Available online at www.sciencedirect.com
ScienceDirect
Organizational Behavior and Human Decision Processes 102 (2007) 154–173
ORGANIZATIONAL BEHAVIOR AND HUMAN DECISION PROCESSES
www.elsevier.com/locate/obhdp

Tests of branch splitting and branch-splitting independence in Allais paradoxes with positive and mixed consequences ☆

Michael H. Birnbaum *

Department of Psychology, California State University, CSUF H-830M, P.O. Box 6846, Fullerton, CA 92834-6846, USA

Received 13 May 2005
Available online 5 June 2006

Abstract

Four experiments with 1391 participants compared descriptive models of risky decision making. The first replicated and extended evidence refuting cumulative prospect theory (CPT) as an explanation of Allais paradoxes. The second and third experiments used a new design to unconfound tests of upper and lower coalescing, which allows tests of branch-splitting independence. Violations of upper coalescing were observed, contrary to an idempotent, lower gains decomposition utility (LGDU) model, but consistent with a transfer of attention exchange (TAX) model. Violations of restricted branch independence refute subjectively weighted utility (SWU), including "stripped" prospect theory. There were also small but significant violations of branch-splitting independence, contrary to SWU but consistent with TAX. The fourth experiment used a new design to test a model in which a gamble's utility is expected utility plus entropy. This model can be rejected because there were significant effects of whether the upper or lower branch is split with equal entropy. Comparing specific numerical models, TAX fit better than LGDU or SWU, which fit better than CPT. Discrepancies from TAX (as fit to previous data) were observed when the number of branches in the two gambles within a choice differed. In these cases, people tend to select the gamble with the greater number of branches leading to favorable consequences.

Keywords: Allais paradox; Prospect theory; Risky decision making; Utility

Introduction

Allais paradoxes constitute evidence against expected utility (EU) theory as a descriptive model of decision making. They were regarded as refutations of Savage's (1954) "sure thing" principle (Kahneman & Tversky, 1979; Wakker, 2001).

Wu and Gonzalez (1998) categorized "constant consequence" paradoxes of Allais (1979) into three types, and showed that cumulative prospect theory (CPT) of Tversky and Kahneman (1992), with its inverse-S shaped probability weighting function, can account for empirical results with all three. Success in describing the different types of Allais paradoxes was considered a strong argument for CPT (Wu, Zhang, & Gonzalez, 2004).

Birnbaum (2001) replicated the results of Wu and Gonzalez with modest but real consequences and noted that his transfer of attention exchange (TAX) model also describes all three types of Allais paradoxes. As noted by Birnbaum (1999a), however, the standard Allais paradoxes are confounded; only by separating these tests into tests of simpler properties is it possible to compare these rival descriptive theories of risky decision making.

This paper will compare models of Allais paradoxes. In addition to TAX and CPT, subjectively weighted utility (SWU) theory (Edwards, 1962) and lower gains

☆ Support was received from National Science Foundation Grants, SBR-9410572, SES 99-86436, and BCS-0129453. Thanks are due R. Duncan Luce and A.A.J. Marley for helpful suggestions on an earlier draft and for useful discussions of these topics; two anonymous reviewers also helped improve the exposition.
* Fax: +1 714 278 7134.
E-mail address: mbirnbaum@fullerton.edu.

0749-5978/$ - see front matter
doi:10.1016/j.obhdp.2006.04.004

decomposition utility (LGDU) (Marley & Luce, 2001) will be evaluated. In order to compare these models, it is necessary to disentangle variables that have been confounded in previous studies.

Decomposition of Allais paradoxes: Behavioral properties

Constant consequence paradoxes can be decomposed into three simpler properties: transitivity, coalescing, and restricted branch independence (Birnbaum, 1999a). If people satisfied these three properties, they would not show Allais paradoxes. These properties are defined as follows:

Transitivity, assumed in all of the models discussed here, holds that $A \succ B$ and $B \succ C \Rightarrow A \succ C$, where $\succ$ denotes the preference relation.

Coalescing is the assumption that if a gamble has two or more (probability-consequence) *branches* yielding identical consequences, those branches can be combined by adding their probabilities. For example, if $G = (\$100, .2; \$100, .2; \$0, .6)$, a gamble with two branches to win \$100 and otherwise win zero, then $G \sim G' = (\$100, .4; \$0, .6)$, where $\sim$ denotes indifference and G' is the *coalesced* form of G. Violations of coalescing combined with transitivity are termed *event-splitting effects* (Birnbaum, 1999a, 1999b; Humphrey, 1995; Starmer & Sugden, 1993). For example, if $G \succ A$ and $G' \prec A$, we say there is an event-splitting effect. Assuming transitivity, event-splitting effects (*branch-splitting* effects) are violations of coalescing.

Upper and lower coalescing are defined as follows. Let $G' = (x, p; y, 1 - p)$, where $x > y > 0$.

Upper coalescing assumes:

$$G' \sim G'' = (x, p - q; x, q; y, 1 - p)$$

Lower coalescing assumes:

$$G' \sim G''' = (x, p; y, q; y, 1 - p - q).$$

When the term "coalescing" is used without qualification, it refers to the assumption of all forms of coalescing. Experiments 2 and 3 will examine upper and lower coalescing separately.

Restricted Branch independence is the assumption that if two gambles have a common probability-consequence (or event-consequence) branch, one can change the common consequence without affecting the preference induced by the other components.

For three-branch gambles with $p + q + r = 1$, $p, q, r > 0$, restricted branch independence can be written as follows for all consequences $x' > x > y > y'$ and all consequences, $z' > z$, with no other restrictions on their order

$$S = (x, p; y, q; z, r) \succ R = (x', p; y', q; z, r) \iff S' = (x, p; y, q; z'; r) \succ R' = (x', p; y', q; z', r) \quad (1)$$

The term "restricted" refers to the constraint that the numbers of branches and probability distributions are the same in all four gambles of Expression 1. Restricted branch independence is weaker than Savage's (1954) "sure thing" axiom. When the rank orders of all consequences are also restricted to be the same in all four gambles, the property is called comonotonic restricted branch independence.

Analysis of Allais paradox

If choices satisfy transitivity, coalescing, and restricted branch independence then they would not show the constant consequence paradoxes of Allais. For example, consider the following series of choices:

A:	\$1M for sure	$\succ$	B:	.10 to win \$2M .89 to win \$1M .01 to win \$0
$\iff$ (*coalescing & transitivity*)				
A':	.10 to win \$1M .89 to win \$1M .01 to win \$1M	$\succ$	B:	.10 to win \$2M .89 to win \$1M .01 to win \$0
$\iff$ (*restricted branch independence*)				
A'':	.10 to win \$1M .89 to win \$0 .01 to win \$1M	$\succ$	B':	.10 to win \$2M .89 to win \$0 .01 to win \$0
$\iff$ (*coalescing & transitivity*)				
C:	.11 to win \$1M .89 to win \$0	$\succ$	D:	.10 to win \$2M .90 to win \$0

From the first to second step, A is converted to a split form, A';A' should be indifferent to A by coalescing, and by transitivity, A' should be preferred to B. From the second to third steps, the consequence on the common branch (.89 to win \$1M) has been changed to \$0 on both sides, so by restricted branch independence, A'' should be preferred to B'. By coalescing branches with the same consequences on both sides, we see that C should be preferred to D.

However, many people choose A over B and D over C; indeed, this behavior is the Allais paradox, so we know that at least one of these three assumptions must be false.

All of the models considered here satisfy transitivity. Different models, described in the next section, attribute the Allais paradoxes to failure of either branch independence or coalescing. By dissecting coalescing and branch independence, we can test among alternative descriptive models (Birnbaum, 1999a, 2004a). CPT implies coalescing and assumes that Allais paradoxes are the result of violations of restricted branch independence.

Branch-splitting independence

Three models attribute Allais paradoxes to violations of coalescing (TAX, LGDU, and SWU); however, these models make different predictions for a property defined by Birnbaum and Navarrete (1998) called *branch-splitting independence* (BSI), which can be defined as follows:

$$G = (y, p - r; y, r; z, 1 - p) \succ G' = (y, p; z, 1 - p) \iff H = (x, 1 - p; y, p - r; y, r) \succ H' = (x, 1 - p; y, p) \quad (2)$$

where $x > y > z > 0$. Conceptually, if splitting the upper branch of G' [i.e., (y,p)] into $(y,p - r;\ y,r)$ improves G then splitting the same branch in the same way should improve H, even though that branch is now the lower branch of H'. Branch-splitting independence (BSI) is the assumption that splitting the same event leading to the same consequence should always have the same effect no matter what the other branches are in the gamble. In other words, splitting the same branch into the same pieces should either improve or diminish the value of any gamble containing that branch.

Branch-splitting independence (BSI) will not be tested in the transparent form of Expression 2, however, but rather these four gambles will be compared indirectly by how they perform against other gambles. In particular, the property will be assessed from four choices as follows:

$$P(G \succ A) > P(G' \succ A) \iff P(H \succ B) > P(H' \succ B) \quad (3)$$

where $P(G' \succ A)$ is the probability of choosing G' over A; G, G', H, H' are as defined above.

If gamble F always has at least as high a probability of winning a given prize or higher as does gamble G, and if F sometimes has a higher probability than G of winning a prize that high, F is said to dominate G by *first order stochastic dominance*. To *satisfy stochastic dominance* means that one should not choose G over F, except by random error. Certain descriptive models satisfy stochastic dominance. Birnbaum (1997), however, constructed an example in which his models implied systematic violations.

These five behavioral properties are listed in Table 1. As shown in the next section, different models satisfy or violate these properties, so one can compare models by testing these properties empirically.

A sixth property, idempotence, also helps to categorize the models. *Idempotence* is the assumption that if a gamble always yields the same consequence, no matter what chance event occurs, that gamble is indifferent to the sure thing of receiving that same consequence. That is

$$G = (x, p_1; x, p_2; \ldots; x, p_n) \sim x, \quad (4)$$

where $\sum_{i=1}^{n} p_i = 1$.

Four descriptive models of decision making

Four models will be compared in the first three experiments: SWU, CPT, TAX, and idempotent, LGDU. All four can describe the original versions of Allais paradoxes, but they make different predictions in these studies. These models will be compared in two ways: first, properties implied by their general forms will be tested; second, predictions will be calculated using parametric assumptions. Whereas refutation of a model is best accomplished by showing systematic violations of that theory's implications (apart from any parametric assumptions), the positive case for a descriptive model can be assessed by evaluating its success in using parameters to predict new results.

A fifth model, which violates idempotence, has been suggested as a theory of the Allais paradoxes (Meginniss, 1976; Yang & Qiu, 2005) and will be tested in Experiment 4.

