setminus C$. Each sub-event leads to its own consequence, so that the gamble has two chance branches (x, C) and (y, \bar{C}) . We write it as $(x, C; y, \bar{C})$.

Two properties of binary gambles that have received some attention are *event commutativity*

$$\left((x, C; y, \bar{C}), D; y, \bar{D} \right) \sim \left((x, D; y, \bar{D}), C; y, \bar{C} \right), \tag{13}$$

and *right autodistributivity*

$$\left((x, C'; y, \bar{C}'), C; z, \bar{C} \right) \sim \left((x, C'; z, \bar{C}'), C; (y, C''; z, \bar{C}'), \bar{C} \right), \tag{14}$$

where the primes refer to independent realizations of the underlying chance experiment. The reason they are of some interest is made clear in Section 1.5.

1.3.2 Unitary gambles and separable representations

One important subclass of binary gambles has $y = e$; these are called *unitary*. One axiomatic issue turns out to be whether such gambles have a multiplicative conjoint representation:

$$U^*(x, C; e, \bar{C}) = U^*(x)S_{\Omega}^*(C), \tag{15}$$

which is called a *separable representation*.

The necessary axioms are well known (Krantz et al. 1971, Ch. 6) to be: transitivity of \succsim , *monotonicity*,⁴ i.e., if for non-null events $C, D, E \in \Omega$,

$$(x, C; e, \bar{C}) \succsim (y, C; e, \bar{C}) \Leftrightarrow (x, D; e, \bar{D}) \succsim (y, D; e, \bar{D}), \tag{16}$$

and the *Thomsen condition*:⁵

$$\begin{aligned} (x, E; e, \bar{E}) \sim (z, D; e, \bar{D}) \ \& \ (z, C; e, \bar{C}) \sim (y, E; e, \bar{E}) \\ \Rightarrow (x, C; e, \bar{C}) \sim (y, D; e, \bar{D}). \end{aligned} \tag{17}$$

The monotonicity assumption has received some empirical study. For example, von Winterfeldt et al. (1997) ran a test of it based on medians and concluded that the evidence for it was mixed. However, Ho et al. (2005) carried out a far more complete statistical analysis, which was based on the entire distribution of responses, which provided strong support for monotonicity. The second axiom for separable representations, the Thomsen condition, (17), Karabatsos (2005) reanalyzed the then unpublished memory data of William H. Batchelder and Jarad Smith using Bayesian methods, and the Thomsen condition was well sustained. There is no question, however, that data on unitary gambles, as such, need to be collected and analyzed in a comparable way.

1.4 Two possible linking laws

At this point, we have two distinct measures of utility: the function U from the JR concatenation structure and U^* from the separable representation of unitary gambles. One hopes that some sort of behavioral law exists that forces them to be the same measure in the sense that there is a unique $\gamma > 0$ such that $U = (U^*)^\gamma$. Luce (2000, Theorem 4.4.6) proved that the following behavioral condition of segregation suffices. The key mathematical result was proved by efforts of Luce (1996) and Aczél et al. (1996).

⁴Actually called *independence* in the measurement literature.

⁵As originally formulated in Luce and Tukey (1964), they assumed double cancellation, which is the Thomsen condition where \sim is replaced by \succsim . The result is the same. Because of experimental “noise” it may be better empirically to evaluate double cancellation than the Thomsen condition.

1.4.1 Segregation

Segregation holds if for $x, y \in X$ with $x \succsim y$ and non-null event C ,

$$(x \oplus y, C; y, \bar{C}) \sim (x, C; e, \bar{C}) \oplus y. \quad (18)$$

Clearly this is empirically testable, and it has been (see discussion below) because of its crucial theoretical importance. To date, the empirical results have been ambiguous (Karabatsos 2005). Whether this would be clarified by taking respondent type into account remains to be seen.

1.4.2 Duplex decomposition

An alternative linking law has been proposed and studied quite a bit both theoretically and experimentally (more details below). As we will soon see, much more needs to be done.

Gambles are said to satisfy *duplex decomposition* (DD) over X if for all $x, y \in X$,

$$(x, C; y, \bar{C}) \oplus (e, C'; e, \bar{C}') \sim (x, C; e, \bar{C}) \oplus (e, C'; y, \bar{C}'), \quad (19)$$

where (C', \bar{C}') is an independent realization of (C, \bar{C}) . Unlike segregation, DD is rather non-rational: on the left one gets either x or y but not both whereas on the right there are four possible consequences: $e \sim e \oplus e$, $x \sim x \oplus e$, $y \sim e \oplus y$, $x \oplus y$.

