Journal of Mathematical Economics 45 (2009) 329–340

Contents lists available at ScienceDirect

Journal of Mathematical Economics

journal homepage: www.elsevier.com/locate/jmateco

Preference for equivalent random variables: A price for unbounded utilities

Teddy Seidenfeld*, Mark J. Schervish, Joseph B. Kadane

Carnegie Mellon University, Statistics, Pittsburgh, PA 15213, United States

ARTICLE INFO

Article history:
Received 25 October 2006
Received in revised form 6 November 2008
Accepted 17 December 2008
Available online 4 February 2009

Keywords:
Unbounded utilities
Equivalent variables
Coherent previsions
St. Petersburg paradox
Non-Archimedean preferences

ABSTRACT

Savage's expected utility theory orders acts by the expectation of the utility function for outcomes over states. Therefore, preference between acts depends only on the utilities for outcomes and the probability distribution of states. When acts have more than finitely many possible outcomes, then utility is bounded in Savage's theory. This paper explores consequences of allowing preferences over acts with unbounded utility. Under certain regularity assumptions about indifference, and in order to respect (uniform) strict dominance between acts, there will be a strict preference between some pairs of acts that have the same distribution of outcomes. Consequently in these cases, preference is not a function of utility and probability alone.

1. Introduction

Our paper explores sufficient conditions under which preferences over unbounded variables preclude indifference between some pairs with the same distributions over outcomes (i.e., pairs of equivalent variables). In order to capture a broad range of such sufficient conditions, we employ a rich measure space, as defined below. That is, in order to formulate a wide variety of circumstances in which our results obtain, we make the following assumptions.

Consider a measure space $\mathcal{M} < \Omega, \mathcal{B}, \mathbf{P} >$, where $\mathbf{P}$ is a countably additive probability. (In Section 4 we generalize our theory to accommodate merely finitely additive probability.) Let $\mathcal{X}$ be a class of M-measurable, real-valued variables. Hereafter, we assume that the class $\mathcal{X}$ contains the constant function $\mathbf{1}$ and is closed under linear operations, i.e., if X and Y belong to $\mathcal{X}$, so too does $\mathbf{a}X+\mathbf{b}Y$, with $\mathbf{a}$ and $\mathbf{b}$ real numbers. This condition insures that the linear span of each finite subset of variables belonging to $\mathcal{X}$ also belongs to $\mathcal{X}$. We assume that $\mathcal{M}$ is sufficiently rich (and Ω is a sufficiently large set) that $\mathcal{B}$ contains various denumerable partitions of Ω, an instance of which we denote by $\prod = \{\pi_n: n = 1, \ldots\}$. Where needed, we assume further that there exist M-measurable partitions $\prod$ with specified geometric distributions, e.g., for $0 < \mathbf{p} < 1$, $\mathbf{P}(\pi_n) = \mathbf{p}(1-\mathbf{p})^{n-1}$. Last, for demonstrating several of our results we assume a discrete random variable with a distribution independent of a set of variables defined on a particular denumerable partition $\prod$. That is, given a denumerable partition $\prod$ and a class of variables defined on $\prod$, we sometimes assume there exists a "randomizer" with respect to these variables.

Definition 1. Variables X and Y are *equivalent*, denoted $X \equiv Y$, provided that for each interval I, $\mathbf{P}(X \in I) = \mathbf{P}(Y \in I)$.

Our investigation focuses on preference among equivalent variables. We assume a preference order over elements of $\mathcal{X}$.

* Corresponding author. Tel.: +1 412 268 2209; fax: +1 412 268 1440.
E-mail address: teddy@stat.cmu.edu (T. Seidenfeld).

0304-4068/$ – see front matter
doi:10.1016/j.jmateco.2008.12.002

- Let $\prec$ be a binary, strict (weak order) preference relation defined over $\mathcal{X} \times \mathcal{X}$, i.e., $\prec$ is asymmetric, and negatively transitive: if not $X \prec Y$ and not $Y \prec Z$ then not $X \prec Z$. Denote by $\sim$ the induced, transitive indifference relation. That is, $X \sim Y$ if and only if neither $X \prec Y$ nor $Y \prec X$.

Definition 2. A variable is *simple* if, with **P**-probability 1, it takes only finitely many values.

When a real-valued utility function for outcomes, $\boldsymbol{U}(X(\omega))$, is bounded over Ω and $\mathcal{X}$, or when all variables in $\mathcal{X}$ are simple, it is consistent with an expected utility theory of preference that the preference order is defined over the equivalence classes of equivalent variables. (See, for example, Fishburn (1979) chapters 8 and 10.) That is, under such circumstances two equivalent variables occupy the same place in the preference order. In this paper, we establish conditions when, using unbounded utilities, preferences over non-simple variables cannot preserve indifference between all pairs of equivalent variables. Then, the preference relation $\prec$ cannot be represented as a function of the probability distribution **P** over Ω and a real-valued utility function $\boldsymbol{U}(X(\omega))$ defined over values of variables. To simplify our presentation, and assuming that utility is measurable, we take the values of variables to be their utilities, $X(\omega) = \boldsymbol{U}(X(\omega))$. Thus, since each real-valued constant belongs to $\mathcal{X}$, $\boldsymbol{U}$ is unbounded both below and above.

We introduce two further requirements that, together with the assumption that preference is a weak order, constitute what we mean by a *coherent preference*.

- *Coherent Indifference*: If $X \sim Y$ then $(X - Y) \sim (Y - X) \sim \boldsymbol{0}$.

This requirement expresses the idea that when variables X and Y are indifferent there is no value in selling one for gaining the other. Such a trade is judged indifferent with the status-quo, which we represent as the constant function $\boldsymbol{0}$.

Definition 3. Variable Y (uniformly strictly) *dominates* variable X if, for some $\varepsilon > 0$ and for each ω, $Y(\omega) - X(\omega) \geq \varepsilon$.

- *Coherent Strict Preference*: If Y dominates X, then $X \prec Y$.

Definition 4. A weak order preference relation is *coherent* when it satisfies both the Coherent Indifference and the Coherent Strict Preference conditions, above.

Note that in the condition of Coherent Strict Preference the dominance relation is required to hold in the finest partition of Ω, the partition of Ω by its singleton states, $\{\omega\}$. We do not require that dominance with respect to an arbitrary M-measurable partition fixes strict preference, unless that dominance obtains also in the privileged partition of Ω by states. Also, we require for dominance that the difference between the two variables is bounded away from 0. These restrictions are designed to accommodate coherent preference based on a more general theory of a finitely additive probability space, rather than a countably additive probability space. We discuss this extension in Section 4. Since our purpose in this essay is to establish sufficient conditions under which coherent preference may not preserve indifference between all pairs of equivalent variables, our results are strengthened by using a limited dominance condition, particularly one that is consistent with a probability that is finitely but not necessarily countably additive.

For our investigation of preference over unbounded variables, we avoid assuming that coherent preference admits a real-valued representation. When preference does not admit a real-valued representation, e.g., because some variables have "infinite" values as with the St. Petersburg variable, still we need to be able to distinguish by strict preference between two "infinite" valued variables in case one dominates another. (See Colyvan (2008) for related considerations.) And we require a notion of indifference between variables that can be used with equivalent variables even when each has an "infinite" value. Coherent preference, as defined above, is consistent with a non-Archimedean preference relation, as explained below, in Section 2.

Our results about the impossibility of indifference between all pairs of equivalent variables take the following general form. When $\mathcal{X}$ includes unbounded variables, we provide sufficient conditions for the existence of a finite set of pairwise equivalent variables, $\{X \equiv Y_1 \equiv \ldots \equiv Y_k\}$ such that the variable $(\sum_i Y_i - kX)$ is strictly preferred to $\boldsymbol{0}$. Hence, if preference is coherent, it cannot be that equivalent variables are pairwise indifferent. This is because, if $\boldsymbol{0} \sim (Y_1 - X) \sim (X - Y_2)$ then by the criterion of coherent indifference, also $\boldsymbol{0} \sim (Y_1 - X) - (X - Y_2)$, which by finite iteration leads to the conclusion that $\boldsymbol{0} \sim (\sum_i Y_i - kX)$. But the equivalent variables we consider also result in a situation where, because of dominance, coherent preference requires that $\boldsymbol{0} \prec (\sum_i Y_i - kX)$. There are two contexts for this construction, involving (Case 1) non-Archimedean and (Case 2) discontinuous coherent preferences, each of which we describe in Section 2.