It is useful to state a generic configural weight model that includes all four of the models to be compared in Experiments 1–3 as special cases. For gambles on positive consequences, $G = (x_1, p_1; x_2, p_2, \ldots x_i, p_i; \ldots; x_n, p_n)$, where the consequences are ranked such that $x_1 > x_2 > \ldots > x_n \geqslant 0$, this generic model can be written:

Table 1
Comparison of predictions of models for five behavioral properties

Model	Behavioral properties				
	RBI	Upper coalescing	Lower coalescing	BSI	SD
SWU	Satisfies	Violates	Violates	Satisfies	Violates
CPT	Violates (*RS*′)	Satisfies	Satisfies	Moot	Satisfies
LGDU	Violates (*SR*′)	Satisfies	Violates (split worse)	Violates	Violates
TAX	Violates (*SR*′)	Violates (split better)	Violates (split worse)	Violates	Violates

Note: RBI, restricted branch independence; BSI, branch-splitting independence; SD, first order stochastic dominance. Expected utility (EU) theory satisfies all of the properties, except BSI, which is moot under EU. SWU, subjectively weighted utility; CPT, cumulative prospect theory. LGDU, idempotent, lower gains decomposition utility. TAX, transfer of attention exchange model. Subjectively weighted average utility is a special case of TAX that satisfies RBI. Predictions shown in parentheses are based on numerical models.

$$\text{RWU}(G) = \sum_{i=1}^{n} w_i(p_1, p_2, \ldots, p_n)u(x_i), \tag{5}$$

where the weight of a given consequence depends on the number of consequences in the partition displayed to the participant, the ranking of the consequences, and the probability distribution on those consequences.

A more general version of this model defined on uncertain events (with unspecified probabilities) is called the rank weighted utility model (Luce & Marley, 2005; Marley & Luce, 2005). The idempotent rank-weighted utility model has been shown to be equivalent to Birnbaum's (1999a) generic TAX model (Marley & Luce, 2005, proposition 6). RWU models satisfy comonotonic, restricted branch independence, but the special cases differ for other properties listed in Table 1. Luce and Marley (2005) deduce relations among special cases.

Subjectively weighted utility (SWU): Stripped prospect theory

The SWU model can be written:

$$\text{SWU}(G) = \sum_{i=1}^{n} w(p_i)u(x_i), \tag{6}$$

where $w(p_i)$ and $u(x_i)$ are the weight of the probability and utility of the consequence of the ith branch, respectively. This formulation was used by Edwards (1962, 1954) and later by Kahneman and Tversky (1979), who restricted it to gambles with no more than two nonzero consequences. Eq. (6) has been called "stripped prospect theory" by Starmer and Sugden (1993), because the editing rules in original prospect theory (OPT) and restriction to no more than two nonzero consequences have been removed. This model violates both upper and lower coalescing and violates stochastic dominance, but it satisfies restricted branch independence and branch-splitting independence.

Expected utility (EU) is a special case of SWU in which $w(p_i) = p_i$. EU satisfies all of the properties in Table 1, except for branch-splitting independence, which is moot, since EU satisfies coalescing. SWU violates idempotence. The model of Lattimore, Baker, and Witte (1992) is a special case of SWU with the following probability weighting function, $w(p_i) = \delta p_i^{\gamma}/[\delta p_i^{\gamma} + (1-p_i)^{\gamma}]$. When $0 < \delta, \gamma < 1$, weight is an inverse-S function of probability.

When Eq. (6) is restricted so that $\sum_{i=1}^{n} w(p_i) = 1$, the model is termed a subjectively weighted average utility (SWAU) model (Birnbaum, 1999a). SWAU satisfies idempotence but violates coalescing when $w(p) \neq \text{p}$. Viscusi's (1989) prospective reference theory (PRT) is a special case of SWAU with $w(p_i) = \omega p_i + (1-\omega)/n$. Neither SWU nor SWAU can account for violations of restricted branch independence, because weight is independent of rank in these models. SWU satisfies branch-splitting independence, but SWAU violates it.

Cumulative prospect theory (CPT)

Let $G = (x_1,p_1; x_2,p_2; \ldots; x_n,p_n)$ refer to a gamble with n distinct branches, where the consequences are ordered such that $x_1 \geqslant x_2 \geqslant \ldots \geqslant x_n \geqslant 0$, where $\sum p_i = 1$ CPT can be written as follows:

$$\text{CPU}(G) = \sum_{i=1}^{n} [W(\sum_{j=1}^{i} p_j) - W(\sum_{j=1}^{i-1} p_j)]u(x_i), \tag{7}$$

where $W(P)$ is a monotonic function transforming decumulative weight to decumulative probability $[P_i = \sum_{j=1}^{i} p_j$ is the (decumulative) probability that a prize is greater than or equal to $x_i]$; $W(0) = 0$ and $W(1) = 1$. When restricted to strictly positive consequences, CPT has the same representation as Quiggin's (1985, 1993) rank-dependent utility (RDU). This model violates restricted branch independence, except when $W(P) = P$, in which case it reduces to EU. However, it must satisfy both upper and lower coalescing and first order stochastic dominance no matter what functions are chosen for $W(P)$ and $u(x)$. Because RDU and CPT satisfy coalescing, branch-splitting independence is moot.

In practice, the decumulative weighting function (Tversky & Wakker, 1995) has been fit with the equation:

$$W(P_i) = \frac{\delta P_i^{\gamma}}{\delta P_i^{\gamma} + (1-P_i)^{\gamma}} \tag{8}$$

where P is decumulative probaility. Tversky and Kahneman (1992) assumed that $u(x) = x^{\beta}$. For the Tversky and Kahneman (1992) data, the parameters $\beta = 0.88$; $\gamma = 0.61$; and $\delta = 0.72$ provided the best fit. With these parameters, CPT implies a particular pattern of violation of restricted branch independence. With $(p,q,r) = (p, p, 1-2p)$; $p \leqslant 1/3$ in Expression 1, it should be possible to select consequences such that $z' > x' > x > y > y' > z > 0$ so that $R \succ S$ and $R' \prec S'$. This pattern is denoted RS' in Table 1. Indeed, this implication follows if the decumulative weighting function is any weakly inverse-S function of decumulative probability (Birnbaum, 2005).

Transfer of attention exchange (TAX) model

A "special" case of the TAX model (Birnbaum & Chavez, 1997) can be written as follows:

$$\text{TAX}(G) = \frac{\sum_{i=1}^{n} u(x_i)[t(p_i) - \frac{\delta}{(n+1)} \sum_{j=i+1}^{n} t(p_i) + \frac{\delta}{(n+1)} \sum_{j=1}^{i-1} t(p_j)]}{\sum_{i=1}^{n} t(p_i)}, \tag{9}$$

where TAX(G) is the utility of the gamble; $t(p)$ is a function of probability; $u(x)$ is the utility function of money, and δ is the single configural parameter; $\delta/(n+1)$ is the proportion of weight taken from a branch with a higher consequence and transferred to each branch with a lower consequence. Note that the amount of weight transferred is proportional to the (transformed) probability of the branch giving up weight.

When $\delta = 0$ the special TAX model reduces to a type of SWAU. PRT is a special case of Eq. (9) in which $\delta = 0$ and $t(p) = \gamma p + (1-\gamma)/n$. When both $\delta = 0$ *and* $t(p) = p$, TAX reduces to EU. The model has been written here with the same ranking conventions as in CPT; consequently, $\delta > 0$ in Eq. (9) corresponds to $\delta < 0$ in Birnbaum and Chavez (1997) and other papers using the earlier conventions.

The special TAX model satisfies idempotence and comonotonic restricted branch independence, but it violates restricted branch independence, upper and lower coalescing, branch-splitting independence, and stochastic dominance. However, this model only rarely violates these properties (Birnbaum, 2004a); therefore, one needs to know the model's parameters in order to design experiments with a reasonable expectation of detecting its predicted violations.

The TAX model was fit to a portion of the Tversky and Kahneman (1992) data. The following approximations were used: $t(p) = p^{\gamma}$, and $u(x) = x^{\beta}$ where $\beta = 1$; $\gamma = 0.7$; and $\delta = 1$. Based on these assumptions, the model violates restricted branch independence with the pattern, $S \succ R$ and $S' \prec R'$, denoted the SR' pattern in Table 1. It also predicts violations of stochastic dominance in the special choices constructed from Birnbaum's (1997) recipe and used here.

Idempotent, lower gains decomposition utility (LGDU)

Marley and Luce (2001) axiomatized the idempotent, LGDU model. The basic idea is that a complex gamble can be decomposed into a series of binary gambles to win the lowest gain or the gamble among higher gains. The binary gambles are represented with the following idempotent form:

$$\mathrm{LGDU}(x,p;y) = W(p)u(x) + [1 - W(p)]u(y) \tag{10}$$

where $x > y > 0$. To calculate values of three-branch gambles, the gains decomposition rule (Luce, 2000, pp. 200–202) is applied as follows:

$$LGDU(G) = W(p+q)LGDU(x, p/(p+q); y) + [1 - W(p+q)]u(z) \tag{11}$$

Expression (11) can be applied iteratively to make predictions for multi-branch gambles, without requiring any new parameters beyond those required for binary gambles.

Eq. (10) (or any idempotent binary model) implies that $\mathrm{GDU}(x,p;x,1-p) = u(x)$; i.e., that $F = (x,p;x,1-p) \sim x$, so this model implies upper coalescing, $G = (x,p;y,q;z,1-p-q) \sim (F,p+q;z,1-p-q) \sim (x, p+q;z,1-p-q)$; however, it need not satisfy lower coalescing, except in the special case where LGDU reduces to RDU (i.e., to CPT).

This model will be fit with $u(x) = x^{\beta}$, and with the weighting function used by Prelec (1998) and Luce (2000):

$$W(p) = \exp[-\gamma(-\ln p)^{\delta}]. \tag{12}$$

These functions match those used by Luce (2000, Chapter 5, p. 200–202) in his illustration of a "less restrictive" model. [There was a rounding error in the calculations in Luce (2000), which appeared to show that this LGDU model can account for Wu's (1994) violations of upper tail independence. However, because upper coalescing follows from lower gains decomposition and idempotence, and because Wu's upper tail independence follows from upper coalescing, this LGDU model must satisfy upper tail independence.] Luce, Ng, Marley, and Aczél (2006) derive a still more general form of gains decomposition utility (GDU) that need not satisfy idempotence and therefore need not satisfy upper coalescing. This more recent version of GDU is not tested here, except for a special case tested in Experiment 4.

Fitting choice proportions

To compute the fit of parametric models to observed choice proportions, choice probabilities are fit as a function of differences in utilities of the gambles, divided by the standard deviation of the difference:

$$P(A,B) = \frac{1}{1 + \exp\left[\frac{U(A)-U(B)}{\sigma_{AB}}\right]}, \tag{13}$$

where $P(A,B)$ is the probability of choosing gamble A over B; $U(A)$ and $U(B)$ are the utilities of A and B (calculated from one of the four utility models), and σ_{AB} is proportional to the standard deviation of the difference in utilities between the gambles. This model is similar to Thurstone's (1927) law of comparative judgment, except it uses the logistic instead of the cumulative normal transformation (Guilford, 1954, p. 144; Luce, 1959, 1994; Luce & Supes, 1965). For simplicity, it is often assumed that the standard deviations are constant, in which case $\alpha = -1/\sigma_{AB}$ is the only new parameter needed.

Busemeyer and Townsend (1993) argue that standard deviations in Eq. (13) should not be assumed to be constant because they might depend on the variance of the difference in the gambles' consequences. If the standard deviations in Eq. (13) are affected by branch splitting, the indirect test of branch-splitting independence in Eq. (3) might show violations produced by those changes rather than by changes in the utilities. Another potential limitation of Eq. (13) is that it assumes an underlying dimension; so it must satisfy weak stochastic transitivity. For these reasons, another approach to representing "error" in the data (described later) will also be used in certain analyses below (Birnbaum, 2004b; Carbone & Hey, 2000).