Duplex decomposition has proved to be critical in some work on the utility of gambling (Luce et al. 2008a, b; Ng et al. 2009a, b) where we have avoided assuming *idempotence*, $(e, C'; e, \bar{C}') \sim e$, and where we have interpreted $(e, C'; e, \bar{C}')$ to be the qualitative structure that underlies the utility of gambling. Of course, we should attempt to verify, in the context of various judgments about broader class of gambles, whether or not idempotence fails.

Although DD has been closely looked at empirically (see Luce 2000, Section 6.2), the most recent empirical study of it and segregation is Cho et al. (2002, 2005). Karabatsos (2005) claims good support for DD, but those results are inherently ambiguous in the light of the theory motivating this article. Most important, the respondents should be classified as to type, as described in Section 2, when evaluating fits to a property. Also, all experimental tests to date have assumed idempotence, which property, as I said earlier, needs to be checked.

1.5 The binary representation

If $x \succsim y$, $\bar{C} = \Omega/C$, and the above assumptions hold under segregation, the binary utility representation is

$$U(x, C; y, \bar{C}) = U(x)S_{\Omega}(C) + U(y)S_{\Omega}(\bar{C}), \quad (20)$$

(for $\delta = 0$, Luce et al. 2008a, Eq. 28; for $\delta \neq 0$, Ng et al. 2009b, Eq. 39). Under duplex decomposition, idempotence, and $\delta = 0$

$$U(x, C; y, \bar{C}) = U(x)S_{\Omega}(C) + U(y)[1 - S_{\Omega}(C)]. \tag{21}$$

(Luce et al. 2008a, Eq. 60), whereas for $\delta \neq 0$

$$U(x, C; y, \bar{C}) = U(x)S_{\Omega}(C) + U(y)S_{\Omega}(\bar{C}) + \delta U(x)U(y)S_{\Omega}(C)S_{\Omega}(\bar{C}). \tag{22}$$

(Ng et al. 2009b, Eq. 47).

These representations feed into the general theory via recursive assumptions given in Section 3.

It is a simple calculation to show that event commutativity, (13), follows from the rank-dependent form (21) but not from (20) unless the weights are finitely additive:

$$S_{\Omega}(\bar{C}) = 1 - S_{\Omega}(C),$$

which is the classic *subjective expected utility* representation. Further, right autodistributivity, (14), holds only in that case.

Chung et al. (1994) evaluated event commutativity and concluded that for the most part it was sustained. Brothers (1990) in his dissertation explored right autodistributivity and found it not consistent with the data, which is consistent with the many studies that showed the inadequacy of subjective expected utility.

2 Empirical classification of individuals

2.1 Events having subjective probability 1/2

If there exist consequences $x \succ y$ and non-empty event E such that

$$(x, E; y, \bar{E}) \sim (x, \bar{E}; y, E) \tag{23}$$

then from either (20) or (21),

$$S_{\Omega}(E) = S_{\Omega}(\bar{E}) = \frac{1}{2}. \tag{24}$$

Such event partitions are called “equally likely.”

2.2 A criterion

For such events, the following criterion is established by Luce (2010): For all $x, y \in X$, with $x \succ x' \succ y \succ y'$, and for E satisfying (23),

$$\delta = \left\{ \begin{array}{c} 1 \\ 0 \\ -1 \end{array} \right\} \Leftrightarrow (x \oplus x', E; y \oplus y', \bar{E}) \left\{ \begin{array}{c} \succ \\ \sim \\ \prec \end{array} \right\} (x \oplus y, E; x' \oplus y', \bar{E}). \tag{25}$$

This is less formidable than it may seem. Luce (2010) mentioned the following examples:

$$\begin{aligned} & \left(160, \frac{1}{2}; 50, \frac{1}{2}\right) \text{ versus } \left(140, \frac{1}{2}; 70, \frac{1}{2}\right) \\ & \left(190, \frac{1}{2}; 30, \frac{1}{2}\right) \text{ versus } \left(120, \frac{1}{2}; 100, \frac{1}{2}\right) \\ & \left(1900, \frac{1}{2}; 60, \frac{1}{2}\right) \text{ versus } \left(1050, \frac{1}{2}; 910, \frac{1}{2}\right). \end{aligned}$$

For example, the first case arises when $x = 100$, $x' = 60$, $y = 40$, $y' = 10$. Notice that the two gambles of each row have the same (subjective) expectation, but that they differ in their (subjective) variance, the one on the left being more risky than the safer one on the right.