2. Coherent preferences for unbounded variables

There are two cases of coherent preferences for unbounded variables relevant to our analysis of preference over equivalent variable:

Case 1. First we consider a coherent preference order $\prec$ that mandates "infinite" values for some variables, e.g., the St. Petersburg variable, W, where $\mathbf{P}(W = 2^n) = 2^{-n}$. Such an order fails the von Neumann–Morgenstern Archimedean Axiom as adapted to coherent preferences over variables:

- *Archimedean Axiom*: If $X \prec Y \prec Z$, then there exist $0 < \mathbf{a}, \mathbf{b} < 1$ such that $\mathbf{a}X + (1-\mathbf{a})Z \prec Y \prec \mathbf{b}X + (1-\mathbf{b})Z$.

Let $X = \boldsymbol{0}$, $Y = \boldsymbol{1}$, and Z be a variable with "infinite" value. Then, as is well known, there is no real number $\mathbf{a}$ $0 < \mathbf{a} < 1$ such that $\mathbf{a}\boldsymbol{0} + (1-\mathbf{a})Z \prec \boldsymbol{1}$.

Case 2. A coherent (possibly) Archimedean preference order may fail to be continuous from below. A criterion of continuity used, e.g., in the definition of the Lebesgue integral for unbounded functions—see Royden [1968, p. 226], adapted to coherent preference is this:

- *Continuity principle*: Let X be a non-negative variable and suppose that $X \sim \boldsymbol{k}$ for some real-valued, constant outcome $\boldsymbol{k}$. Let $\{X_n\}$ be a sequence of non-negative variables converging (pointwise over Ω) from below to the variable X. That is, for each state ω, $X_n(\omega) \le X(\omega)$ $(n = 1, \ldots)$ and $lim_n X_n = X$. If $X_n \sim \boldsymbol{k_n}$ with $\{\boldsymbol{k}_n : n = 1, \ldots\}$ a sequence of real-valued constant outcomes, then $lim_n \boldsymbol{k_n} = \boldsymbol{k}$.

Wakker [1993, Lemma 1.8] establishes the similar property of "truncation continuity" of finite Choquet integrals for unbounded variables. In the balance of this section we illustrate how to define coherent preference in each of these two cases.

2.1. Case 1: Coherent preferences that cannot be represented by real values

This is the more familiar of the two situations with unbounded quantities, which we illustrate with the St. Petersburg variable.

Example 2.1. Let N be a random variable with a Geometric($\mathbf{p}$) distribution. For each positive integer n, let $\pi_n = \{N = n\}$, the event that $N = n$. Let $\mathcal{B}$ be the sigma field generated by N. Then $\mathbf{P}(\pi_n) = \mathbf{p}(1-\mathbf{p})^{n-1}$, for $n = 1, \ldots$. Suppose $\mathcal{X}$ contains all the bounded variables that are $\mathcal{B}$ measurable. It is an elementary result of expected utility theory that ordering $\mathcal{X}$ by the $\mathbf{P}$-expected value, $\mathbf{E_P}[X]$, of its elements yields a coherent preference ordering. Denote this strict preference $\prec$. Let the *value* $\boldsymbol{V}$ of the (bounded) variable X be its $\mathbf{P}$-expected value, $\boldsymbol{V}(X) = \mathbf{E_P}[X]$, and then $\boldsymbol{V}$ represents $\prec$ over $\mathcal{X}$. That is, for X and Y elements of $\mathcal{X}$, $\boldsymbol{V}(X) < \boldsymbol{V}(Y)$ if and only if $X \prec Y$.

For simplicity, let $\mathbf{p} = 1/2$. Define the unbounded, discrete variable $W = 2^N$, the St. Petersburg variable, and extend $\mathcal{X}$ to $\mathcal{X}^*$ by adding W and closing the class under linear combinations. If $\prec^*$ is a coherent order over $\mathcal{X}^*$ that extends $\prec$, then $\boldsymbol{V}^*(W) > \mathbf{r}$ for each real number $\mathbf{r}$, and $\boldsymbol{V}^*$ is not real-valued. Thus, as is well known, $\prec^*$ fails the Archimedean axiom. Nonetheless, there is a coherent preference order over $\mathcal{X}^*$ based on the following lexicography. (See Hausner (1954) for a general theory of non-Archimedean preferences over variables that obey the "Independence" axiom of von Neumann–Morgenstern theory.)

Write $X^* \in \mathcal{X}^*$ as $X^* = (\mathbf{a}W + \mathbf{b}X)$ with $X \in \mathcal{X}$ and real numbers $\mathbf{a}$ and $\mathbf{b}$. Then, define the weak order $\prec^*$ by:

$X_1^* \prec^* X_2^*$ if and only if either $\mathbf{a}_1 < \mathbf{a}_2$, or else $\mathbf{a}_1 = \mathbf{a}_2$ and $\mathbf{b}_1\boldsymbol{V}(X_1) < \mathbf{b}_2\boldsymbol{V}(X_2)$.

That $\prec^*$ is coherent follows because—

- For the first criterion: If $X_1^* \sim X_2^*$ then $\mathbf{a}_1 = \mathbf{a}_2$ and $\mathbf{b}_1\boldsymbol{V}(X_1) = \mathbf{b}_2\boldsymbol{V}(X_2)$. Hence, denoting $(X_1^* - X_2^*)$ by X_3^*, we see that $X_3^* \in \mathcal{X}$ as $\mathbf{a}_3 = (\mathbf{a}_1 - \mathbf{a}_2) = 0$. But $\boldsymbol{V}[\mathbf{b}_1(X_1) - \mathbf{b}(X_2)] = \mathbf{b}_1\boldsymbol{V}(X_1) - \mathbf{b}_2\boldsymbol{V}(X_2) = 0$. Therefore $(X_1^* - X_2^*) \sim^* \mathbf{0}$, as required.
- For the second criterion: If X_2^* dominates X_1^* then, as the dominance must hold on all states in the "tail" of the Geometric($\mathbf{p}$) where W is unbounded, $\mathbf{a}_1 < \mathbf{a}_2$ and therefore $X_1^* \prec^* X_2^*$ as required.

Next, we introduce the class of generalized St. Petersburg variables that we use in Section 3. Let m be a nonnegative integer, and let N have the Geometric$(1-2^{-m})$ distribution. Let $\mathbf{p} = 1 - 2^{-m}$, and define $\mathbf{q} = (\mathbf{1} - \mathbf{p})/\mathbf{p}$. Let U be independent of N and have the Bernoulli($\mathbf{q}$) distribution. Define π_{ni}: $n = 1, \ldots$; $i = 0, 1$ by $\pi_{ni} = \{N = n\} \cap \{U = i\}$. Then, $\mathbf{P}(\pi_{n1} \cup \pi_{n2}) = \mathbf{p}(1-\mathbf{p})^{n-1}$, with $\mathbf{P}(\pi_{n1}) = (1-\mathbf{p})^n$ and $\mathbf{P}(\pi_{n0}) = (2\mathbf{p}-1)(1-\mathbf{p})^{n-1}$.

Definition 5. The *generalized St. Petersburg(*$\boldsymbol{p}$*) variable*, W_m is constant on each element of the partition $\prod = \{\pi_{ni} : n = 1, \ldots;$ $i = 1, 0\}$ with values $W_m(\omega) = (1-\mathbf{p})^{-n}$ for ω in π_{n1} and $W_m(\omega) = 0$ for ω in π_{n0}.

Note, for $m = 1$ then W_1 is the familiar St. Petersburg variable, W, from Example 2.1.

With Theorem 1 (formulated in Section 3) we establish that, given $\mathbf{p}$, coherent preferences cannot preserve indifference between all pairs in a specific finite set of variables, each of which is equivalent to a generalized St. Petersburg($\mathbf{p}$) variable.

2.2. Case 2: A coherent, Archimedean preference ordering that, though represented by real values, is not given by the expected values of its (non-simple) variables

Consider the class $\mathcal{X}$ of all the bounded variables and let $\boldsymbol{V}(X) = \mathbf{E_P}[X]$. Define a coherent preference order $\prec$ in terms of values of $\boldsymbol{V}$. Let Z be an unbounded random variable, bounded below, and with finite expectation $-\infty < \mathbf{E_P}[Z] < \infty$. Extend $\mathcal{X}$ to

$\mathcal{X}^*$ by including Z and closing under linear combinations. As de Finetti has argued [1974, p. 131], the preference order $\prec$ may be extended to an order $\prec^*$ over $\mathcal{X}^*$ that respects the *dominance* criterion if and only if Z is assigned an extended real value $\boldsymbol{V}^*(Z)$ at least as great as its expectation. Choose $\boldsymbol{V}^*(Z) = \mathbf{E_P}[Z] + \boldsymbol{\beta}(Z)$, with $0 \le \boldsymbol{\beta}(Z) < \infty$. We let $\boldsymbol{\beta}(Z)$ denote the real-valued ***boost*** that Z receives in excess of its expected value.