Each of the four models has four free parameters: α, β, γ, and δ. The SWU model is fit with the same two-parameter probability weighting function as in CPT. The same symbols are used here for analogous parameters in SWU, CPT, TAX, and LGDU models. But keep in mind that the parameters have different meanings (and different values) in different models. LGDU (with $\beta = 1$, $\gamma = .38$ and $\delta = 1.34$), prior TAX (with $\beta = 1$, $\gamma = .7$ and $\delta = 1$), prior CPT (with $\beta = 0.88$, $\gamma = .61$, and $\delta = .72$), and prior SWU (with $\beta = 0.88$, $\gamma = .61$, and $\delta = .72$) make similar predictions for binary gambles of the form, $F = (x,p;0,1-p)$, where $0 < x < \$150$.

Testing the models

Birnbaum's (2004a) analysis of the Allais paradox refuted both versions of prospect theory, with or without the editing principles of combination and coalescing. Because of the popularity of prospect theories (Starmer, 2000; Wakker, 2001; Wu et al., 2004), however, it seems reasonable that every effort should be made to either salvage them or to convincingly refute them. The goal is to identify where descriptive theories have flaws and to describe the underlying psychology that is contrary to these models.

One defense of CPT is that Birnbaum's (1999b, 2001, 2004a, 2004b) studies involved small, real chances to win modest prizes (e.g., \$100), whereas the studies of Kahneman and Tversky (1979), for example, dealt with hypothetical choices among gambles to win fairly large prizes (comparable to a month's salary). Perhaps CPT remains viable for hypothetical decisions involving large consequences.

There is a second set of reasons that large prizes are interesting. Birnbaum's (1999a) TAX model correctly predicted the modal choices in Birnbaum's (2004a) experiment without estimating any new parameters from the data. Birnbaum's "prior TAX model" (which uses parameters estimated from previous data) used a linear utility function for positive consequences ranging up to \$150. Would this same model and utility function work for gambles on consequences in the millions?

The results of the first experiment largely replicate previous findings; in particular, the first experiment will show strong violations of coalescing that agree with the Allais paradoxes and which contradict CPT. The second and third experiments test among three models that imply violations of coalescing, TAX, LGDU, and SWU ("stripped" prospect theory).

According to the TAX model, with parameters estimated from previous data, splitting the branch leading to the highest consequence within a gamble should increase the utility of a gamble, but splitting the branch leading to the lowest consequence of a gamble should decrease the utility of the gamble. According to SWU, as fit to previous data, splitting a branch leading to a positive consequence should always improve a gamble. According to any version of LGDU, splitting the upper branch should have no effect.

The purpose of the second and third experiments is therefore to decompose the Allais paradoxes into still finer steps than has yet been done, in order to provide separate and independent tests of upper and lower coalescing. Thus, these experiments allow tests of *branch-splitting independence* (BSI), Expression 2, which has apparently been examined in only one previous study (Martin, 1998) that had inconclusive results (see Birnbaum & Martin, 2003). Table 1 indicates that SWU satisfies BSI, that TAX and LGDU violate it, and that BSI is moot in CPT and EU, since they satisfy coalescing.

In particular, the TAX model with its prior parameters, plus the assumption Eq. (13) with equal standard deviations implies the following violation of BSI:

$$P[(\$50,.1;\$50,.05;\$7,.85) \succ (\$100,.1;\$7,.9)] > P[(\$50,.15;\$7,.85) \succ (\$100,.1;\$7,.9)]$$

and yet

$$P[(\$100,.85;\$50,.1;\$50,.05) \succ (\$100,.95;\$7,.05)] < P[(\$100,.85;\$50,.15) \succ (\$100,.95;\$7,.05)]$$

LGDU satisfies upper coalescing but violates lower coalescing. It implies that the first two probabilities (involving splitting of the upper branch) should be equal, but the second two need not be equal.

Experiment 4 tests a specific hypothesis related to branch splitting. It uses a special design to test a model that has been proposed by Yang and Qiu (2005) as a theory of the Allais paradox.

Experiment 1

Intuitively, it seems that the difference in happiness at winning \$1 Million rather than \$1 is psychologically greater than the difference between winning \$2 Million and \$1 Million. Therefore, the linear utility function used in the prior TAX model for small prizes seems unlikely to extrapolate to such large prizes.

Furthermore, it seems plausible that very large prizes might induce additional psychological considerations or processes, such as satisfying an aspiration level (as in Lopes & Oden, 1999; Payne, 2005). For example, $1 Million might allow one to retire from that 9 to 5 job, whereas prizes from $1 to $100 are not life-changing. And if TAX fails here, perhaps another model, like CPT or the Lopes and Oden (1999) model might work better.

The first experiment therefore investigates Allais paradoxes with large hypothetical prizes to check whether models that worked for small consequences can be extended to account for hypothetical choices involving very large prizes.

Experiment 1 also employed a second variation from previous work. Evidence of violation of coalescing is what most strongly contradicts CPT. Therefore, a new type of event-framing was applied in an attempt to facilitate coalescing. Consider the format and event framing in the choices below. The decision-maker selects an urn from which a marble will be blindly and randomly drawn, and the color of marble determines the prize. Each urn contains exactly 100 marbles in the following choices:

A:	10 blue marbles to win $1M 01 blue marble to win $1M 89 white marbles to win $2	*B*:	10 red marbles to win $2M 01 white marbles to win $2 89 white marbles to win $2
C:	11 blue marbles to win $1M 89 white marbles to win $2	*D*:	10 red marbles to win $2M 90 white marbles to win $2

With this framing, it should be easy to see that in *A*, 10 blue marbles plus 1 blue marble is the same as 11 blue marbles to win $1M in *C*. Similarly, it should be easy to see that 89 plus 1 white marbles to win $2 in *B* is the same as 90 white marbles to win $2 in *D*. So, if people tend to combine branches with the same consequences within gambles (Kahneman & Tversky, 1979; Kahneman, 2003), this method of event framing (i.e., labeling the events with identical marble colors for the same consequences) ought to facilitate it. If so, this *event framing* should make coalescing more likely to be satisfied.

In addition, if people tend to cancel common branches (Kahneman & Tversky, 1979), it should be easy to see that *A* and *B* both have a common branch of 89 white marbles to win $2. So if people use cancellation, this framing should facilitate it, and they should satisfy restricted branch independence.

Method of Experiment 1

The method and procedures were the same as in Birnbaum (2004a) with five exceptions, all designed to "help" CPT. First, consequences were large and hypothetical; second, events were framed within gambles as well as within choices, to promote coalescing. Third, the majority of choices used event-framing (16 of 20 decisions), to allow people to learn to use this feature. Fourth, event-framing was consistent in its mapping of colors to magnitude of consequences (red was used for very large consequences, blue and green for medium consequences, and white for small consequences). Fifth, each participant completed the task twice, allowing participants to learn to use the editing rules, and allowing the experimenter to fit models involving random "errors".

The design also included tests of stochastic dominance in split and coalesced forms. Complete materials can be viewed at the following URL: http://psych.fullerton.edu/mbirnbaum/decisions/Allais_big_fu.htm.

Participants viewed the instructions and choices via the WWW. They clicked a button beside the gamble that they would prefer in each choice. Instructions read (in part) as follows: "Although you can't win or lose money in this study, you should imagine that these decisions are for real, and make the same decisions that you would if very large, real stakes were involved."

There were 200 undergraduates who completed Experiment 1 twice, separated by three intervening judgment tasks that required about 15 min. Of these, 68% were female and 93% were 22 years of age or younger.

Within each replication, there were two series of tests, each of which included two tests of coalescing, a test of restricted branch independence, and all three types of Allais paradoxes. For example, in series A (Table 2), choices 6 and 9 differ with respect to coalescing/splitting only. Choices 9 and 16 differ only in the value on the common branch (branch independence). Choices 16 and 19 form another test of coalescing. Choices 6 and 12 are a Type 1 Allais paradox, choices 12 and 19 are a Type 2, and choices 6 and 19 are a Type 3 paradox (Birnbaum, 2001; Wu & Gonzalez, 1998).

The arrangement in series B (Table 3) is a variation of the same design, except counterbalanced for positions of the safe and risky gambles and using different probabilities and consequences. Table 4 shows tests of stochastic dominance and coalescing. In choice 5, the gamble shown on the left ($G+$) dominates the gamble on the right ($G-$). Choice 11 is the same as choice 5, except in choice 11 branches have been split so that corresponding branches have the same probabilities.

Results of Experiment 1

The percentages of participants who chose the "risky" gamble, *R*, in each choice are shown in Tables 2 and 3. According to EU, the decision-maker should choose either *R* in all rows or choose *S* in all rows, because EU satisfies both coalescing and branch independence. For example, choices 6 and 9 in Table 2 differ

Table 2
Dissection of Allais paradox with large consequences (series A)

No.	Type	Choice		Choice %		Prior TAX		Prior CPT	
		Second gamble, *S*	First gamble, *R*	Rep 1	Rep 2	*S*	*R*	*S*	*R*
6		11 blues to win $1,000,000 89 whites to win $2	10 reds to win $2,000,000 90 whites to win $2	76	77	125 K	**236 K**	132 K	**248 K**
9	Split #6	10 blues to win $1,000,000 01 blues to win $1,000,000 89 whites to win $2	10 reds to win $2,000,000 01 whites to win $2 89 whites to win $2	35	39	155 K	172 K	132 K	248 K
12	RBI #9	100 blues to win $1,000,000	10 reds to win $2,000,000 89 blues to win $1,000,000 01 white marble to win $2	40	44	**1000 K**	810 K	1000 K	1065 K
16	RBI #9, 12	89 reds to win $2,000,000 10 blues to win $1,000,000 01 blues to win $1,000,000	89 reds to win $2,000,000 10 reds to win $2,000,000 01 whites to win $2	43	45	1400 K	1449 K	1714 K	1825 K
19	Coal #16	89 reds to win $2,000,000 11 blues to win $1,000,000	99 reds to win $2,000,000 01 whites to win $2	35	38	**1541 K**	1282 K	1714 K	1825 K

Each entry is the percentage choosing *R*. The right most columns show predicted certainty equivalents of the gambles from prior TAX and prior CPT, based on parameters estimated from data of Tversky and Kahneman (1992). Bold font indicates correct prediction. The second column (Type) shows the relation between each choice and a preceding one in the table. For example, choice 9 is a split version of choice 6, choice 16 differs from choice 9 in that the consequence on the common branch is $2,000,000 instead of $2 (RBI, restricted branch independence); choice 19 is a coalesced version of choice 16.