Ng et al. (2009b) have shown that the co-domains (images) of U are

$$\left\{ \begin{array}{l}] - 1, \infty[\\] - \infty, \infty[\\] - \infty, 1[\end{array} \right\} \Leftrightarrow \delta = \left\{ \begin{array}{l} 1 \\ 0 \\ -1 \end{array} \right\}. \quad (26)$$

This fact together with the criterion makes clear that the $\delta = 0$ people are oblivious to the (subjective) variance difference and so may be called *risk neutral (RN)*; those for whom $\delta = -1$ by contrast like the safe gamble and so are called *risk averse (RA)*; and those for whom $\delta = 1$ prefer the riskier gambles and so are called *risk seeking (RS)*. Such concepts have repeatedly appeared in discussions of risk taking behavior (e.g., Tversky and Kahneman 1992; Schneider and Lopes 1986) with criteria that are special cases of (25), which itself is an “if and only if” form. They typically pit a gamble against a pure consequence equivalent to the expected value as the gamble. This criterion is the special case of (25) where $x \oplus y \sim x' \oplus y'$ and idempotence holds.

It is noteworthy that (26) does not have a fourth case of doubly bounded utility functions. Why is this case missing? I simply do not know what serves to rule it out. This requires further investigation.

2.3 Arrow-Pratt measure of local risk

Other authors, mostly economists, have characterized the type of risk attitude as embodied in properties of a utility u function for money, not as a property of people themselves. Unlike my and some other psychologists' work, the forms for utility functions are usually ad hoc but with apparently desired properties. Moreover, the domain is usually assumed to represent total wealth, not increments in it. So the independent variable can never be negative. That is not true of my work.

The most famous measure, now called the Arrow-Pratt measure of *local risk aversion*, is

$$r(x) = -\frac{u''(x)}{u'(x)}. \tag{27}$$

(Pratt 1964; Arrow 1965, 1974). An excellent summary up to 1975 of both theory and applications centered on (27) is Keeney and Raiffa (1976/1993).⁶ Later discussions of this and related measures are extensive and include Holt and Laury (2002), Levy and Levy (2002), Ross (1981), and references cited in these articles.

In Section 2.4 the utility forms derived from my theory and the resulting $r(x)$ for them are stated.

As mentioned, Karabatsos’ (2005) reanalysis of other authors’ data makes clear that $\delta = 0$ is not sustained over the sample of people studied. So, a major empirical problem is to collect data from a “representative sample” of people that is not restricted to students and academics to see, first, if each person is consistent in choosing either the risky over the safe, the safe over the risky, or is indifferent between the two. Assuming that people are consistent, then what proportions of each type seem to be found? And is there any substantial correlation with social roles. Judging by informal “data” collected at several of my lectures, such academics tended to be type $\delta = -1$, i.e., risk averse with a smaller fraction of risk seekers, and a few apparently are risk neutral ($\delta = 0$).

Presumably, type may vary over different domains of activity. A person may have one risk attitude toward financial matters and a different one about the risks entailed in such apparently dangerous sports as mountain climbing and skiing.

A more general criterion, due to C. T. Ng, which does not depend upon finding an event satisfying (23), is also reported in Luce (2010). Although more general, it is not nearly as transparent as is (25).

An important empirical issue is whether there really are any people who are of type $\delta = 0$? If not, then all the earlier theories of utility simply cannot be descriptive. This would be a pretty shocking discovery, as is made clear by Regenwetter and Davis-Stober (2008); Regenwetter et al. (2010).

2.4 Utility of money

Under his assumptions, Luce (2010) showed that there must be an increasing function $g : X \rightarrow \mathbb{R}$ and a constant $\alpha > 0$ such that

$$U(x) = \left\{ \begin{array}{l} e^{\alpha g(x)} - 1 \\ \alpha g(x) \\ 1 - e^{-\alpha g(x)} \end{array} \right\} \Leftrightarrow \delta = \left\{ \begin{array}{l} 1 \\ 0 \\ -1 \end{array} \right\}. \tag{28}$$

⁶A. A. J. Marley reminded me of this reference. The 1993 edition is mainly an update of applications.