Aside: It is coherent, also, to let $\boldsymbol{\beta}(Z) = \infty$, which results in a discontinuous, non-Archimedean preference order. We do not investigate such orders in this paper.

More generally, let $\prec$ be a coherent Archimedean order over the class $\mathcal{Y}$ of random variable with finite absolute $\mathbf{P}$-expectation. That is, for $Y \in \mathcal{Y}$, $\mathbf{E_P}[|Y|] < \infty$. Let $\boldsymbol{V}$ be a real-valued representation of this order. The boost function $\boldsymbol{\beta}(\cdot)$ is defined as follows.

Definition 6. $\boldsymbol{\beta}(Y) = \boldsymbol{V}(Y) - \mathbf{E_P}[Y]$.

It is straightforward to show that over the class $\mathcal{Y}$ the boost function $\boldsymbol{\beta}(\cdot)$ is a finitely additive linear operator (Dunford and Schwartz, 1988, p. 36) that has the value 0 on all bounded variables. That is, $\boldsymbol{\beta}(X+Y) = \boldsymbol{\beta}(X) + \boldsymbol{\beta}(Y)$, $\boldsymbol{\beta}(\boldsymbol{a}Y) = \boldsymbol{a}\boldsymbol{\beta}(Y)$, and when X is a bounded variable then $\boldsymbol{\beta}(X) = 0$. Since, by de Finetti's result, in order to respect dominance the boost function for variables bounded below is non-negative, it follows that the boost is non-positive for variables bounded above.

The next example illustrates a coherent preference order that involves positive boost for an unbounded geometric variable.

Example 2.2. Let Z have the Geometric($\mathbf{p}$) distribution, and define $\pi_n = \{Z = n\}$ for $n = 1, \ldots$ so that $\mathbf{P}(\pi_n) = \mathbf{p}(1-\mathbf{p})^{n-1}$ for $n = 1, \ldots$. Let $\prod$ denote the partition $\{\pi_n\colon n = 1, \ldots\}$. Let $\mathcal{X}$ contain all the bounded $\prod$-measureable variables and let $\boldsymbol{V}(X) = \mathbf{E_P}[X]$. Extend $\mathcal{X}$ to $\mathcal{X}^*$ by including the unbounded, discrete variable Z with $\mathbf{E_P}[Z] = \mathbf{p}^{-1}$. Close $\mathcal{X}^*$ under linear combinations, which entails that all variables in $\mathcal{X}^*$ have finite absolute expectations. Choose a finite positive boost for Z, $\boldsymbol{\beta}(Z) = \mathbf{b} > 0$. Then, define $\boldsymbol{V}^*(Z) = \mathbf{E_P}[Z] + \mathbf{b}$, and generally, for $X^* \in \mathcal{X}^*$ where $X^* = \mathbf{a}Z + \mathbf{c}X$, with $X \in \mathcal{X}$, define $\boldsymbol{V}^*(X^*) = \boldsymbol{V}^*(\mathbf{a}Z + \mathbf{c}X) = \mathbf{a}\boldsymbol{V}^*(Z) + \mathbf{c}\boldsymbol{V}(X) = \boldsymbol{V}(X^*) + \mathbf{ab}$.

The resulting preference order, $\prec^*$, is Archimedean because it is represented by a real-valued function $\boldsymbol{V}^*(\cdot)$ that is a linear operator on elements of $\mathcal{X}^*$. Since $0 < \mathbf{b}$, $\prec^*$ is discontinuous as Z is the pointwise limit (from below) of bounded variables $Z_n \in \mathcal{X}$. But $lim_{n\to\infty}$ $\boldsymbol{V}^*(Z_n) = \mathbf{E_P}[Z] < \boldsymbol{V}^*(Z)$. Nonetheless, $\prec^*$ is coherent. This follows because:

- If $X_1^* \sim^* X_2^*$ then $\boldsymbol{V}^*(X_1^*) = \boldsymbol{V}^*(X_2^*)$ and $\boldsymbol{V}^*(X_1^* - X_2^*) = 0$.
- If X_2^* dominates X_1^* then evidently $\mathbf{E_P}[X_1^*] < \mathbf{E_P}[X_2^*]$ and also the dominance must hold on all states in the "tail" of the Geometric($\mathbf{p}$) distribution, where Z is unbounded. Hence, $\mathbf{a}_1 < \mathbf{a}_2$ and therefore $\boldsymbol{V}^*(X_1^*) < \boldsymbol{V}^*(X_2^*)$ as required.

In Section 3.2, with Theorem 2 we establish that unless the boost function is identically 0 on all Geometric($\mathbf{p}$) variables, and therefore unless preference is continuous from below for all variables (bounded below) whose tail is "thin" relative to some Geometric($\mathbf{p}$) variable, a coherent strict preference obtains between some pair of equivalent variables.

3. Mandatory strict preference between some equivalent variables

We formulate our results when coherent preference cannot preserve indifference between all pairs of equivalent variables separately for Case 1 and Case 2.

Regarding the former case, we have the following result:

Theorem 1. *Let $\prec$ be a coherent preference order over $\mathcal{X}$, and let $\mathbf{p} = \mathbf{1} - \mathbf{2}^{-m}$.*
Assume that there is at least one Geometric($\mathbf{p}$) variable Y and that there exists a random variable T with the uniform distribution on the interval [0,1] *that is independent of Y. Then some pair of equivalent variables are not indifferent.*

The condition in Theorem 1, that there exists a uniform random variable on the interval [0,1] that is independent of Y could be replaced by a slightly weaker condition requiring the existence of a large collection of discrete random variables that are independent of Y, but the precise statement of such a condition would be more complicated than the added generality justifies.

With coherent preference that is discontinuous from below, we have the following result.

Theorem 2. *Let $\prec$ be a coherent preference order over the class $\mathcal{Y}$ of variables with finite absolute* $\mathbf{P}$*-expectations. Assume that there exist at least two independent Geometric($\mathbf{p}$) variables Y_1 and Y_2, and assume that $\boldsymbol{\beta}(Y_i) > 0$ for at least one of $i = 1$ or 2. Then some pair of equivalent variables are not indifferent.*

Proofs of the two theorems are given in Appendix A and Appendix B; however, in Subsections 3.1 and 3.2 we provide elementary illustrations.

Table 1
Displays the values of the equivalent variables, X_1, X_2, and W, used in Example 3.1.

	π_1	$\pi_2\ldots$	$\pi_n\ldots$
B	$W=2$ $X_1=4$ $X_2=2$	$W=4$ $X_1=8$ $X_2=2$	$W=2^n$ $X_1=2^{n+1}$ $X_2=2$
B^c	$W=2$ $X_1=2$ $X_2=4$	$W=4$ $X_1=2$ $X_2=8$	$W=2^n$ $X_1=2$ $X_2=2^{n+1}$

3.1. Strict preferences among generalized St. Petersburg($\boldsymbol{p}$) variables

In Theorem 1, for each Geometric($1-2^{-m}$) distribution ($m=2,3,\ldots$), we construct a set of 2^{m-1} equivalent variables, $X_1 \equiv X_2 \equiv \ldots \equiv X_{2^{m-1}}$, each one equivalent to a generalized St. Petersburg($1-2^{-m}$) variable, W_m, such that:

$$\sum_{i=1}^{2^{m-1}}(X_i - W_m) \sim 2^{m-1}. \quad (*)$$

That is, though the X_i ($i=1,\ldots,2^{m-1}$) are pairwise equivalent variables, and equivalent to W_m, their pairwise differences with W_m cannot all be indifferent with $\boldsymbol{0}$.

Example 3.1 is a simplified version of the construction for the St. Petersburg(1/2) variable (Table 1).