Table 3
Dissection of Allais paradox into branch independence and coalescing (series B)

No.	Type	Choice		Choice % *R*		Prior TAX model		Prior CPT model	
		First gamble, *S*	Second gamble, *R*	Rep 1	Rep 2	*S*	*R*	*S*	*R*
10		15 blues to win $500,000 85 whites to win $11	10 reds to win $1,000,000 90 whites to win $11	73	67	76 K	**118 K**	81 K	**124 K**
17	Split #10	10 blues to win $500,000 05 blues to win $500,000 85 whites to win $11	10 reds to win $1,000,000 05 whites to win $11 85 whites to win $11	24	30	**100 K**	82 K	81 K	124 K
20	RBI #17	10 blues to win $500,000 85 blues to win $500,000 05 blues to win $500,000	10 reds to win $1,000,000 85 blues to win $500,000 05 whites to win $11	28	31	**500 K**	378 K	**500 K**	470 K
14	RBI #17, 20	85 reds to win $1,000,000 10 blues to win $500,000 05 blues to win $500,000	85 reds to win $1,000,000 10 reds to win $1,000,000 05 whites to win $11	40	47	**684 K**	674 K	**834 K**	791 K
8	Coalesce #14	85 reds to win $1,000,000 15 blues to win $500,000	95 reds to win $1,000,000 05 whites to win $11	13	14	**757 K**	591 K	**834 K**	791 K

The right most columns show predicted certainty equivalents of the gambles from prior TAX and prior CPT. Bold font indicates correct prediction of modal choice.

only with respect to coalescing. The branch of 11 blue marbles to win $1 Million in *S* of choice 6 has been split into 10 blue marbles to win $1 Million and 1 blue marble to win $1 Million in choice 9. On the right, the 90 white marbles to win $2 in choice 6 has also been split in choice 9. Contrary to EU, there are large changes from row to row in both tables.

Tests of coalescing

The percentages who chose the "risky" (*R*) gamble (with 10 red marbles to win $2M and 90 whites to win $2) rather than *S* (the "safe" gamble with 11 blue marbles to win $1M and 89 whites to win $2) were 76% and 77% in the first and second repetitions of choice 6. However, when the upper branch

Table 4
Violations of stochastic dominance and coalescing

No.	Choice		% violations	
	G+	*G*−	Rep 1	Rep 2
5	90 reds to win $960,000 05 blues to win $140,000 05 whites to win $120,000	85 reds to win $960,000 05 blues to win $900,000 10 whites to win $120,000	79	76
11	85 reds to win $960,000 05 blues to win $960,000 05 greens to win $140,000 05 whites to win $120,000	85 reds to win $960,000 05 blues to win $900,000 05 greens to win $120,000 05 whites to win $120,000	8	12
15	90 reds to win $960,000 05 yellows to win $140,000 05 pinks to win $120,000	85 blacks to win $960,000 05 blues to win $900,000 10 whites to win $120,000	77	71
7	94 blacks to win $1,100,000 03 yellows to win $80,000 03 purples to win $60,000	91 reds to win $1,100,000 03 blues to win $1,000,000 06 whites to win $60,000	74	69
13	91 blacks to win $1,100,000 03 pinks to win $1,100,000 03 yellows to win $80,000 03 purples to win $60,000	91 reds to win $1,100,000 03 blues to win $1,000,000 03 greens to win $60,000 03 whites to win $60,000	15	12
18	94 reds to win $1,100,000 03 blues to win $80,000 03 whites to win $60,000	91 reds to win $1,100,000 03 blues to win $1,000,000 06 whites to win $60,000	63	70

In choices 5, 11, and 15, G+ was first, and in choices 7, 13, and 18 it was second. Choices 5, 11, and 18 are framed.

of *S* and the lower branch of *R* are split in choice 9, only 35% and 39% chose the *R* gamble. If people "spontaneously transformed" (Kahneman, 2003) choice 9 by coalescing (to choice 6), they should make the same decisions in choices 9 and 6, except for error.

Instead, 93 of 200 participants switched preference from *R* to *S* in choices 6 and 9, compared with only 11 who made the opposite switch in the first replicate. In the second replicate, the numbers are 90 versus 14. Both splits are statistically significant by the binomial test, $z = 8.04$ and 7.45. They are also statistically significant by the following, more conservative test: Significantly more than half chose *R* in choice 6 ($z = 7.50$ and 7.78), and significantly more than half chose *S* in choice 9 ($z = 4.10$ and 3.05). [For $n = 200$, percentages outside the interval from 43% to 57% are significantly different from 50% by a two-tailed test with $\alpha = .05$ level of significance.]

In Table 3, splitting again produced significant reversals of the majority choice. In choice 10, 73% and 67% chose *R*, whereas only 24% and 30% chose *R* in choice 17. These differ significantly from 50% in opposite directions. These results violate CPT, EU, and the editing rule of combination, which imply coalescing.

Stochastic dominance and coalescing

According to any version of CPT (with any functions and parameters), people should satisfy coalescing and stochastic dominance (Birnbaum & Navarrete, 1998). Instead, Table 4 shows that from 63% to 79% violate stochastic dominance (significantly more than 50% in all 8 tests) in the coalesced form. However, only 8% to 15% (significantly less than 50% in all 4 tests) violate it when gambles are appropriately split so that equal probabilities appear on corresponding branches (choices 11 and 13). The prior TAX model correctly predicted all majority preferences in Table 4, including the reversal of the modal choices produced by branch splitting. No version of CPT can account for majority violations of stochastic dominance in choices 5, 7, 15, and 18, nor can it explain preference reversals produced by splitting.

There were 8 tests of stochastic dominance in coalesced form, counting four choices over two replications. Of the 200 participants, there were 47 who violated stochastic dominance on all 8 tests, and only 5 who satisfied it on all 8 tests. The average percentage of violations on a single test was 72.4%. The effects of having marble colors that matched (choices 5 and 18) or did not match (choices 7 and 15) on corresponding branches seems to have had little effect, if any, on these tests of stochastic dominance (72.0% versus 72.9%, respectively).

Analysis with "true and error" model

The use of replications permits testing of a true and error model to the data (Birnbaum, 2004b). This error model differs from that in Eq. (13).

Let $a =$ the probability that a person "truly" satisfies dominance, and $e =$ the probability of making an "error" in reporting a preference on one such test. The probability that a person would show a particular sequence of m violations in n identical tests, is as follows:

$$P(m,n) = ae^{m}(1-e)^{n-m} + (1-a)e^{n-m}(1-e)^{m}. \quad (14)$$

This two parameter model was fit to the 16 observed frequencies for choices 5 and 15 with two repetitions, solving for a and e, to minimize the χ^2 between observed and fitted (predicted) frequencies of each choice combination. The model indicates $a = 0.16$ and $e = 0.16$, with $\chi^2(13) = 18.4$; thus, 84% of participants "truly" violated stochastic dominance on these choices, and participants made "errors" 16% of the time when expressing their preferences.

To illustrate the connection between this error model and another way to conceptualize "error," participants were divided into two groups according to their internal consistency of responding. Overall, the average agreement between choices in the first and second replicates was 76.3%, similar to previous reported levels of agreement (Birnbaum, 1999b). There were 87 participants who agreed with their own decisions (between replicates) fewer than 15 times out of 20 (75%), and there were 113 who made identical choices on 15 or more decisions. The observed rates of violation of stochastic dominance on choices 5 and 15 were 0.65 and 0.84 in the two groups (both significantly greater than 0.5), where the more consistent group showed the higher rate of observed violation.

When the true and error model (Eq. (14)) is applied to data from the 87 less consistent participants, estimates were $1 - a = 0.85$ and $e = 0.30$. In a separate analysis of the 113 with higher self-consistency, $1 - a = 0.86$ and $e = 0.09$. The estimated rates of "true" violation are nearly equal in the two groups (85% and 86%), but their inferred error rates are (quite) different (30% and 9%). These findings further strengthen the case that these observed violations of CPT (and EU) are not attributable to random error or carelessness, but rather to systematic intention. Table 5 shows the observed and "predicted" (fitted) frequencies for this analysis. More detailed analyses and raw data for these studies are available from the following URL: http://psych.fullerton.edu/mbirnbaum/archive.htm.

Table 5
Numbers of participants showing each choice combination in tests of stochastic dominance on choices 5 and 15

Choice pattern	Inconsistent group ($n = 87$)		Consistent group ($n = 113$)	
	Observed	Fitted (87)	Observed	Fitted (113)
SSSS	3	3.82	6	11.14
SSSV	3	2.76	0	1.12
SSVS	2	2.76	2	1.12
SSVV	3	3.83	1	0.72
SVSS	3	2.76	2	1.12
SVSV	1	3.83	0	0.72
SVVS	3	3.83	0	0.72
SVVV	9	7.83	5	6.45
VSSS	5	2.76	3	1.12
VSSV	3	3.83	0	0.72
VSVS	5	3.83	1	0.72
VSVV	3	7.83	6	6.45
VVSS	5	3.83	0	0.72
VVSV	10	7.83	4	6.45
VVVS	10	7.83	8	6.45
VVVV	19	17.84	75	67.27

Participants were separated into groups according to their internal consistency.
S, satisfied; V, violated stochastic dominance on choices 5, 15, in replicates 1 and 2, respectively. Participants were separated according to their internal consistency over 20 choices. Inconsistent participants agreed in fewer than 15 of 20 decisions, consistent participants agreed in 15 or more of 20 decisions. The data were fit to the "true and error" model (Eq. (14)) to minimize the χ^2 between observed and fitted (best-fit, predicted) frequencies.

Prior CPT and TAX models

According to the prior CPT model (as fit by Tversky & Kahneman, 1992), the majority should have chosen the "risky" gamble in all five rows of Table 2. This CPT model therefore makes a correct prediction in only one row of Table 2, even failing to predict the basic Allais paradox in the reversal between choices 6 and 12. Prior TAX model with its linear utility function correctly predicted three choices, but it failed to predict the reversal between choices 6 and 9, and it predicted a reversal between 9 and 16 that did not materialize.

Although no version of CPT can account for the observed reversal between choices 10 and 17 (splitting) in Table 3, prior CPT correctly predicted the type 1 paradox (choices 10 and 20) and the type 3 paradox (choices 10 and 8) in Table 3. Prior TAX correctly predicted all five modal choices in Table 3.

Parametric model fitting

Table 6 shows results of estimating parameters to minimize the sum of squared differences between observed and predicted choice proportions using Eq. (13). Both LGDU and TAX do a much better job of predicting empirical choice proportions than either SWU or CPT with the same number of free parameters. This should not be a surprise, because no version of CPT can predict violations of stochastic dominance or coalescing, and no version of SWU can violate restricted branch independence. Whereas the prior TAX model with its linear utility function failed to predict two choices (#9 and 16), by estimating one new parameter from the data (the exponent of the utility function), TAX(2)

Table 6
Fit of models to the group choice proportions (Experiment 1)

Model	Estimated parameters				SS deviations	Failures to predict majority choices
	α	β	γ	δ		
Prior CPT (1)	3.99 E−05	(0.88)	(0.61)	(0.72)	0.761	9 (#9, 16, 17, 5, 15, 7, 18, 12, 19)
SWU(2)	0.12	0.27	(0.61)	(0.72)	0.567	4 (#12, 20, 19, 10)
CPT (4)	0.20	0.23	0.47	0.93	0.508	6 (#9, 17, 5, 15, 7, 18)
Prior LGDU (1)	5.23 E−06	(1)	(1.382)	(0.542)	0.475	4 (#9, 16, 14, 19)
Prior TAX (1)	5.41 E−06	(1)	(.7)	(1)	0.241	2 (#9, 16)
TAX(2)	7.26 E−05	0.802	(.7)	(1)	0.236	0
SWU(4)	0.039	0.39	0.57	1.37	0.209	2 (#12, 20)
TAX(3)	6.91 E−06	0.961	0.295	(1)	0.185	0
LGDU (4)	2.74 E−03	0.574	0.917	0.315	0.139	0
TAX (3)	9.70 E−03	0.498	0.347	(0)	0.098	0

Note: Parameter values in parentheses were fixed. Failures to predict majority choices correspond to choice numbers in Tables 2–4.

perfectly reproduces all of the majority choices in Tables 2–4. Consistent with intuition, the estimated exponent is less than one, $u(x) = x^{.802}$. LGDU(4), with all parameters estimated from the data, is also able to correctly reproduce all of the modal choices.