So the Arrow-Pratt measure (27) is easily calculated to be

$$r(x) = -\frac{g''(x)}{g'(x)} - \delta\alpha g'(x), \quad \delta = -1, 0, 1. \quad (29)$$

For money amounts, it seems most plausible that

$$x \oplus y \sim x + y, \quad (30)$$

from which it follows that $g(x) = x$. If so, then (29) reduces to

$$r(x) = -\delta\alpha, \quad \delta = -1, 0, 1, \quad (31)$$

which, of course, is constant risk aversion for $\delta = -1$ or constant risk seeking for $\delta = 1$ (Keeney and Raiffa (1976), explored (28) with $g(x) = x$ and without the ± 1 term).

Notice that

1. For $\delta = 1$ as $x \rightarrow -\infty$, then $U(x) = e^{\alpha x} - 1 \rightarrow -1$, and
2. For $\delta = -1$ as $x \rightarrow \infty$, then $U(x) = 1 - e^{-\alpha x} \rightarrow 1$.

The empirical issues of estimating these asymptotes and the parameter α have yet to be tackled systematically.

Given estimates of U , one acceptable way⁷ to average over people of the same type is for $\delta = 1$ to average $\ln(1 + U(x))$ and for $\delta = -1$ to average $-\ln(1 - U(x))$ vs. x .

3 General gambles

So far we have discussed only binary gambles, but clearly a satisfactory utility theory must deal with gambles having $n > 2$ branches as well. So far, this seems to have best been done via recursive forms.

Let a general gamble, where the consequences are ranked from best to worst, be denoted

$$g_{[n]} := (x_1, C_1; \dots; x_i, C_i; \dots; x_n, C_n). \quad (32)$$

Branching is the recursion in which the first two branches are combined into a single first-order binary gamble $g_{[2]} := (x_1, C_1; x_2, C_2)$, i.e.,

$$\begin{aligned} g_{[n]} &\sim (g_{[2]}, C_1 \cup C_2; x_3, C_3; \dots; x_n, C_n) \\ &= ((x_1, C_1; x_2, C_2), C_1 \cup C_2; x_3, C_3; \dots; x_n, C_n). \end{aligned} \quad (33)$$

⁷As noted earlier, it is well known that for linear functions, and only for them, the average function is also linear. Further, the slope of the average is the average of the individual slopes.

Upper gamble decomposition (UGD) is the recursion that treats as indifferent a binary gamble consisting of the branch with the best consequence and all of the remaining branches combined as a single first-order gamble, i.e.,

$$g_{[n]} \sim (x_1, C_1; g_{[n],-1}, \Omega \setminus C_1) \tag{34}$$

where $g_{[n],-1} := (x_2, C_2; \dots; x_i, C_i; \dots; x_n, C_n)$.

Some of the theoretical implications of the representations resulting from these properties are explored in Luce et al. (2008a, b) and Ng et al. (2009a) for the additive case $\delta = 0$ and in Ng et al. (2009b) for the cases $\delta \neq 0$. Idempotence was not assumed, but when idempotence is assumed the results described here follow readily. To my knowledge, no experiments have been run on either branching or UGD. There are many opportunities for research here, both empirical and theoretical. For example, if both branching and UGD are shown to be empirically incorrect, then theorists must seek alternative recursive properties.

4 Experimental program

For the experiments listed below in Section 4.1, type does not matter and so these studies need not be rerun because of our classification. Of course, that does not mean the existing studies are beyond criticism. For example, commutativity and associativity of \oplus , the Thomsen condition, event commutativity, and right autodistributivity are all formulated in terms \sim which invites having the respondent do matches, which is known to be a problematic procedure.

The several experiments discussed in the subsequent sections were run without any apparent awareness of the classification of people into 3 types. So, when that distinction is important to a property, it must be rerun taking risk type into account.

4.1 Tests independent of risk type

Of course, even when risk type does not directly matter, we must be most cautious about averaging the data from several individuals because there are usually parametric differences as, e.g., in the utility of functions.