Example 3.1. Let N have the Geometric(1/2) distribution, and let $\pi_n=\{N=n\}$ for $n=1,2,\ldots$ Let Π be the partition $\{\pi_n$: $n=1,\ldots)$ so that, $\mathbf{P}(\pi_n)=2^{-n}$ ($n=1,2,\ldots$). Let U be independent of N with U having the Bernoulli(1/2) distribution. Partition each event π_n into two equi-probable states using the independent, probability 1/2 events $B=\{U=1\}$ and $B^{\mathbf{c}}=\{U=0\}$, and define three equivalent variables, X_1, X_2, and W as follows:

- the canonical St. Petersburg variable: $W=2^N$, which is independent of U.
- the variable $X_1(\omega)=2^{n+1}$ for $\omega \in B\cap\pi_n$ and $X_1(\omega)=2$ for $\omega\in B^{\mathbf{c}}\cap\pi_n$.
- the variable $X_2(\omega)=2$ for $\omega\in B\cap\pi_n$ and $X_2(\omega)=2^{n+1}$ for $\omega\in B^{\mathbf{c}}\cap\pi_n$.

Though $X_1\equiv X_2\equiv W$, for each $\omega\in\Omega$, $X_1(\omega)+X_2(\omega)-2W(\omega)=2>0$. By dominance, this contradicts the hypothesis that the difference between equivalent variables is indifferent to $\boldsymbol{0}$, which would entail that $(X_1+X_2-2W)\sim\boldsymbol{0}$.

3.2. Strict preference among equivalent variables whose values differ from their expectations

Consider a coherent, Archimedean order $\prec$ over the space of variables with finite absolute expectations, $\mathcal{Y}$. Let $\boldsymbol{V}(\cdot)$ be a real-valued representation of $\prec$, which for a bounded variable is its expectation. That is, $Y_1\prec Y_2$ if and only if $\boldsymbol{V}(Y_1)<\boldsymbol{V}(Y_2)$, and if X is bounded, $\boldsymbol{V}(X)=\mathbf{E_P}(X)$. Let Y have a Geometric($\mathbf{p}$) distribution, so it is bounded below. Assume that $\boldsymbol{V}(Y)$ is finite but greater than its expectation, $\mathbf{E_P}[Y]=\mathbf{p}^{-1}$. So $\boldsymbol{\beta}(Y)>0$. For Theorem 2, we show there exist equivalent variables W_1 and W_2, that cannot have the same $\boldsymbol{V}$-value. Example 3.2 illustrates a simplified version of this construction for the special case of the Geometric(1/2) distribution, where at least two (among three) equivalent variables cannot be indifferent (Table 2).

Example 3.2. Let Y be a Geometric(1/2) variable measurable, and define $\pi_n=\{Y=n\}$ for $n=1,2,\ldots$. Hence, $\mathbf{P}(\pi_n)=2^{-n}$, for $n=1,2,\ldots$, and $\mathbf{E_P}[Y]=2$. Let U be independent of Y with U having the Bernoulli(1/2) distribution. Let $B=\{U=1\}$ and $B^{\mathbf{c}}=\{U=0\}$, so that $\mathbf{P}(B\cap\pi_n)=\mathbf{P}(B^{\mathbf{c}}\cap\pi_n)=2^{-(n+1)}$, for $n=1,2,\ldots$. Define two other variables W_1 and W_2 as follows:

$W_1(\omega)=n+1$ for $\omega\in B\cap\pi_n$; $W_1(\omega)=1$ for $\omega\in B^{\mathbf{c}}\cap\pi_n$ ($n=1,2,\ldots$)and
$W_2(\omega)=1$ for $\omega\in B\cap\pi_n$; $W_2(\omega)=n+1$ for $\omega\in B^{\mathbf{c}}\cap\pi_n$ ($n=1,2,\ldots$).

Table 2
Displays the values of the equivalent variables, Y, W_1 and W_2, used in Example 3.2.

	π_1	$\pi_2\ldots$	$\pi_n\ldots$
B	$Y=1$ $W_1=2$ $W_2=1$	$Y=2$ $W_1=3$ $W_2=1$	$Y=n$ $W_1=n+1$ $W_2=1$
B^c	$Y=1$ $W_1=1$ $W_2=2$	$Y=2$ $W_1=1$ $W_2=3$	$Y=n$ $W_1=1$ $W_2=n+1$

Obviously, W_1 and W_2 are equivalent. Moreover, each has the Geometric(1/2) distribution; hence, $Y \equiv W_1 \equiv W_2$. However, for each $\omega \in \Omega$, $W_1(\omega) + W_2(\omega) - Y(\omega) = 2$.

Thus, $\mathbf{V}(W_1 - Y) + \mathbf{V}(W_2 - Y) = 0$ if and only if $\mathbf{V}(W_1) = \mathbf{V}(W_2) = \mathbf{V}(Y) = 2$. Then the value $\boldsymbol{V}$ for a Geometric(1/2) variable is its expectation, and $\boldsymbol{\beta}(W_1) = \boldsymbol{\beta}(W_2) = \boldsymbol{\beta}(Y) = 0$, and the boost function is 0 for each of these unbounded variables.

Theorem 2 has a Corollary relating to continuous preference orders, which we express in terms of tail-dominance between variables.

Definition 7. For variables X and Y, bounded below, measurable with respect to a denumerable partition $\Pi = \{\pi_1, \ldots\}$, Y *tail-dominates* X if, for some $\boldsymbol{k}$ and for all $n \geq \boldsymbol{k}$, $Y(\omega) \geq X(\omega)$ for all $\omega \in \pi_n$.

When Y tail dominates X, their difference $Y - X$ is bounded below. By de Finetti's (1974, p. 131) result, then $\boldsymbol{\beta}(Y - X) \geq 0$ and we have $\boldsymbol{\beta}(Y) \geq \boldsymbol{\beta}(X)$. Thus, Theorem 2 has this corollary.

Corollary. *Suppose that for a given value of* $\boldsymbol{p}$*, a coherent Archimedean preference order respects indifference between all pairs of equivalent Geometric(*$\boldsymbol{p}$*) variables. If variable X is tail-dominated by one of the Geometric(*$\boldsymbol{p}$*) variables, then preference for X is continuous from below.*

4. de Finetti's theory of coherent previsions

In this section we show that the finitely additive version of *coherent preference*, the variant of the theory from Section 1 modified to permit the use of merely finitely additive probabilities, applies to de Finetti's [1974] theory of *Coherent Previsions*. In de Finetti's theory, for each real-valued variable, $X \in \mathcal{X}$, measurable with respect to a common measurable space $< \Omega, \mathcal{B} >$, the decision maker has an extended real-valued prevision, **Prev**(X). He allows that, in particular, when X is unbounded its prevision may be infinite, negative or positive de Finetti's [1974, Sections 3.12.4 and 6.5.4–6.5.9].

When the prevision for X is real-valued, it is subject to a *two-sided*, real-valued payoff $\boldsymbol{c_X}(X - \mathbf{Prev}(X))$, where $\boldsymbol{c}_{\mathrm{X}}$ is a real number that depends upon X and **Prev**(X) and which is chosen by a rival gambler. The prevision is said to be *two-sided*, as $\boldsymbol{c}_{\mathrm{X}}$ may be chosen by the rival gambler either positive or negative (or 0), corresponding informally to the decision maker being required to buy or to sell the payoff X for the amount **Prev**(X), scaled by the magnitude $|\boldsymbol{c_X}|$. When $\boldsymbol{c_X} = 0$, there is no transaction involving X and the decision maker remains at her/his status-quo wealth, which is judged indifferent to a null-gain, $\boldsymbol{0}$. In short, the decision maker is committed to using **Prev**(X) as the "fair price" for buying or selling each unit of the quantity X as chosen by a rival.

When the prevision for X is infinite-positive, i.e., when X has a value to the decision maker greater than any finite amount, then we interpret de Finetti's theory of previsions to mean that for each real constant $\boldsymbol{k_X}$ and for each $\boldsymbol{c_X} > 0$ that may be chosen by the rival gambler, the decision maker is willing to accept (i.e., is committed to "buy") a *one-sided* payoff $\boldsymbol{c_X}(X - \boldsymbol{k_X})$. Likewise, when the prevision for X is infinite-negative, with value less than any finite amount, then for each real constant $\boldsymbol{k_X}$ and for $\boldsymbol{c_X} < 0$ that may be chosen by the rival gambler, the decision maker is willing to accept (i.e., is committed to sell) a *one-sided* payoff $\boldsymbol{c_X}(X - \boldsymbol{k_X})$.

In accord with de Finetti's theory, the decision maker is required to accept an arbitrary, finite sum of such real-valued payoffs as fixed by the rival gambler's selection of coefficients $\boldsymbol{c_X}$ and, where one-sided payoffs are involved, constants $\boldsymbol{k_X}$.

Definition 8. Previsions are de Finetti-***Coherent*** if there is no finite selection of non-zero constants, $\boldsymbol{c_X}$ (and where one-sided previsions are involved also constants $\boldsymbol{k_X}$) with the sum of the payoffs *uniformly* dominated by $\boldsymbol{0}$ in the partition Ω.