Conclusions of Experiment 1

Experiment 1 shows that strong violations of coalescing are obtained even with very large consequences and even with event-framing (marble colors) designed to facilitate coalescing. Despite changes in procedure designed to "help" CPT, results continue to violate CPT, extending previous results. Apparently, the use of very large, hypothetical consequences appears to require no new principle aside from the need to estimate a utility function with diminishing marginal returns. These results also violate the model of Lopes and Oden (1999) that satisfies coalescing and stochastic dominance.

In Experiment 1, all tests of coalescing confounded upper and lower coalescing. For example, from choices 6 to 9 of Table 2, both the upper branch of S and the lower branch of R were coalesced in choice 6 and both were split in choice 9. So, by violating lower coalescing, LGDU was able to account for the data of Experiment 1. In Experiments 2 and 3, these two types of coalescing are teased apart.

Experiments 2 and 3

These experiments compare three models that violate coalescing, which appears to be at the root of violations of stochastic dominance and of Allais paradoxes. These three models are SWU, which violates both upper and lower coalescing and satisfies branch-splitting independence, LGDU, which violates lower coalescing but satisfies upper coalescing, and TAX, which violates both upper and lower coalescing in opposite ways and violates branch-splitting independence.

Method of Experiments 2 and 3

As in Experiment 1, people participated via the Internet, and clicked a button beside the gamble in each choice that they preferred. Unlike Experiment 1, participants in Experiment 2 were informed that three participants would be chosen at random to receive the prize of one of their chosen gambles. The consequences in Experiment 2 ranged from \$0 to \$108.

Participants in Experiment 3 were told that 3 participants per 100 would receive a gift of \$50 plus the result of one of their chosen gambles. Experiment 3 had the same choices as in Experiment 2, except that each consequence was reduced by \$45. This procedure created gambles with framed losses, in which participants could lose some of their endowment of \$50, but they did not risk their own money. As promised, prizes were awarded to six winners (one declined the prize).

Experiments 2 and 3 included 22 choices between gambles. The first four choices were the same as in Birnbaum (2004a), but the other 18 trials were constructed to allow tests of branch-splitting independence. There were two series with 9 choices each, shown in Tables 7 and 8. The position of the "safe" and "risky" gambles within choices was counterbalanced between series A and B. The consequences were non-negative in all three conditions of Experiment 2, but they were framed as mixed gambles in Experiment 3.

There were three studies within Experiment 2, with different participants in each. Study A2 was run first. Next, participants were randomly assigned to Conditions A3 or A4, where A3 used the same consequences as A2. Condition A4, however, used slightly altered values of the consequences when a branch was split, designed to see if people might be more or less likely to coalesce if the branches led to exactly equal

Table 7
Dissection of Allais paradox (series A): percentages choosing "Risky" gamble in conditions A2, A3, or A4 of Experiment 2 and in mixed gambles of Experiment 3 (AM)

No.	Type		Choice		Condition				Prior TAX model		Prior LGDU model	
	S	*R*	Gamble, *S*	Gamble, *R*	203 A2	315 A3	325 A4	152 AM	*S*	*R*	*S*	*R*
6	*C*	*C*	20% ($40, $40, −$5) 80% ($2, $2, −$43)	10% ($98, $98, $53) 90% ($2, $2, −$43)	65	60	60	84	9.0	13.3	8.4	12.9
11	*US*	*C*	10% ($40, $40, −$5) 10% ($40, $*39*, −$5) 80% ($2, $2, −$43)	10% ($98, $98, $53) 90% ($2, $2, −$43)	36	35	33	***80***	11.1	13.3	8.4	12.9
21	*C*	*LS*	20% ($40, $40, −$5) 80% ($2, $2, −$43)	10% ($98, $98, $53) 10% ($2 $*3*, −$43) 80% ($2, $2, −$43)	53	55	63	78	9.0	9.6	8.4	7.2
9	*US*	*LS*	10% ($40, $40, −$5) 10% ($40, $*39*, −$5) 80% ($2, $2, −$43)	10% ($98, $98, $53) 10% ($2 $*3*, −$43) 80% ($2, $2, −$43)	39	39	44	***70***	11.1	9.6	8.4	7.2
12			100% ($40, $40, −$5)	10% ($98, $98, $53) 80% ($40, $40, −$5) 10% ($2, $2, −$43)	47	49	50	61	40.0	30.6	40.0	31.9
16	*LS*	*US*	80% ($98, $98, $53) 10% ($40, $*39*, −$5) 10% ($40, $40, −$5)	80% ($98, $98, $53) 10% ($98, $*97*, $53) 10% ($2, $2, −$43)	59	56	59	73	59.8	62.6	65.0	65.8
7	*LS*	*C*	80% ($98, $98, $53) 10% ($40, $*39*, −$5) 10% ($40, $40, −$5)	90% ($98, $98, $53) 10% ($2, $2, −$43)	24	21	25	38	59.8	54.7	65.0	65.8
13	*C*	*US*	80% ($98, $98, $53) 20% ($40, $40, −$5)	80% ($98, $98, $53) 10% ($98, $*97*, $53) 10% ($2, $2, −$43)	57	58	67	65	68.0	62.6	71.4	65.8
19	*C*	*C*	80% ($98, $98, $53) 20% ($40, $40, −$5)	90% ($98, $98, $53) 10% ($2, $2, −$43)	22	23	31	25	68.0	54.7	71.4	65.8

Note: Values in parentheses show consequences of conditions A2 (and A3), A4, and AM, respectively; italics indicate values altered in A4. The common branch is 80% to win $2 in choices 6, 11, 21, and 9, 80% to win $40 in choice 12, and 80% to win $98 in choices 16, 13, 7, and 19. The notation *C* under *S* or *R* indicates a gamble in coalesced form; *US* or *LS* indicate upper or lower branch is split, respectively. Bold font shows where mixed and positive gambles yield different modal choices.

consequences (A2 or A3) as opposed to slightly different consequences (A4). The cash consequences in Tables 7 and 8 show the framed consequences in Study A2 (A3), A4, and Experiment 3 (labeled AM = Allais mixed), respectively.

There were 203, 315, and 325 participants in conditions A2, A3, and A4, respectively. Of the 843 in Experiment 2, 67% were 22 years of age or younger, 8% were 40 or older, and 64% were female.

Experiment 3 had 152 participants in condition AM, who received the same choices as in A3, except all consequences were reduced by $45, and each winner would receive an endowment of $50. This modification creates framed "mixed" gambles. Participants in Experiment 3 were college undergraduates who were tested in the laboratory and completed the entire task twice. Of the 152, 75% were female and 92% were 22 years of age or less. Complete materials, with instructions, can be viewed from the following URL: http://psych.fullerton.edu/mbirnbaum/archive.htm.

Within each series, there are two, 2 × 2, lower split by upper split, factorial designs in which the lower branch of one gamble and the higher branch of the other gamble are either split or coalesced. For example, in Table 7, choices #6, 11, 21, and 9 constitute one 2 × 2, upper branch split in *S* by lower branch split in *R*, factorial design. The two levels of split are either coalesced (*C* under "type") or split into two branches (*US* or *LS* designate splitting of upper or lower branches, respectively). For example, choices 6 and 11 of Table 7 differ

Table 8
Dissection of Allais paradox (series B): percentages choosing "Risky" gamble

No.	Type		Choice		Condition				Prior TAX model		Prior GDU model	
	S	*R*	Gamble, *S*	Gamble, *R*	A2	A3	A4	AM	*S*	*R*	*S*	*R*
10	*C*	*C*	15% ($50, $50, $5) 85% ($7, $7, −$38)	10% ($100, $100, $55) 90% ($7, $7, −$38)	78	78	75	81	13.6	18.0	13.1	17.6
15	*US*	*C*	10% ($50, $50, $5) 05% ($50, $*49*, $5) 85% ($7, $7, −$38)	10% ($100, $100, $55) 90% ($7, $7, −$38)	52	44	42	***62***	15.6	18.0	13.1	17.6
22	*C*	*LS*	15% ($50, $50, $5) 85% ($7, $7, −$38)	10% ($100, $100, $55) 05% ($7, $*8*, −$38) 85% ($7, $7, −$38)	70	68	73	62	13.6	14.6	13.1	12.6
17	*US*	*LS*	10% ($50, $50, $5) 05% ($50, $*49*, $5) 85% ($7, $7, −$38)	10% ($100, $100, $55) 05% ($7, $*8*, −$38) 85% ($7, $7, −$38)	44	42	46	46	15.6	14.6	13.1	12.6
20			100% ($50, $50, $5)	10% ($100, $100, $55) 85% ($50, $50, $5) 05% ($7, $7, −$38)	50	48	52	71	50.0	40.1	50.0	44.0
14	*LS*	*US*	85% ($100, $100, $55) 10% ($50, $*51*, $5) 05% ($50, $50, $5)	85% ($100, $100, $55) 10% ($100, $*99*,$55) 05% ($7, $7, −$38)	69	60	65	***42***	68.4	69.7	74.9	77.5
5	*LS*	*C*	85% ($100, $100, $55) 10% ($50, $*51*, $5) 05% ($50, $50, $5)	95% ($100, $100, $55) 05% ($7, $7, −$38)	37	33	41	20	68.4	62.0	74.9	77.5
18	*C*	*US*	85% ($100, $100, $55) 15% ($50, $50, $5)	85% ($100, $100, $55) 10% ($100, $*99*, $55) 05% ($7, $7, −$38)	70	65	70	***48***	75.7	69.7	79.8	77.5
8	*C*	*C*	85% ($100, $100, $55) 15% ($50, $50, $5)	95% ($100, $100, $55) 05% ($7, $7, −$38)	25	26	30	14	75.7	62.0	79.8	77.5

Note: Values in parentheses show consequences in conditions A2, A4, and Experiment 3, respectively. Italics show altered consequences in condition A4. The common consequence is either 85% to win $7 in choices 10, 15, 22, and 17, 85% to win $50 in choice 20, or 85% to win $100 in choices 14, 5, 18, and 8. The notation *C* indicates coalesced form, *US* or *LS* indicate upper or lower branch is split, respectively. Bold font is used to show cases where mixed gambles show different modal choices from strictly positive gambles.

only in that the upper branch of *S* (gamble on the left) has been split in Trial 11 and it is coalesced in choice 6. Choices #6 and 21 differ only in that the lower branch of *R* (gamble on the right) has been split in choice 21 and it is coalesced in choice 6. In choice #9, both gambles have these branches split.

Similarly, choices 16, 7, 13, and 19 (last four rows in Table 7) also constitute a 2×2 factorial design; which in this case is a lower split in *S* by upper split in *R*, factorial design. The common branch in series A is either 80% to win $2 (first four rows), $40 (choice 12), or $98 (last four rows).