- Transitivity of \succsim , (1), appears well supported empirically at this point and is not in urgent need of further evaluation.
- Commutativity, (3), and Associativity, (4), of JR have not been directly evaluated, but few doubt that they hold. Indeed, if money satisfies (30), then they must hold by virtue of properties of arithmetic.
- Monotonicity (or independence), (7), of JR was explored by Cho and Fisher (2000) and was mostly sustained.
- The Thomsen condition, (17), has not, to my knowledge, been directly evaluated using unitary gambles. Because practically every theory that has

been proposed implicitly or explicitly assumes it to be correct, it is almost certainly worthy of some empirical study.

- Consequence monotonicity, (16), of binary gambles seems first to have been explored empirically by von Winterfeldt et al. (1997) and reanalyzed by Ho et al. (2005), as described in Section 1.3.2. They concluded that the von Winterfeldt et al. (1997) analysis was not sufficiently thorough and that consequence monotonicity was actually very well sustained. Again, methods invoking \sim rather than \succsim have the usual matching difficulties (Birnbaum and Sutton 1992; Birnbaum 1992).
- Event commutativity, (13), and right autodistributivity, (14), were both discussed in Section 1.5, and there seems no need for further data collection.

4.2 Classification by risk type

Given the p-additive representation, (8), it is clearly critical to know the risk type of each experimental respondent and to analyze their data separately. To do this, one should find an event E with binary symmetry, (23), and determine type by the criterion (25).

- One obvious empirical question is whether or not a person is consistent in adhering to the criterion. At this time, we simply do not know—it seems to be an empirically virgin topic.
- A second thorny issue is whether the criterion is just for gains and losses separately or whether it also works for mixed gains and losses. This topic needs extensive exploration in order to guide future theory construction.

4.3 Linking for risk neutral people

The properties of segregation, (18), and duplex decomposition, (19), have proved important theoretically, and they have received a fair amount of empirical investigation (Sneddon and Luce 2001; Cho et al. 2002; Karabatsos 2005) with somewhat mixed and confusing results. However, at the time, no one recognized the risk type distinction of the p-additive form, and so the data were not so partitioned. This should be done.

- Those people exhibiting $\delta = 0$ should have, according to the theory (see Tables 1 of the Ng et al. (2009a, b) articles), a rank-dependent representation under segregation and a linear weighted one under duplex decomposition.
- Once that is done, the recursive properties of branching, (33), and UGD, (34), need direct empirical investigation.

4.4 Linking for risk averse or risk seeking people

Ng et al. (2009b) have studied the representations that follow from segregation, (18), and duplex decomposition, (19).

- For $\delta \neq 0$, segregation leads to the rank-dependent form, which according to some data (of course, not partitioned by type), was rejected because it implies coalescing (Luce 2000, pp. 92, 180), which seems empirically wrong, at least when not partitioned. Again, this needs to be restudied with respondent type playing a role.
- With $\delta \neq 0$, duplex decomposition implies that, in essence, the weights do not vary with events, which is absurd. So, DD should fail for these 2 types.
- Should the evidence, when partitioned by risk types and carefully checked for experimental flaws, show systematic failures of segregation and/or duplex decomposition, then theorists will be forced to devise alternatives and work out their consequences.

4.5 Inductive conditions of branching and UGD

To my knowledge, no empirical work has ever been attempted to check directly either branching, (33), or upper gamble decomposition, (34). Given their current theoretical importance, it is important to do so with respondents partitioned by type.

4.6 Fitting utility and weighting functions to data

Considerable work has been done on fitting functions for both utility and weighting functions to appropriate data (see Luce 2000, Sections 3.3 and 3.4 for a summary to 1999).

- Mostly this has been done under the implicit assumption of additive joint receipts ($\delta = 0$), and no attempt has been made to fit the form (28) which has been recently derived and its special case (30).
- To do so, we need to develop ways to estimate from finite sets of data the upper bound for $\delta = -1$ and lower bound for $\delta = 1$ types.
- Of the weighting functions, the generally most successful class of functions is the Prelec one for events with known probabilities

$$W(p) = \exp[-\beta(-\ln p)^\alpha],$$

which includes power functions as the special case $\alpha = 1$ (e.g., Sneddon and Luce 2001).⁸ Luce (2001) presents a fairly simple axiomatic condition that is equivalent to a Prelec function, which should be explored again but partitioned by type. Aczél and Luce (2007) generalized that to the case $W(1) \neq 1$.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

⁸The expression for the Prelec function in this article, Eq. 26, was incorrectly stated using $-p$ where it should have been p .