The previsions are (de Finetti-) ***Incoherent*** otherwise.

This criterion is related to "overtaking" between variables as used by Becker and Boyd [1997, p. 67] in intertemporal choice with unbounded variables.

> ***Theorem*** (de Finetti [1974, 3.10 & 3.12]): Previsions over the set of bounded variables are (de Finetti-) Coherent if and only if they are the expectations of some finitely additive probability **P** that makes $< \Omega, \mathcal{B}, \mathbf{P} >$ into a finitely additive measure space.

Note that when the variables in question are the indicator functions for events, then their coherent previsions are their probabilities under the finitely additive measure that satisfies the theorem above.

In order to allow that all finitely additive expectations are de Finetti-Coherent, it is necessary that:

(i) dominance is limited to uniform dominance, and
(ii) dominance is formulated with respect to a privileged partition, e.g., Ω.

To see why the first condition is necessary, consider this simple counter example. Let Ω be the positive integers and let $X(\omega) = -1/\omega$. Let **P** be any (purely) finitely additive probability with $\mathbf{P}(\omega) = 0$ for all ω. Then $\mathbf{E_P}(X) = 0$; hence, $X \sim \boldsymbol{0}$. Nonetheless, $\boldsymbol{0}$ dominates X in Ω, though not uniformly.

For motivating the second condition, observe that uniform dominance in a partition other than Ω may fail to determine even the ordinal relation of which of two previsions is greater. For example, let $\Omega = \{0,1\} \times \{1, 2, \ldots\}$. Name the events $B = \{1\} \times \{1, 2, \ldots\}$ and $\pi_n = \{(0,n),(1,n)\}$. Let $\mathbf{P}(B) = \mathbf{P}(B^{\mathbf{c}}) = 1/2$, $\mathbf{P}(B \cap \pi_n) = 2^{-(n+1)}$ and $\mathbf{P}(B^{\mathbf{c}} \cap \pi_n) = 0$, for $n = 1, \ldots$. That is,

$\mathbf{P}(\pi_n|B)=2^{-n}$ is a countably additive conditional distribution; whereas, $\mathbf{P}(\pi_n|B^{\mathbf{c}})=0$ for $n=1,\ldots$, is a purely finitely additive conditional distribution. Since $\mathbf{P}(\pi_n)=2^{-(n+1)}>0$, the conditional probability given each π_n also is well defined and satisfies: $\mathbf{P}(B|\pi_n)=1$, $n=1,\ldots$. Thus, even though $\mathbf{P}(B)=\mathbf{P}(B^{\mathbf{c}})$, with respect to the partition $\prod=\{\pi_1,\pi_2,\ldots\}$ the conditional probabilities satisfy: $\mathbf{P}(B|\pi_n)=\mathbf{P}(B^{\mathbf{c}}|\pi_n)+1$, $n=1,\ldots$. The conditional probability for B is greater than the conditional probability for $B^{\mathbf{c}}$ given each $\pi\in\prod$. Moreover, the differences are bounded away from 0. De Finetti [1974] calls this "non-conglomerability" of conditional probability. (See Kadane et al. (1986) for additional discussion.)

These probabilities and conditional probabilities are the values of the respective de Finetti previsions, and conditional previsions given $\prod$. Thus, $\mathbf{Prev}(B-B^{\mathbf{c}}|\pi_n)=1$ $(n=1,\ldots)$ despite the fact that $\mathbf{Prev}(B)=\mathbf{Prev}(B^{\mathbf{c}})=1/2$. But note that the uniform dominance of $(B-B^{\mathbf{c}})$ over $\boldsymbol{0}$ for the conditional prevision, given $\prod$, is not duplicated in the privileged partition by elements of Ω, where $\mathbf{E_P}[B-B^{\mathbf{c}}|\omega)=1$ or -1 according as $\omega\in B$ or $B^{\mathbf{c}}$. It is an elementary fact of finitely additive probabilities that always they are "conglomerable" in the privileged partition of Ω by it elements, the partition comprised by the states of the measure space.

Next, consider a measurable space $<\Omega,\mathcal{B}>$ and a class $\mathcal{X}$ of variables over which de-Finetti-Coherent previsions are given. We define a finitely additive coherent preference order $\prec^*$ over $\mathcal{X}$ based on the prevision function $\mathbf{Prev}(\cdot)$. As before, we assume that $\mathcal{X}$ is closed under linear spans for each finite subset of variables, and contains the constant $\boldsymbol{1}$.

Definition 9. $X\prec^* Y$ if and only if $\mathbf{Prev}(Y-X)>0$.

- ***Proposition***: The preference relation $\prec^*$ is a coherent weak order if the extended real-value previsions are de Finetti-Coherent.

Proof. For the class of simple variables the Proposition, and more, is immediate from de Finetti's ([1972, section 5.9] or [1974, section 3.10]) principal result about the existence of coherent (real-valued) previsions. Specifically, de Finetti shows that for a constant c, $\mathbf{Prev}(\boldsymbol{c})=c$; for variables X and Y in $\mathcal{X}$, $\mathbf{Prev}(X+Y)=\mathbf{Prev}(X)+\mathbf{Prev}(Y)$; and if $\mathbf{Prev}(X)=0$, then $\boldsymbol{0}$ does not uniformly dominate X in the partition Ω. His reasoning extends to bounded variables, as de Finetti notes [1972, section 5.33].

To show that the Proposition holds for classes of unbounded variables, where previsions might be infinite, assume that extended-valued de Finetti-Coherent previsions exist over $\mathcal{X}$. (We do not provide the existence proof here.) If $\mathbf{Prev}(Y-X)\le 0$ and $\mathbf{Prev}(Z-Y)\le 0$, then by de Finetti-Coherence of previsions, $\mathbf{Prev}(Z-X)=\mathbf{Prev}(Z-Y+Y-X)\le 0+0=0$, and $\prec^*$ is negatively transitive. If $\mathbf{Prev}(X-Z)=0$, then since real-valued previsions are two-sided, $\mathbf{Prev}(Z-X)=\mathbf{Prev}(-[X-Z])=-0=0$, and $\prec^*$ satisfies the criterion of Coherent Indifference. Last, assume that Y (uniformly) dominates X in the partition Ω. Then there exists $\varepsilon>0$ such that, for each $\omega\in\Omega$, $X(\omega)+\varepsilon\le Y(\omega)$. By de Finetti-Coherence, then $\mathbf{Prev}(Y-X)\ge\varepsilon>0$, so $X\prec^* Y$, as required by the criterion of Coherent Strict Preference. □

The theory of coherent preferences for finitely additive measure spaces is more general than de Finetti's theory of (de Finetti-) Coherent previsions. This can be seen from the fact that our account of coherent preference does not require an Archimedean order even over the class of bounded variables. That is, a coherent preference over the class of bounded variables may fail to have a real-valued representation; however, de Finetti-coherent previsions are real-valued for this same class. Though this aspect of our theory is not relevant to the two Theorems of Section 3, it is important for the development of conditional preference given null events.

It is old news that within de Finetti's theory, coherent conditional previsions given a null event cannot be defined from the (unconditional) coherent previsions using the device of called-off gambles. (See, e.g., Levi [1980, chapter 5].) We conjecture that, in our framework, coherent conditional preferences given a null event may be defined from a coherent, non-Archimedean preference order. (For related discussion involving one-sided conditional previsions, see Troffaes (2006) and the references given there.) Nonetheless, as the results of Section 3 apply to unconditional preferences, those findings stand whether or not this conjecture is accurate.

5. Conclusions and further questions

We have shown that coherent preference orderings over unbounded variables cannot satisfy indifference between pairs of equivalent quantities when either:

(i) the preference order is non-Archimedean as a result of including, e.g., St. Petersburg variables, or
(ii) the preference ordering, though Archimedean, is not continuous (from below) as a result of a positive "boost" for some variable, bounded below, that is tail-dominated by a geometric distribution.

These results conflict with the usual approach to theories of Subjective Expected Utility, such as Savage's theory [1972], where preference is defined over the equivalence classes of equivalent lotteries. The contrast with de Finetti's theory is a subtle one, however.