Combining series A and B (Tables 7 and 8), there are 8 tests of splitting/coalescing of upper branches and 8 tests of splitting/coalescing lower branches. Each series provides one test of restricted branch independence (choices 9 and 16 in Table 7 and choices 17 and 14 in Table 8). Finally, each series has all three types of Allais paradoxes [(Type 1: choices 6 & 12, Type 2: 12 & 19, and Type 3: 6 & 19) and (Type 1: choices 10 & 20, Type 2: 20 & 8, and Type 3: 10 & 8)]; in these three "basic" paradoxes, TAX, LGDU, and CPT all agree in their predictions.

Results of Experiment 2

The percentages of participants who chose the "risky" gamble in each choice are displayed in Tables 7 and 8. The results for conditions A2, A3, and A4 (which used slightly different consequences when a branch was split) of Experiment 2 are virtually identical. (Results from Experiment 3 are shown in the column labeled AM, and are described below).

According to EU, there should be no differences between any two rows in Tables 7 or 8. Instead, the percentages choosing the "risky" gamble vary from significantly greater than 50% in the first row (60% to 65% in choice 6) to significantly less than 50% (22% to 31%) in the last row (choice 19) in Table 7. In Table 8, percentages vary from 75% or above (choice 10) to 30% or below (choice 8). Clearly, these large changes represent systematic Allais paradoxes, violating EU.

Note that the first four choices listed in Table 7 involve identical prospects. Choices 6, 11, 21, and 9 differ only in how branches are coalesced or split. According to CPT/RDU, all decisions in the first four rows should be the same, apart from error. In violation of these predictions (A2, A3, and A4), percentages choosing the "risky" second gamble (R) changed from significantly greater than 50% (in choice 6), to significantly less than 50% (in choice 11).

All choice percentages in Table 7, averaged over 843 participants in Experiment 2 (A2–A4) are significantly different from 50%, except for choice 12. By the (more sensitive) within-subjects test of correlated proportions, difference between choice 6 and 11 is also significant when tested separately in all three studies, $z = 6.18$, 6.86, and 7.55 in A2, A3, and A4, respectively.

According to CPT, choice percentages should also be the same within the last four rows of Table 7, within first four rows of Table 8, and within the last four rows of Table 8. Instead, the results of Experiment 2 significantly refute CPT in all four 2×2 tests.

According to LGDU, there should be no difference between any two rows in which only the upper consequence is split or coalesced. Thus, LGDU does not predict a change between choices 6 and 11 (first two rows in Table 7). In Table 7, choices 6 & 11, 21 & 9, 16 & 7, and 13 & 19 should also be the same within each pair. In Table 8, choices 10 & 15, 22 & 17, 14 & 5, and 18 & 8 should also be the same within pairs. Predictions for the parameterized form of LGDU are shown in the last two columns of Tables 7 and 8.

According to the TAX model with parameters estimated from previous data, there should be violations of both lower and upper coalescing, as shown by the changing cash equivalent values for the gambles (A2), which are presented in columns under "prior TAX model."

The new feature of Experiment 2 is that it separately tests the effects of splitting the branch with the higher or lower valued consequence. According to prior TAX predictions, splitting the higher consequence should improve a gamble and splitting the branch with the lowest consequence should lower the value of the gamble, violating event-splitting independence. There are 8 tests of each of these predictions.

Splitting the branch with the higher consequence significantly increased the proportion choosing that gamble in all eight tests. In choice 6×11, 21×9, 19×13, 7×16, 10×15, 22×17, 8×18, and 5×14, the binomial test of correlated proportions yielded $z = 11.8$, 7.6, 14.3, 14.9, 14.0, 11.8, 15.8, and 11.8, respectively. (The critical value of $|z|$ for $\alpha = .05$ level of significance is 1.96.). All eight are significant and all are in the direction predicted by TAX. All produce reversals of the majority preference as well. These results are not consistent with EU, CPT, or LGDU, which imply no effects of splitting the upper branch. They are also consistent with SWU.

Splitting the branch with the lower consequence had effects that were smaller and less consistent. Of the eight tests on the combined data, three were statistically significant. The four tests with the largest values of $|z|$ were choices 6×21, 11×9, 10×21, and 8×5, with $z = 1.94$, -3.01, 3.18, and 5.49, respectively, with three of these four in the direction predicted by TAX. According to SWU, branch-splitting independence holds; if splitting the upper branch improved a gamble, then splitting the same branch in the lower position should have also improved the gamble.

In choices 10 and 15 of Table 8, for example, \$50 was the highest consequence in the S gamble; splitting the 15% branch to win \$50 significantly increased the percentage choosing this gamble (a shift of about 30%). In choices 8 and 5 the lowest consequence of S is \$50; splitting this branch makes people more likely to choose the other gamble, consistent the idea that this split made the gamble worse, as predicted by TAX. Therefore, choices 10, 15, 8, and 5 represent a case where splitting the same probability (0.15) to receive the same consequence (\$50) can either significantly increase or significantly decrease the proportion who choose the gamble. The effect of the split depends on whether that consequence (in this case, \$50) is the highest or lowest consequence in the gamble. The effect of splitting \$50 in the lower branch, was statistically significant when tested in each of A2, A3, and A4 separately, but it produced modest shifts of only about 10%.

The prior TAX model correctly predicted the majority choice in all but four choices of Experiment 2: choices 11 and 13 in Table 7, and choices 15 and 18 in Table 8. In all cases of discrepancy, the choices involve unequal numbers of branches in the two gambles, and the discrepancy from TAX can be described as follows: the majority chose the gamble with the greater number of (positive) branches. This finding suggests that the configural weight transfer might be better represented by δ/n or even $\delta/(n-1)$, rather than by $\delta/(n+1)$.

In each table, there is one test of restricted branch independence. Note that in Table 7, choice 9 has a common branch of 80 marbles to win \$2 and in choice 16, there is a common branch of 80 marbles to win \$98. In Table 8, choices 17 and 14 differ only in the consequence on the common branch of 85 marbles to win (either \$7 in choice 17 or \$100 in choice 14). The marble colors on the common branches were the same in Table 7 and different in Table 8, so choices 9 and 16 are "framed" (supposedly making it easier to notice and cancel the common branch). choices 17 and 14 are not "framed" because there were different colors on the common branch.

In both tests, there are significant violations of restricted branch independence ($z = 8.03$ and 8.87). These significant violations refute SWU and SWAU, which imply branch independence. The observed violations are in the opposite direction from the pattern predicted by the prior CPT model. They are also opposite what is required by CPT to account for the Allais paradox between choices 6 and 19 and between choices 10 and 8, respectively, which are supposedly due to violations of branch independence, according to CPT. Both trends agree, however, with the prior TAX model; and all six tests also show reversals of the modal choice.

There was no discernable evidence of event-framing (marble colors). The failure to find event-framing effects does not disprove their possible existence, of course, but the continued failure to find any substantial effect in any of several experiments (see also Birnbaum, 2004a, 2004b) suggests that this type of framing is of minor importance, if any.

Fit of parametric models to Experiment 2

The parametric models were fit to the combined data of Experiment 2 to minimize the sum of squared discrepancies between predicted and observed choice proportions. The results are shown in Table 9, which also lists choices where the models failed to predict the modal choices. Although SWU(4) model (4 free parameters) achieved a better fit than prior TAX(1), SWU cannot account for violations of restricted branch independence nor can it account for the small, but significant violations of branch-splitting independence. The prior TAX model fits nearly as well as SWU(4) using only one free parameter (α), and it does a better job than CPT or LGDU models with four free parameters. However, TAX did not always predict the majority choice in cases when the number of branches differed within the choice, even with all four parameters were estimated from the data.

Results of Experiment 3

Results for mixed gambles (Experiment 3) are shown under column AM in Tables 7 and 8. At first glance, results for framed "mixed" gambles appear similar to their corresponding choices in Experiment 2. Indeed, the percentages choosing the "risky" gamble show large decreases from the top to bottom rows in both tables, resembling the main trends in Experiment 2. However, there are some important differences, indicated in italicized bold font. Note that the modal responses in choices 11 and 9 in Table 7 differ between AM and the conditions with positive consequences, and modal choices in #15, 14, and 18 in Table 8 also differ between the two experiments. In four of these cases, the modal choice in "mixed" gambles favors the gamble with the smaller number of branches leading to negative consequences (i.e., to losses).

In addition, there is a significant majority preference in both tables for the "mixed" gamble over the sure thing in both choices #12 and #20, suggesting that these participants are less "loss averse" than has been found when participants make hypothetical choices among gambles that risk their own money (Tversky & Kahneman, 1992).

Birnbaum and Bahra extrapolated the prior TAX model to mixed gambles by the assumptions that the utility function is $u(x) = x$ for $-\$100 \leqslant x \leqslant \$100$ and that the same configural parameter (δ) applies for gambles with mixed consequences as for gambles on strictly positive consequences. These assumptions imply that people should make same choices in Experiment 3 as in Experiment 2; that is, these simplifying assumptions imply no framing effects. There were five cases where the prior TAX model with these assumptions failed to

Table 9
Fit of models to group choice proportions (Experiment 2)

Model	Estimated parameter				SS deviations	Failures to predict modal choices
	α	β	γ	δ		
Prior CPT(1)	0.323	(0.88)	(0.61)	(0.72)	0.769	8 (#11, 9, 16, 13 15, 17, 20, 18)
Prior LGDU(1)	0.077	(1)	(1.382)	(0.542)	0.434	9 (#11, 21, 13, 7, 22, 15, 20, 5, 18)
LGDU(4)	1.864	0.462	0.863	0.909	0.396	8 (#11, 21, 16, 13, 15, 17, 5, 8)
CPT(4)	0.920	0.555	0.810	0.998	0.378	7 (#11, 9, 16, 13, 15, 17, 18)
SWU(1)	0.100	(0.88)	(0.61)	(0.72)	0.370	8 (# 11, 9, 12, 19, 15, 17, 5, 8)
Prior TAX(1)	0.049	(1)	(.7)	(1)	0.335	4 (#11, 13, 15, 18)
SWU(4)	1.000	0.65	0.85	1.24	0.100	4 (#12, 16, 15, 17)

Note: Failures to predict modal choices in combined data of Experiment 2 are listed in parentheses by the choice numbers in Tables 7 and 8. TAX model could be improved by estimating parameters from the data, but even with estimated parameters it did not account for all choices.

predict the modal choice: choices 9, 12 and 13 in Table 7 and choices 20 and 14 in Table 8.

Two of these cases of discrepancy involve restricted branch independence, which is tested in choices 9 and 16 of Table 7 and in choices 17 and 14 in Table 8. There are two tests of each case, in first and second replicates. None of the four tests of violations of branch independence in AM are very large and none were significant in Experiment 3, in contrast to Experiment 2, where all six tests (three conditions by two tests) were substantial and significant.

Because there are two replicates, there are 16 tests of upper coalescing and 16 tests of lower coalescing. In 12 of these tests, the upper consequence was positive; splitting the upper consequence in these cases increased the proportion choosing that gamble (which decreased the proportion choosing the other gamble). In all four tests where the upper consequence was negative (−\$5), splitting this consequence lowered the inferred value of the gamble, but in only one case was it significant.

Of 16 tests of splitting the lower consequence, there were 12 tests where the lower consequence was negative. In all 12 cases, splitting the lower (negative) consequence made the gamble less likely to be chosen; in 8 of these 12 cases it was significant. Of the four tests where the lower consequence was positive (\$5), splitting made it better in two cases and worse in two cases. Only one of the four tests was significant, where splitting the lower consequence of \$5 made the gamble worse (choices 5 versus 8). When the upper consequence was \$5, splitting this consequence always made the gamble significantly better, so we again have a small, but significant violation of branch-splitting independence (choices 10, 15, 5, and 8).