References

- Aczél, J., & Luce, R. D. (2007). Remark: A behavioral condition for Prelec's weighting function without restricting its value at 1. *Journal of Mathematical Psychology*, *51*, 126–129.
- Aczél, J., Luce, R. D., & Maksa, G. (1996). Solutions to three functional equations arising from different ways of measuring utility. *Journal of Mathematical Analysis and Applications*, *204*, 451–471.
- Arrow, K. J. (1965). *Aspects of the theory of risk bearing*. Helsinki: Academic Bookstores.
- Arrow, K. J. (1974). *Essays in the theory of risk-bearing*. Chicago: Markham.
- Birnbaum, M. H. (1992). Issues in utility measurement. *Organizational Behavior and Human Decision Processes*, *52*, 319–330.
- Birnbaum, M. H. (2008). Evaluation of the priority heuristic as a descriptive model of risky decision making: Comment on Brandstätter, Gigerenzer, and Hertwig (2006). *Psychological Review*, *115*, 253–262.
- Birnbaum, M. H., & Gutierrez, R. J. (2007). Testing for intransitivity of preferences predicted by a lexicographic semi-order. *Organizational Behavior and Human Decision Making*, *104*, 97–112.
- Birnbaum, M. H., & Sutton, S. E. (1992). Scale convergence and utility measurement. *Organizational Behavior and Human Decision Processes*, *52*, 183–215.
- Brothers, A. (1990). *An empirical investigation of some properties that are relevant to generalized expected-utility theory*. Unpublished doctoral dissertation, University of California, Irvine.
- Cho, Y.-H., & Fisher, G. R. (2000). Receiving two consequences: Test of monotonicity and scale invariance. *Organizational Behavior and Human Decision Processes*, *83*, 81.
- Cho, Y.-H., & Luce, R. D. (1995). Tests of assumptions about certainty equivalents and joint receipt of lotteries. *Organizational Behavior and Human Decision Processes*, *64*, 229–248.
- Cho, Y.-H., Luce, R. D., & Truong, L. (2002). Duplex decomposition and general segregation of lotteries of a gain and a loss. *Organizational Behavior and Human Decision Processes*, *89*, 1176–1193.
- Cho, Y.-H., Luce, R. D., & von Winterfeldt, D. (1994). Tests of assumptions about the joint receipt of gambles in rank- and sign-dependent utility theory. *Journal of Experimental Psychology: Human Perception and Performance*, *20*, 931–943.
- Cho, Y.-H., Truong, L., & Haneda, M. (2005). Testing the indifference between a binary lottery and its edited components using observed estimates of variability. *Organization Behavior and Human Decision Processes*, *97*, 82–89.
- Chung, N.-K., von Winterfeldt, D., & Luce, R. D. (1994). An experimental test of event commutativity in rank-dependent utility theory. *Psychological Science*, *5*, 394–400.
- Ho, M.-H., Regenwetter, M., Niederée, R., & Heyer, D. (2005). An alternative perspective on von Winterfeldt et al.'s (1997) test of consequence monotonicity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *31*, 365–373.
- Hölder, O. (1901). Die axiome der quantität und die Lehre vom Mass. *Ber. Verh. Kgl. Sächs. Ges. Wiss. Leipzig, Math.-Phys. Classe*, *53*, 1–64.
- Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *The American Economic Review*, *92*, 1644–1655.
- Iverson, G., & Falmagne, J.-C. (1985). Statistical issues in measurement. *Mathematical Social Science*, *10*, 131–153.
- Karabatsos, G. (2005). The exchangeable multinomial model as an approach to testing deterministic axioms of choice and measurement. *Journal of Mathematical Psychology*, *49*, 51–69.
- Keeney, R. L., & Raiffa, H. (1976). *Decisions with multiple objectives: Preferences and value trade-offs*. New York: Wiley. Reprinted with some revisions (1993), New York: Cambridge University Press.
- Krantz, D. H., Luce, R. D., Suppes, P., & Tversky, A. (1971). *Foundations of measurement* (Vol. I). New York: Academic Press. Reprinted by Dover Publications in 2007.
- Levy, H., & Levy, M. (2002). Arrow-Pratt risk aversion, risk premium and decision weights. *Journal of Risk and Uncertainty*, *25*, 265–290.
- Luce, R. D. (1996). When four distinct ways to measure utility are the same. *Journal of Mathematical Psychology*, *40*, 297–317.
- Luce, R. D. (2000). *Utility of gains and Losses*. Mahwah, NJ: Erlbaum.

- Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology*, 45, 167–179.
- Luce, R. D. (2003). Rationality in choice under certainty and uncertainty. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives in judgment and decision making* (pp. 64–83). Cambridge, England: Cambridge University Press.
- Luce, R. D. (2010). Interpersonal comparisons of utility for 2 of 3 types of people. *Theory and Decision*, 68, 5–24.
- Luce, R. D., Ng, C. T., Marley, A. A. J., & Aczél, J. (2008a). Utility of gambling I: Entropy-modified linear weighted utility. *Economic Theory*, 36, 1–33. (See erratum at Luce web site: Utility of gambling I: entropy-modified linear weighted utility and utility of gambling II: Risk, paradoxes, and data).
- Luce, R. D., Ng, C. T., & Marley, A. A. J., & Aczél, J. (2008b). Utility of gambling II: Risk, paradoxes, and data, economic. *Economic Theory*, 36, 165–187. (See erratum at Luce web site: Utility of gambling I: Entropy-modified linear weighted utility and utility of gambling II: Risk, paradoxes, and data).
- Luce, R. D., & Tukey, J. W. (1964). Simultaneous conjoint measurement: A new type of fundamental measurement. *Journal of Mathematical Psychology*, 1, 1–27.
- Marley, A. A. J. & Luce, R. D. (2005). Independence properties vis-à-vis several utility representations. *Theory and Decision*, 58, 77–143.
- Ng, C. T., Luce, R. D., & Marley, A. A. J. (2009a). Utility of gambling when events are valued: An application of inset entropy. *Theory and Decision*, 67, 23–63. See erratum at Luce web site: Utility of gambling I: entropy-modified linear weighted utility and utility of gambling II: risk, paradoxes, and data.
- Ng, C. T., Luce, R. D., & Marley, A. A. J. (2009b). Utility of gambling under p-additive joint receipt and segregation or duplex decomposition. *Journal of Mathematical Psychology*, 53, 273–286.
- Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica*, 12, 122–136.
- Regenwetter, M., Dana, J., & Davis-Stober, C. P. (2010). Transitivity of preferences. *Psychological Review* (in press).
- Regenwetter, M., & Davis-Stober, C. P. (2008). There are many models of transitive preference: A tutorial review and current perspective. In T. Kugler, J. C. Smith, T. Connolly, & Y. J. Son (Eds.), *Decision modeling and behavior in uncertain and complex environments*. New York: Springer.
- Ross, S. A. (1981). Some stronger measures of risk aversion in the small and the large with applications. *Econometrica*, 49, 621–638.
- Savage, L. J. (1954). *Foundations of statistics*. New York: Wiley.
- Schneider, S. L., & Lopes, L. (1986). Reflection in preferences under risk: Who and when may suggest why. *Journal of Experimental Psychology: Human Perception and Performance*, 12, 535–548.
- Sneddon, R., & Luce, R. D. (2001). Empirical comparisons of bilinear and non-bilinear utility theories. *Organizational Behavior and Human Decision Processes*, 84, 71–94.
- Steingrimsson, R. (2009). Evaluating a model of global psychophysical judgments for brightness: I. Behavioral properties of summations and productions. *Attention, Perception, & Psychophysics*, 71, 273–286.
- Steingrimsson, R. (2010). Evaluating a model of global psychophysical judgments for brightness: II: Behavioral properties linking summations and productions. *Attention, Perception, & Psychophysics* (in press).
- Steingrimsson, R., & Luce, R. D. (2005). Evaluating a model of global psychophysical judgments I: Behavioral properties of summations and productions. *Journal of Mathematical Psychology*, 49, 290–306.
- Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, 5, 297–323.
- von Winterfeldt, D., Chung, N.-K., Luce, R. D., & Cho, Y.-H. (1997). Tests of consequence monotonicity in decision making under uncertainty. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 23, 406–426.