Like de Finetti's theory, Savage's theory permits merely finitely additive personal probability, i.e., preference in Savage's theory is not required to be continuous in the sense that we use here. But in contrast with de Finetti's theory, in Savage's theory the problems with unbounded variables discussed in this paper are sidestepped entirely. In his theory given by seven postulates, P1-P7, utility is bounded. (See Savage [1972, p. 80].) If the theory comprised by Savage's postulates P1-P6 is considered, instead, the resulting weakened theory admits unbounded utility and an expected utility representation for preference over simple lotteries. But it does not ensure an expected utility representation for preference over non-simple lotteries, even when variables are bounded. Nor does the theory P1-P6 entail that uniform dominance is reflected in strict preference. (See Seidenfeld and Schervish (1983) for details.)

We understand the results of this paper as pointing to the need for developing a normative theory that fits between Savage's P1-P7 and de Finetti's theory of coherent previsions. The former is overly restrictive, we think, in requiring that utility is bounded. The latter is overly generous in allowing finite but discontinuous previsions for unbounded quantities, even when all bounded quantities have continuous previsions and probability is countably additive. We hope to find a theory that navigates satisfactorily between these two landmarks.

Acknowledgements

We thank Fabio Cozman and Matthias Troffaes for helpful discussions. Also, we thank the participants at the *8th Brazilian Meeting on Bayesian Statistics* (March 2006), the participants at the University of Pittsburgh Center for Philosophy of Science's Workshop *Bayesianism, Fundamentally* (October 2006), and an anonymous referee for helpful comments. Seidenfeld gratefully acknowledges support by the State of Sao Paulo, Brazil, grant FAPESP #05/59541-4, during his stay at the University of Sao Paulo, Spring term 2006.

Appendix A. Proof of Theorem 1

Let $\mathbf{p} = \mathbf{1} - \boldsymbol{2}^{-m}$, and let Y be a Geometric($\mathbf{p}$) variable as assumed in the statement of Theorem 1. The proof given here works for arbitrary m. We construct a generalized St. Petersburg variable W_m and another 2^{m-1} equivalent variables $X_1 \equiv X_2 \equiv \ldots \equiv X_{2^{m-1}} (\equiv W_m)$, such that, if preferences are coherent and pairwise differences between equivalent variables are indifferent to $\boldsymbol{0}$, then we obtain the contradiction:

$$\sum_{i=1}^{2^{m-1}} (X_i - W_m) \sim 2^{m-1}. \quad (*)$$

Fix $m \geq 2$. For each n define $\pi_n = \{Y = n\}$ so that:

$$\mathrm{P}(\pi_n) = \mathbf{p}(1 - \mathbf{p})^{n-1}.$$

Let T be independent of Y and let it have the uniform distribution on the interval [0,1]. Partition the interval [0,1] into the subintervals $I_0 = [0,1/2]$, and $I_1, \ldots, I_{2^{m-1}}$, where the last 2^{m-1} intervals are all of equal length, 2^{-m}. Define the events $B_i = \{T \in I_i\}$ for $i = 0, 1, \ldots, 2^{m-1}$.

Partition each I_i into two subintervals whose lengths are in the ratio of $(1 - \mathbf{p})$: $(2\mathbf{p} - 1)$. Call the first subinterval J_{i1} and call the second J_{i2}. Let $K_1 = \cup_i J_{i1}$ and let $K_2 = \cup_i J_{i2}$. Define the events $t_{n1} = \pi_n \cap \{T \in K_1\}$ and $t_{n2} = \pi_n \cap \{T \in K_2\}$, so that $t_{n1} \cup t_{n2} = \pi_n$.

The marginal probabilities for these newly defined events are:

$$\mathbf{P}(t_{n1}) = (1 - \mathbf{p})^n, \mathbf{P}(t_{n2}) = (2\mathbf{p} - 1)(1 - \mathbf{p})^{n-1},$$

$$\mathbf{P}(B_i) = 1 - \mathbf{p} \quad (i = 1, \ldots, 2^{m-1}), \quad \text{and} \quad \mathbf{P}(B_0) = \frac{1}{2}.$$

Each t_{n1} is independent of each B_i, so that for $i = 1, \ldots, 2^{m-1}$,

$$\mathbf{P}(B_i \cap t_{n1}) = (1 - p)^{n+1}.$$

Next, define W_m, the generalized St. Petersburg variable as follows:

$$W_m(\omega) = (1 - \mathbf{p})^{-n} \quad \text{for } \omega \in t_{n1} \quad \text{and} \quad W_m(\omega) = 0 \quad \text{for } \omega \in t_{n2}.$$

Note that W_m does not depend on B_i, and W_m has infinite $\mathbf{P}$-expectation.

For the remainder of the proof, we use the following notational shortcut. For a random variable X and an event B, we use $X(B) = c$ to indicate that X is constant on B and equals c. Define the variables X_i so that for $i = 1, \ldots 2^{m-1} - 1$,

$$X_i(B_i \cap t_{n1}) = (1 - \mathbf{p})^{-(n+1)}$$

$$X_i(B_i \cap t_{n2}) = 0$$

$$X_i(B_{i+1} \cap t_{n1}) = X_i(B_{i+1} \cap t_{n2}) = (1 - \mathbf{p})^{-1}$$

for other states ($j \neq i, i+1$):

$$X_i(B_j \cap t_{n1}) = X_i(B_j \cap t_{n2}) = 0$$

Table 3
The partition of π_n into $2^{m-1}+1$ rows and 2 columns, with the values of the $2^{m-1}+1$ equivalent variables displayed within the table.

	t_{n1}	t_{n2}
B_1	$W_m=(1-\mathbf{p})^{-n}$ $X_1=(1-\mathbf{p})^{-(n+1)}$ $X_2=0$ $\ldots$ $X_{2^{m-1}-1}=0$ $X_{2^{m-1}}=(1-\mathbf{p})^{-1}$	$W_m=0$ $X_1=0$ $X_2=0$ $\ldots$ $X_{2^{m-1}-1}=0$ $X_{2^{m-1}}=(1-\mathbf{p})^{-1}$
B_2	$W_m=(1-\mathbf{p})^{-n}$ $X_1=(1-\mathbf{p})^{-1}$ $X_2=(1-\mathbf{p})^{-(n+1)}$ $X_3=0$ $\ldots$ $X_{2^{m-1}}=0$	$W_m=0$ $X_1=(1-\mathbf{p})^{-1}$ $X_2=0$ $X_3=0$ $\ldots$ $X_{2^{m-1}}=0$
$\vdots$		
B_i	$W_m=(1-\mathbf{p})^{-n}$ $X_1=0$ $\ldots$ $X_{i-2}=0$ $X_{i-1}=(1-\mathbf{p})^{-1}$ $X_i=(1-\mathbf{p})^{-(n+1)}$ $X_{i+1}=0$ $\ldots$ $X_{2^{m-1}}=0$	$W_m=0$ $X_1=0$ $\ldots$ $X_{i-2}=0$ $X_{i-1}=(1-\mathbf{p})^{-1}$ $X_i=0$ $X_{i+1}=0$ $\ldots$ $X_{2^{m-1}}=0$
$\vdots$		
$B_{2^{m-1}}$	$W_m=(1-\mathbf{p})^{-n}$ $X_1=0$ $\ldots$ $X_{2^{m-1}-2}=0$ $X_{2^{m-1}-1}=(1-\mathbf{p})^{-1}$ $X_{2^{m-1}}=(1-\mathbf{p})^{-(n+1)}$	$W_m=0$ $X_1=0$ $\ldots$ $X_{2^{m-1}-2}=0$ $X_{2^{m-1}-1}=(1-\mathbf{p})^{-1}$ $X_{2^{m-1}}=0$
B_0	$W_m=(1-\mathbf{p})^{-n}$ $X_1=0$ $\ldots$ $X_{2^{m-1}}=0$	$W_m=0$ $X_1=0$ $\ldots$ $X_{2^{m-1}}=0$

and

$$X_i(B_0\cap t_{n1})=X_i(B_0\cap t_{n2})=0.$$

For $X_{2^{m-1}}$, modify only the third line of this definition, as follows:

$$X_{2^{m-1}}(B_{2^{m-1}}\cap t_{n1})=(1-\mathbf{p})^{-(n+1)}$$

$$X_{2^{m-1}}(B_{2^{m-1}}\cap t_{n2})=0$$

$$X_{2^{m-1}}(B_1\cap t_{n1})=X_{2^{m-1}}(B_1\cap t_{n2})=(1-\mathbf{p})^{-1}$$

for other states ($j\neq 2^{m-1}$, 1):

$$X_{2^{m-1}}(B_j\cap t_{n1})=X_{2^{m-1}}(B_j\cap t_{n2})=0$$

and

$$X_{2^{m-1}}(B_0\cap t_{n1})=X_{2^{m-1}}(B_0\cap t_{n2})=0.$$

The X_i are pairwise equivalent variables, as is evident from the symmetry of their definitions and the fact that the first 2^{m-1} rows have equal probability. Each X_i is equivalent to W_m as well, since the probability that each assumes the value $(1-\mathbf{p})^{-n}$ equals $(1-\mathbf{p})^n$ for $n=1,2,\ldots$. Table 3, below, displays these $2^{m-1}+1$ variables defined over the $(2^{m-1}+1)\times 2$ matrix partition of a single π_n.