Experiment 4

Experiment 4 tests a specific prediction of a model by Meginniss (1976), whose work went largely unheralded until it received new attention and generalization by Luce et al. (2006). Meginniss (1976) developed a "new class of utility models" that violate idempotence and coalescing. For gambles, $G = (x_1, p_2; x_2, p_2; \ldots; x_n, p_n)$, Meginniss utility (MU) can be written as follows:

$$\mathrm{MU}(G) = \sum_{i=1}^{n} p_i^{\kappa} u(x_i) + \lambda H^{(\kappa)}(p_1, p_2, \ldots, p_n), \kappa > 0, \tag{15}$$

where

$$H^{(\kappa)}(p,p,\ldots,p) = \begin{cases} \frac{1}{\ln 2}(-\sum_{i=1}^{n} p_i \ln p_i), & \kappa = 1, \\ \frac{1}{1-2^{1-\kappa}}(1 - \sum_{i=1}^{n} p_i^{\kappa}), & \kappa \neq 1, \ \kappa > 0. \end{cases}$$

In the case where $\kappa = 1$, Meginniss utility (MU) reduces to expected utility plus Shannon information entropy. This case ($\kappa = 1$) was recently proposed independently by Yang and Qiu (2005) as a descriptive model of Allais paradoxes. One can think of the entropy term as a utility for gambling, apart from the consequences of the gamble. Experiment 4 tests an implication of this model. In this model, there is no influence of rank on weight. Therefore, in a series of choices between low range ("safe") and higher range ("risky") gambles with the same EU and equal entropy, there should be no effect of whether it is upper or lower branches that are split in both gambles in each choice. In contrast, TAX implies that splitting the upper branches of both gambles should improve R relative to S and splitting the lower branches of both gambles should improve S relative to R.

Method of Experiment 4

Experiment 4 included a series of 5 experimental choices, intermixed among 27 other "filler" choices involving mixed gambles. These were presented to 196 participants, using the same procedures as in Experiment 3. Each participant completed each choice twice, separated by intervening tasks requiring about 15 min. These choices are shown in Table 10. Note that all of the gambles involve the same objective choice, between a fifty-fifty gamble to win either \$43 or \$47 ("safe") and a fifty-fifty gamble to win either \$11 or \$89 ("risky"). That is, all of the choices are the same as choice 7 in Table 10, assuming coalescing.

Results of Experiment 4

According to the TAX model, splitting the upper branch of both gambles should improve R relative to S, (choices 7, 23, and 13), whereas splitting the lower branch of both gambles should make R worse relative to S (choices 7, 25, and 9). According to the Yang and Qiu (2005) model (the special case of Meginniss utility with $\kappa = 1$), however, there should be no effect of rows because EU of S and R are unaffected by splitting, and entropy is equal (in S and R) within each choice.

Table 10 shows the choice percentages. Note percentages who chose the risky (R) gamble increase from 31% to 39% in choice 9, when lower branches are split, to 58–63% in choice 7, where both gambles are coalesced, to 73–75% in choice 13, where the upper branches are split. Each difference between successive rows is significant in at least one replicate by the test of correlated proportions. For choices 9×25, $z = 2.67$, 3.62; choices 25×7, $z = 4.00$, 2.29; choices 7×23, $z = 3.30$, 1.03, for choices 23×13, $z = 0.57$, 3.26, in replicates 1 and 2, respectively. All two-step and higher differences were significant in both replicates.

Table 10
Test of EU plus entropy model

No.	Choice		Repetition		Prior TAX model	
	Gamble, S	Gamble, R	Rep 1	Rep 2	S	R
9	50 black win \$47 15 purple win \$43 15 white win \$43 10 gray win \$43 10 brown win \$43	50 black win \$89 15 purple win \$11 15 white win \$11 10 gray win \$11 10 brown win \$11	31	39	43.5	21.4
25	50 black win \$47 25 purple win \$43 25 white win \$43	50 black win \$89 25 purple win \$11 25 white win \$11	42	53	43.9	28.5
7	50 black win \$47 50 purple win \$43	50 black win \$89 50 purple win \$11	58	63	44.3	37.0
23	25 black win \$47 25 purple to win \$47 50 brown win \$43	25 black win \$89 25 purple win \$89 50 brown win \$11	71	67	44.7	43.3
13	10 black win \$47 10 purple win \$47 15 white win \$47 15 gray win \$47 50 brown win \$43	10 black win \$89 10 purple win \$89 15 white win \$89 15 gray win \$89 50 brown win \$11	73	75	45.0	50.1

Percentages choosing "Risky" gamble ($n = 196$).

This systematic trend is not consistent with the Yang and Qiu (2005) model (i.e., the Meginniss model with $\kappa = 1$). However, this trend would be consistent with the Meginniss model if $\kappa < 1$ and $\lambda = 0$. [With $\kappa = 0.52$, $u(x) = x$ and $\lambda = 0$, Eq. (13) with $\alpha = 0.037$, this model fits the trend in Table 10]. In that case, the model becomes a type of SWU model. It cannot account for violations of restricted branch independence, however, since the entropy terms are all equal in that property (Expression 1) and the model's weights are independent of rank. Experiment 4 does not test the generalization by Luce et al. (2006), which allows an another term for utility associated with the outcomes of the chance experiment, apart from the utility of their consequences.

Predictions of the TAX model with parameters estimated from previous data are shown in the right columns of Table 10. These predict that the "risky" gamble improves from row to row whereas the "safe" gamble changes little, so the percentage choosing the R gamble should increase as the lower branches are coalesced and the upper ones are split. However, the previous parameters predict that the majority should have reversed preference between choices 23 and 13, whereas the observed modal data switched between choices 25 and 7 in the first replicate and between choices 9 and 25 in the second replicate. These data thus show less risk aversion than predicted by the prior model. The TAX model provides a better fit with the assumption that δ is about 0.4 instead of 1.0.

Discussion

There are eight conclusions that can be drawn.

- First, the evidence against expected utility theory (provided by the significant changes from row to row within each of Tables 2, 3, 7, 8 and 10) is overwhelming.
- Second, evidence of event-splitting effects and violations of stochastic dominance, which refute rank-dependent utility theories, including CPT, is also overwhelming. This evidence also contradicts EU. There are significant effects of splitting in all five experiments. These effects are not only significant, but are large enough to reverse the majority preference in choices 6 and 9 (Tables 2 and 7), 10 and 17 (Tables 3 and 8), 5 and 11 (Table 4), 7 and 13 (Table 4), 16 and 19 (Table 7) and 14 and 8 (Table 8), 9 and 13 (Table 10). Significant majority violations of stochastic dominance were observed in choices 5, 15, 7, and 18 in Table 4, and these could also be reversed by appropriate splitting (choices 11 and 13 in Table 4).
- Third, there is evidence of significant violations of restricted branch independence, but these violations among gambles on positive gambles are significantly more frequent in opposition to the effects predicted by CPT with an inverse-S weighting function than in agreement with that model (choices 9 and 16 in Table 7 and 17 and 14 in Table 8). The empirical violations

are also contrary to the direction of the Allais paradox, so it is hard to defend the position (required by any version of CPT) that Allais paradoxes are produced by violations of restricted branch independence. They are consistent, however, with prior TAX and the hypothesis that Allais paradoxes are produced by violations of coalescing.

- Fourth, there is strong evidence that splitting the branch with the higher valued consequence improves a gamble. This effect contradicts the idempotent LGDU model, which satisfies upper coalescing as well as EU and CPT, which satisfy all forms of coalescing.
- Fifth, there is statistically significant, but far less dramatic, evidence that splitting the branch with the lower-valued consequence can make a gamble worse. Splitting a .2 branch to win \$50 significantly improved the probability of choosing the gamble when \$50 is the highest consequence in the gamble (choices 10 and 15 in Table 8), and splitting the same branch significantly lowers the proportion of choosing that gamble when \$50 was the lowest consequence in the gamble (choices 5 and 8 in Table 8). Although significant, these violations of branch-splitting independence are not as impressive as the effects of splitting the higher consequence. A possible confounding factor in Experiments 2–3 is the fact that the design involves choices between gambles with differing numbers of branches. A tendency to prefer gambles with the greater number of branches yielding positive prizes might compete with the effect of splitting the lower valued branch of a gamble. A second possible confounding factor is the possibility that standard deviations in Eq. (13) might not be equal, which could bias the test (Expression 3) of branch-splitting independence.
- Sixth, the prior TAX model predicted successfully the modal choices among gambles with equal numbers of branches, but did not always make correct predictions in cases of gambles with unequal numbers of branches.
- Seventh, the effects of splitting and coalescing appear similar for gambles involving large and small consequences, and the prior TAX model in most cases correctly describes the main trends in the data, unlike the CPT model. The only complication for TAX to fit data involving large consequences was to use a nonlinear utility function.
- Eighth, splitting the upper branches of both R and S improves R relative to S and splitting the lower branches of both R and S diminishes R relative to S. This finding is not consistent with the Yang and Qiu (2005) model, which is a special case of Meginniss utility (Eq. (15)) with $\kappa = 1$.

In summary, the present data add to the growing case against RDU and CPT models as descriptive theories of risky decision making. Those theories attribute the Allais paradoxes to violations of branch independence rather than to violations of coalescing. Violations of coalescing, such as those reported here, refute those models with any functions and parameters. Furthermore, the observed violations of restricted branch independence go in the opposite direction from the Allais paradoxes and in the opposite direction from the pattern required by the parametric model of CPT.

RDU and CPT, unlike Birnbaum's configural weight models, assume that decumulative weight is a monotonic function of decumulative probability; this assumption forces coalescing and stochastic dominance. Indeed, that is why Quiggin (1985, 1993) adopted this "rank-dependent" representation. The assumption of coalescing had strong intuitive appeal to Kahneman and Tversky (1979; see Kahneman, 2003). Because RDU and CPT satisfy coalescing and transitivity, they are forced to explain Allais paradoxes as the result of violations of restricted branch independence. The empirical data, however, give clear answers regarding these prior intuitions: people do not obey coalescing. Instead, splitting a branch appears to give that branch greater weight. The data indicate that violations of branch independence oppose the Allais paradoxes, contradicting the implication of CPT that it is violations of branch independence that cause the Allais paradoxes.

The idempotent lower gains decomposition model (Marley & Luce, 2001) is more accurate than RDU or CPT because it violates coalescing. Therefore, it can account for the violations of coalescing and stochastic dominance in the first experiment. However, this model is contradicted by violations of upper coalescing, which were tested separately of lower coalescing in Experiments 2 and 3. A parametric version of this model was fit to the data, and was found to be more accurate than CPT, but not quite as accurate as TAX or SWU, even when it had more free parameters. The non-idempotent GDU models (Luce et al., 2006; Marley & Luce, 2005) have not yet been tested, except for the special case tested in Experiment 4.

The SWU model, like TAX, assumes that weights are a function of branch probabilities, so it violates coalescing and stochastic dominance. SWU cannot account for violations of restricted branch independence, however, nor can it account for violations of branch-splitting independence. Because there were only small violations of branch splitting independence and only two tests of restricted branch independence in Experiment 2, the SWU(4) model was able to achieve a fit comparable to TAX(1) in that study when it was allowed enough free parameters.