We establish a contradiction with the hypothesis that the difference between pairs of equivalent variables is indifferent to ***0***, as follows. Consider the variable obtained by the finite sum:

$$Z_m = \sum_{i=1}^{2^{m-1}} (X_i - W_m).$$

Then

$$Z_m(B_i \cap t_{n1}) = \frac{(1-\mathbf{p})-n}{2} + (1-\mathbf{p})^{-1}$$

$$Z_m(B_i \cap t_{n2}) = (1-\mathbf{p})^{-1}$$

$$Z_m(B_0 \cap t_{n1}) = -2^{m-1}(1-\mathbf{p})^{-n} = -\frac{(1-\mathbf{p})^{-n}}{2}$$

$$Z_m(B_0 \cap t_{n2}) = 0.$$

Note that Z_m does not distinguish among the B_i, which we may now collapse into a single row of cells, denoted by their union B_0^c, with combined probability 1/2.

Write Z_m as a sum of three variables, T_m, U_m, and V_m, defined as follows on the four events in $\{B_0, B_0^c\} \times \{t_{n1}, t_{n2}\}$ that partition π_n:

$$T_m(B_0^c \cap t_{n1}) = -U_m(B_0 \cap t_{n1}) = \frac{(1-\mathbf{p})^{-n}}{2}$$

$$T_m(B_0^c \cap t_{n2}) = T_m(B_0 \cap t_{n1}) = T_m(B_0 \cap t_{n2}) = 0$$

$$U_m(B_0^c \cap t_{n1}) = U_m(B_0^c \cap t_{n2}) = U_m(B_0 \cap t_{n2}) = 0$$

$$V_m(B_0^c \cap t_{n1}) = V_m(B_0^c \cap t_{n2}) = (1-\mathbf{p}^{-1})$$

$$V_m(B_0 \cap t_{n1}) = V_m(B_0 \cap t_{n2}) = 0.$$

Note that as $P(B_0^c) = 1/2$ and V_m is simple, $V_m \sim 2^{m-1}$. Observe also that T_m and $-U_m$ are equivalent, though unbounded variables. By the hypothesis that the difference between two equivalent variables is indifferent to ***0***, then $(T_m + U_m)$ is indifferent to ***0***. Thus, equation (*) follows, which contradicts the hypothesis that the 2^{m-1} many variables $((X_i - W_m)$, for $i = 1, \ldots 2^{m-1}$, each is indifferent to ***0***.

Aside: The pairwise equivalences among the $2^{m-1}+1$ many variables $W_m, X_1, X_2, \ldots, X_{2^{m-1}}$, obtain over all values of $\mathbf{p}$ for which the construction above is well defined, i.e., the equivalence obtains for all $1 > \mathbf{p} \geq 1 - 2^{-(m-1)}$. However, in order to avoid appeal to the following extra assumption, we apply the construction solely to the case where $\mathbf{p} = 1-2^{-m}$, when the proof of the theorem does not require an extra assumption. The additional assumption needed to apply the construction to the other values of $\mathbf{p}$ is that, if (*i*) X is simple with $\mathbf{V}(X) = 0$ and (*ii*) X and Y are independent, then $\mathbf{V}(XY) = 0$.

Appendix B. Proof of Theorem 2

We offer an indirect proof, assuming the hypothesis that equivalent variables with finite absolute expectations carry equal prevision. The argument is presented in 3 parts: Part 1 of the proof defines the equivalent variables whose previsions, in the end, cannot all be equal. The construction begins with two *iid* Geometric($\mathbf{p}$) variables. Part 2 develops two results about how previsions for independent variables relate to their expected values, assuming the hypothesis. Part 3 puts the pieces together.

Part 1 of the proof: Let $\mathbf{V}(X) = \mathbf{E}[X] + \boldsymbol{b} = \boldsymbol{t} > \mathbf{p}^{-1}$; so $\boldsymbol{\beta}(X) = \boldsymbol{b} > 0$. Consider two, *iid* draws, X_1, X_2, from this Geometric($\mathbf{p}$) distribution. By the hypothesis $\mathbf{V}(X_i) = \boldsymbol{t}\,(i = 1, 2)$.

Define the variable $W = X_1 + X_2$, which has the NegBin(2,$\mathbf{p}$) distribution. By coherence, then $\boldsymbol{V}(W) = 2\boldsymbol{t}$.

Note that the conditional distribution $\mathbf{P}(X_1|W = n) = (n-1)^{-1}$ for $(1 \leq X_1 \leq n-1)$ is uniform, because $\mathbf{P}(X_1 = k|W = n) = \mathbf{P}(X_1 = k, X_2 = n-k, W = n)/\mathbf{P}(W = n)$. This follows as

$$\begin{aligned}\mathbf{P}(X_1 = k, X_2 = n-k, W = n) &= \mathbf{P}(X_1 = k, X_2 = n-k)\\ &= \mathbf{p}(1-\mathbf{p})^{k-1}\mathbf{p}(1-\mathbf{p})^{n-k-1}\\ &= \mathbf{p}^2(1-\mathbf{p})^{n-2}\end{aligned}$$

which is constant (and positive) for $1 \le k \le n-1$.Hence,

$$\frac{\mathbf{P}(X_1 = i|W = n)}{\mathbf{P}(X_1 = j|W = n)} = 1 \quad \text{for } 1 \le i, j \le n-1.$$

Write W as a sum of three variables: $W = W_1 + W_2 + W_3$, as defined below. The first two of these will be equivalent but they will have different boosts. Each of these variables equals W for approximately 1/3 of all ω and equals 0 for the other approximately 2/3 of all ω. There are $n-1$ ω for which $W(\omega) = n$. When $n-1$ is not a multiple of 3, $W_3 = n$ for one or two ω more than each of W_1 and W_2. In the (X_1, X_2)-plane, the sample space for W is the set of points in the first quadrant with positive integer values in both coordinates. In the sample space, W is constant along each finite line segment with a slope -1.The random variables W_i ($i = 1, 2, 3$) are defined as follows. Let k_n denote the greatest integer less than or equal to $(n-1)/3$ for each integer n.

For each n

$W_1 = n$ for the k_n points satisfying $\{W = n \text{ and } 1 \le X_2 \le k_n\}$, and $W_1 = 0$ for all other points.
$W_2 = n$ for the k_n points satisfying $\{W = n \text{ and } k_n + 1 \le X_2 \le 2k_n\}\}$, and $W_2 = 0$ for all other points.
$W_3 = n$ for the $n - 1 - 2k_n$ points satisfying $\{W = n \text{ and } 2k_n + 1 \le X_2 \le n-1\}\}$, and $W_3 = 0$ for all other points.

It is evident that $W = W_1 + W_2 + W_3$, so $\boldsymbol{V}(W) = \sum_i \boldsymbol{V}(W_i)$. Also it is evident that $W_1 \equiv W_2$, as these two variables are constructed so that, for each $n = 1, 2, \ldots$, the event $W_i = n$ ($i = 1, 2$) obtains for the same number of $\mathbf{P}$-non-null points, and $\mathbf{P}(\cdot|W = n)$ has a uniform distribution on its support of $n-1$ points.

Part 2 of the proof: Next, we develop two general claims about the $\boldsymbol{V}$-values for independent variables, Lemmas 1 and 2, from which, in Part 3, we derive that $\boldsymbol{V}(W_2) < \boldsymbol{V}(W_1)$, in contradiction with the hypothesis that equivalent variables are equally preferred. Both lemmas assume the hypothesis that equivalent variables are indifferent.

Lemma 1. *Let Y be a nonnegative integer variable with finite mean, $\boldsymbol{E}(Y) = \mu < \infty$, and finite value $\boldsymbol{V}(Y) = \pi < \infty$. Coherence assures that $\mu \le \pi$. Let F be the indicator for an event, independent of Y, with $\boldsymbol{P}(F) = \alpha$. Then $\boldsymbol{V}(FY) = \alpha\pi$.*

Proof. If α is a rational fraction, $\alpha = k/m$, the lemma follows using the hypothesis that equivalent variables are indifferent, applied to the m-many equivalent variables F_iX, where $\{F_1, \ldots, F_m\}$ is a partition into equiprobable events F_i. That is, from the hypothesis, $\boldsymbol{V}(F_iY) = c$ ($i = 1, \ldots, m$), and by finite additivity of previsions, then $c = \pi/m$, so that $\boldsymbol{V}(FY) = k\pi/m = \alpha\pi$. If α is an irrational fraction, the lemma follows by dominance applied to two sequences of finite partitions of equally probable events. One sequence provides bounds on $\boldsymbol{V}(FY)$ from below, and the other sequence provides bounds from above. □

Next, let X and Y be independent variables, with X bounded below, defined on the positive integers N. Consider a function $\boldsymbol{g}(i) = j$, $\boldsymbol{g}{:}N \to N$, with the sole restriction that for each value j, $\boldsymbol{g}^{-1}(\{j\})$ is a finite (and possibly empty) set. The graph of the function $\boldsymbol{g}$ forms a binary partition of the positive quadrant of the (X, Y)-plane into events G and G^c, with G defined as: $G = \{(x, y){:}\ \boldsymbol{g}(x) \le y\}$. G is the region at or above the graph of $\boldsymbol{g}$. Then, on each horizontal line of points in the positive quadrant of the (X, Y)-plane, on a line satisfying $\{Y = j\}$, only finitely many points belong to the event G.

Let GX denote the variable that equals X on G and 0 otherwise, and likewise for the variable $G^{\mathbf{c}}X$. So, $X = GX + G^{\mathbf{c}}X$. The next lemma shows how, under the hypothesis that equivalent variables are indifferent, the *boost* $\boldsymbol{\beta}(X)$ for the variable X divides over the binary partition formed by the event G.

Lemma 2. *With X, Y, and G defined above,*

$$\boldsymbol{V}(GX) = \mathbf{E}(GX), \quad \text{whereas } \boldsymbol{V}(G^cX) = \mathbf{E}(G^cX) + \boldsymbol{\beta}(X).$$

*That is, **all** of the boost associated with the prevision of X attaches to the event G^c, regardless the probability of $G^{\mathbf{c}}$.*

Proof. For each value of $j = 1, 2, \ldots$, write the variable $\{Y = j\}X$ as a sum of two variables, using G (respectively G^c) also as its indicator function:

$$\{Y = j\}X = \{Y = j\}GX + \{Y = j\}G^cX.$$

So,

$$\boldsymbol{V}(\{Y = j\}G^cX) = \boldsymbol{V}(\{Y = j\}X) - \boldsymbol{V}(\{Y = j\}GX)$$

and

$$\mathbf{E}(\{Y = j\}G^cX) = \mathbf{E}(\{Y = j\}X) - \mathbf{E}(\{Y = j\}GX).$$

But $\{Y = j\}GX$ is a simple variable, as the event G contains only finitely many points along the strip $\{Y = j\}$. Thus, $\boldsymbol{V}(\{Y = j\}GX) = \mathbf{E}(\{Y = j\}GX)$.

Since X and Y are independent, by Lemma 1,

$$\boldsymbol{V}(\{Y = j\}X) = \mathbf{P}(Y = j)(\mathbf{E}[X] + \boldsymbol{\beta}(X))$$

So,

$$\begin{aligned}\boldsymbol{V}(\{Y=j\}G^cX) &= \mathbf{P}(Y=j)(\mathbf{E}[X]+\boldsymbol{\beta}(X)) - \mathbf{E}(\{Y=j\}GX)\\ &= \mathbf{P}(Y=j)\boldsymbol{\beta}(X) + \mathbf{E}(\{Y=j\}X) - \mathbf{E}(\{Y=j\}GX)\\ &= \mathbf{P}(Y=j)\boldsymbol{\beta}(X) + \mathbf{E}([Y=j]G^cX).\end{aligned}$$

Thus, the prevision for $\{Y=j\}G^cX$ contains a boost equal to $\mathbf{P}(Y=j)\boldsymbol{\beta}(X)$. But as $\sum_j \mathbf{P}(Y=j)\boldsymbol{\beta}(X)=\boldsymbol{\beta}(X)$, we have $V(G^cX)=\sum_j(\mathbf{E}(\{Y=j\}G^cX)+\mathbf{P}(Y=j)\boldsymbol{\beta}(X))=\mathbf{E}[G^cX]+\boldsymbol{\beta}(X)$ and there is no boost associated with GX, $\boldsymbol{V}(GX)=\mathbf{E}[GX]$. □

Part 3 of the proof: Recall that, by hypothesis, since $X_1 \equiv X_2$, then $\mathbf{V}(X_i)=\boldsymbol{t}$ and $\boldsymbol{\beta}(X_i)=\mathbf{b}>0$, for $i=1,2$. Apply Lemma 2 with X being X_1, Y being X_2, and $g(i)=(i-1)/2$. One can verify that G^c is the set where $W_1=W$. It follows from Lemma 2 that G^cX_1 gets all of the boost of X_1 while GX_1 gets none of the boost. Apply Lemma 2 again, this time with X being X_2, Y being X_1, and $g(i)=(i+5)/2$ for odd i, and $g(i)=(i+2)/2$ for even i. This time, let H denote the set called G in Lemma 2 so that we can distinguish it from the set found in the first application of Lemma 2. Now, H^c is the set where $W_3=W$, and H^cX_2 gets all of the boost of X_2 while HX_2 gets none of the boost. We can write $W_1=G^cX_1+G^cX_2$, so that W_1 gets all of the boost of X_1. It is clear that $G^c\subset H$, hence G^cX_2 gets none of the boost of X_2, and neither does W_1. Similarly, we can write $W_3=H^cX_1+H^cX_2$, so that W_3 gets all of the boost of X_2. Since $H^c\subset G$, W_3 gets none of the boost of X_1. In summary

W_1 gets all of the boost of X_1 and none of the boost of X_2, $\boldsymbol{\beta}(W_1)=\mathbf{b}>0$;
W_3 gets all of the boost of X_2 and none of the boost of X_1, $\boldsymbol{\beta}(W_3)=\mathbf{b}>0$.

There is no boost left for W_2, hence W_2 gets none of the boost of either X_1 or X_2, and $\boldsymbol{\beta}(W_2)=0$.

Since $W_1\equiv W_2$, then $\mathbf{E}[W_1]=\mathbf{E}[W_2]$. Therefore, by adding the respective boosts, $\boldsymbol{V}(W_2)<\boldsymbol{V}(W_1)$, which establishes the Theorem.

References

Becker, R.A., Boyd, J.H., 1997. Capital Theory, Equilibrium Analysis, and Recursive Utility. Blackwell, Oxford.
Colyvan, M., 2008. Relative Expectation Theory. Journal of Philology 105, 37–44.
de Finetti, B., 1972. Probability, Induction, and Statistics. Wiley, New York.
de Finetti, B., 1974. Theory of Probability, Vols. I and II. Wiley, New York.
Dunford, N., Schwartz, J.T., 1988. Linear Operators. Part 1. General Theory. Wiley, New York.
Fishburn, P.C., 1979. Utility Theory for Decision Making. Krieger, New York.
Hausner, M., 1954. Multidimensional utilities. In: Thrall, R.M., Coombs, C.H., Davis, R.L. (Eds.), Decision Processes. Wiley, New York.
Kadane, J.B., Schervish, M.J., Seidenfeld, T., 1986. Statistical implications of finitely additive probability. In: Goel, P.K., Zellner, A. (Eds.), Bayesian Inference and Decision Techniques. Elsevier, Amsterdam, pp. 59–76.
Levi, I., 1980. The Enterprise of Knowledge. MIT Press, Cambridge.
Royden, H.L., 1968. Real Analysis. Macmillan, New York.
Savage, L.J., 1972. The Foundations of Statistics, 2nd revised edition. Dover, New York.
Seidenfeld, T., Schervish, M.J., 1983. A conflict between finite additivity and avoiding Dutch Book. Philosophy of Science 50, 398–412.
Troffaes, M., 2006. Conditional lower previsions for unbounded random quantities. In: Jonathan Lawry, Enrique Miranda, Alberto Bugarin, Shoumei Li, Mariá Ángeles Gil, Przemyslaw Grzegorzewski, and Olgierd Hryniewicz (Eds.), Soft Methods in Probability for Integrated Uncertainty Modelling, Advances in Soft Computing: 201–209, Springer.
Wakker, P., 1993. Unbounded utility for Savage's 'Foundations of Statistics' and other models. Mathematics of Operations Research 18, 446–485.