Birnbaum's (1997, 2005) rank affected multiplicative weights (RAM) model is virtually identical to the TAX model in this experiment. RAM and TAX can be distinguished by testing properties such as distribution independence (Birnbaum, 2005; Birnbaum & Chavez, 1997). Such tests have favored TAX over RAM.

Like EU or CPT, the TAX model assumes that the utility of a gamble is a weighted average of the utilities of its consequences. Unlike CPT, however, TAX assumes that (1) attention to a branch is initially a function of branch probability (not decumulative probability), and (2) weight is drawn from branches leading to good consequences and transferred to branches with bad consequences. Although some might use the term ''pessimism'' for this transfer of weight, I do not think that people believe that the lower consequence is more likely to happen; I think they simply consider this consequence of greater importance. Indeed, when we place judges in the seller's point of view, the configural weighting pattern is reversed (Birnbaum & Stegner, 1979).

The configural transfers of weight account for risk aversion in fifty-fifty gambles and also produce the observed pattern of violation of branch independence. Birnbaum and Stegner (1979) represented the difference between buying and selling prices in terms of a change in this transfer parameter, δ, in Eq. (9). Whereas buyers (like people in this study who choose gambles) shift more attention to the bad consequences, sellers tend to place relatively more weight on higher consequences (Birnbaum & Sutton, 1992; Birnbaum & Beeghley, 1997; Birnbaum & Zimmermann, 1998). These two psychological distinctions between TAX and CPT account for the advantages in descriptive accuracy of TAX over RDU or CPT as accounts of choices and of the so-called ''endowment effect'' in judged prices (Birnbaum & Zimmermann, 1998).

Although the prior TAX model provided better predictions than other prior models, there are three findings that suggest possible improvements to that model. First, to account for large consequences in TAX, it was necessary to utilize a nonlinear utility function. This result is hardly surprising, nor is it really a contradiction to the TAX model.

Second, when framed ''mixed'' gambles were created by subtracting a constant from all consequences and adding that same constant as an endowment, people made different decisions from those they made in the strictly positive case. These consequence framing effects contradict the simplifying assumption used by Birnbaum and Bahra (in press) that the same linear utility function and configural transfer parameter apply in the (framed) ''mixed'' case as in the strictly positive case. Predictions of TAX might be improved by allowing utility or configural weighting parameters to differ for branches leading to positive and negative consequences, as suggested by Edwards (1962) and used in both versions of prospect theory.

Third, the prior TAX model made systematic errors of prediction when the numbers of branches in the two gambles compared were unequal. If the configural weight transfer were proportional to $\delta/(n-1)$ instead of $\delta/(n+1)$, then splitting positive branches should tend to improve gambles and splitting negative branches should tend to make gambles worse.

Aside from theoretical conclusions, there are very clear practical implications of these results for negotiation and bargaining. Sellers (and others who want to make an option seem more attractive) should divide branches (or events) that lead to its best consequence and coalesce branches (events) leading to its worst consequence. Those who want to make an option seem less attractive, however, should split branches leading to worst consequences and coalesce those leading to the best consequences.

References

Allais, M. (1979). The foundations of a positive theory of choice involving risk and a criticism of the postulates and axioms of the American School. In M. Allais & O. Hagen (Eds.), *Expected utility hypothesis and the Allais paradox* (pp. 27–145). Dordrecht, The Netherlands: Reidel.

Birnbaum, M. H. (1997). Violations of monotonicity in judgment and decision making. In A. A. J. Marley (Ed.), *Choice, decision, and measurement: Essays in honor of R. Duncan Luce* (pp. 73–100). Mahwah, NJ: Erlbaum.

Birnbaum, M. H. (1999a). Paradoxes of Allais, stochastic dominance, and decision weights. In J. Shanteau, B. A. Mellers, & D. A. Schum (Eds.), *Decision science and technology: Reflections on the contributions of Ward Edwards* (pp. 27–52). Norwell, MA: Kluwer Academic Publishers.

Birnbaum, M. H. (1999b). Testing critical properties of decision making on the Internet. *Psychological Science, 10*, 399–407.

Birnbaum, M. H. (2001). A Web-based program of research on decision making. In U.-D. Reips & M. Bosnjak (Eds.), *Dimensions of Internet science* (pp. 23–55). Lengerich, Germany: Pabst Science Publishers.

Birnbaum, M. H. (2004a). Causes of Allais common consequence paradoxes: an experimental dissection. *Journal of Mathematical Psychology, 48*(2), 87–106.

Birnbaum, M. H. (2004b). Tests of rank-dependent utility and cumulative prospect theory in gambles represented by natural frequencies: effects of format, event framing, and branch splitting. *Organizational Behavior and Human Decision Processes, 95*, 40–65.

Birnbaum, M. H. (2005). Three new tests of independence that differentiate models of risky decision making. *Management Science, 51*, 1346–1358.

Birnbaum, M. H., & Bahra, J. (in press). Gain-loss separability and coalescing in risky decision making. Management Science, xx, xxx–xxx.

Birnbaum, M. H., & Beeghley, D. (1997). Violations of branch independence in judgments of the value of gambles. *Psychological Science, 8*, 87–94.

Birnbaum, M. H., & Chavez, A. (1997). Tests of theories of decision making: violations of branch independence and distribution independence. *Organizational Behavior and Human Decision Processes, 71*(2), 161–194.

Birnbaum, M. H., & Martin, T. (2003). Generalization across people, procedures, and predictions: violations of stochastic dominance and coalescing. In S. L. Schneider & J. Shanteau (Eds.), *Emerging perspectives on decision research* (pp. 84–107). New York: Cambridge University Press.

Birnbaum, M. H., & Navarrete, J. B. (1998). Testing descriptive utility theories: violations of stochastic dominance and cumulative independence. *Journal of Risk and Uncertainty, 17*, 49–78.

Birnbaum, M. H., & Stegner, S. E. (1979). Source credibility in social judgment: bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology, 37*, 48–74.

Birnbaum, M. H., & Sutton, S. E. (1992). Scale convergence and utility measurement. *Organizational Behavior and Human Decision Processes, 52*, 183–215.

Birnbaum, M. H., & Zimmermann, J. M. (1998). Buying and selling prices of investments: configural weight model of interactions predicts violations of joint independence. *Organizational Behavior and Human Decision Processes, 74*(2), 145–187.

Busemeyer, J. R., & Townsend, J. T. (1993). Decision Field Theory: a dynamic cognition approach to decision making. *Psychological Review, 100*, 432–459.

Carbone, E., & Hey, J. D. (2000). Which error story is best?. *Journal of Risk and Uncertainty 20*(2), 161–176.

Edwards, W. (1962). Subjective probabilities inferred from decisions. *Psychological Review, 69*, 109–135.

Guilford, J. P. (1954). *Psychometric methods* (2nd ed.). NY: McGraw-Hill.

Humphrey, S. J. (1995). Regret aversion or event-splitting effects? More evidence under risk and uncertainty. *Journal of risk and uncertainty, 11*, 263–274.

Kahneman, D. (2003). Experiences of collaborative research. *American Psychologist, 58*(9), 723–730.

Kahneman, D., & Tversky, A. (1979). Prospect theory: an analysis of decision under risk. *Econometrica, 47*, 263–291.

Lattimore, P. K., Baker, J. R., & Witte, A. D. (1992). The influence of probability on risky choice. *Journal of Economic Behavior and Organization, 17*, 377–400.

Lopes, L. L., & Oden, G. C. (1999). The role of aspiration level in risky choice: a comparison of cumulative prospect theory and SP/A theory. *Journal of Mathematical Psychology, 43*, 286–313.

Luce, R. D. (1959). *Individual choice behavior*. New York: Wiley.

Luce, R. D. (1994). Thurstone and sensory scaling: Then and now. *Psychological Review, 101*, 271–277.

Luce, R. D. (2000). *Utility of gains and losses: measurement-theoretical and experimental approaches*. Mahwah, NJ: Lawrence Erlbaum Associates.

Luce, R. D., & Marley, A. A. J. (2005). Ranked additive utility representations of gambles: old and new axiomatizations. *Journal of Risk and Uncertainty, 30*, 21–62.

Luce, R. D., Ng, C. T., Marley, A. A. J., & Aczél, J. (2006). Utility of Gambling: Entropy-modified linear weighted utility. Available from R. Duncan Luce, Mathematical Social Science, University of California, Irvine, CA 92717 (in preparation).

Luce, R. D., & Supes, P. (1965). Preference, utility, and subjective probability. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (pp. 249–410). New York: Wiley.

Marley, A. A. J., & Luce, R. D. (2001). Rank-weighted utilities and qualitative convolution. *Journal of Risk and Uncertainty, 23*(2), 135–163.

Marley, A. A. J., & Luce, R. D. (2005). Independence properties vis-à-vis several utility representations. *Theory and Decision, 58*, 77–143.

Martin, T. (1998) Comparing rank dependent, subjective weight, and configural weight utility models: Transitivity, monotonicity, coalescing, stochastic dominance, and event splitting independence. Master's Thesis, California State University, Fullerton.

Meginniss, J. R. (1976). A new class of symmetric utility rules for gambles, subjective marginal probability functions, and a generalized Bayes' rule. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 471–476.

Payne, J. W. (2005). It is whether you win or lose: The importance of the overall probabilities of winning or losing in risky choice. *Journal of Risk and Uncertainty, 30*, 5–19.

Prelec, D. (1998). The probability weighting function. *Econometrica, 66*, 497–527.

Quiggin, J. (1985). Subjective utility, anticipated utility, and the Allais paradox. *Organizational Behavior and Human Decision Processes, 35*, 94–101.

Quiggin, J. (1993). *Generalized expected utility theory: The rank-dependent model*. Boston: Kluwer.

Savage, L. J. (1954). *The foundations of statistics*. New York: Wiley.

Starmer, C. (2000). Developments in non-expected utility theory: the hunt for a descriptive theory of choice under risk. *Journal of Economic Literature, 38*, 332–382.

Starmer, C., & Sugden, R. (1993). Testing for juxtaposition and event-splitting effects. *Journal of Risk and Uncertainty, 6*, 235–254.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review, 34*, 273–286.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Tversky, A., & Wakker, P. (1995). Risk attitudes and decision weights. *Econometrica, 63*, 1255–1280.

Viscusi, K. W. (1989). Prospective reference theory: toward an explanation of the paradoxes. *Journal of risk and uncertainty, 2*, 235–264.

Wakker, P. (2001). Testing and characterizing properties of nonadditive measures through violations of the sure-thing principle. *Econometrica, 69*, 1039–1075.

Wu, G. (1994). An empirical test of ordinal independence. *Journal of Risk and Uncertainty, 9*, 39–60.

Wu, G., & Gonzalez, R. (1998). Common consequence conditions in decision making under risk. *Journal of Risk and Uncertainty, 16*, 115–139.

Wu, G., Zhang, J., & Gonzalez, R. (2004). Decision under risk. In D. Koehler & N. Harvey (Eds.), *Blackwell handbook of judgment and decision making* (pp. 399–423). Oxford: Oxford University Press.

Yang, J., & Qiu, W. (2005). A measure of risk and a decision-making model based on expected utility and entropy. *European Journal of Operational Research, 164*, 792–799.