



Munich Personal RePEc Archive

**Asymptotic Theory Of Stochastic Choice
Functionals For Prospects With
Embedded Comotonic Probability
Measures**

Cadogan, Godfrey

27 April 2010

Online at <https://mpra.ub.uni-muenchen.de/22380/>
MPRA Paper No. 22380, posted 29 Apr 2010 00:38 UTC

Asymptotic Theory Of Stochastic Choice Functionals For Prospects With Embedded Comotonic Probability Measures

Godfrey Cadogan *
Working Paper

April 27, 2010

*Corresponding address: 3401-B N.W. 72nd Ave, Ste # T-419, Miami, FL 33122; e-mail: go-cadog@gmail.com. I am grateful to Richard Gonzales, and Mark Zanecki for comments on a prior draft of this paper. I thank Jan Kmenta for the idea of embedding parameters of interest in economic variables. The research assistance of Oliver Martin and Sampson Cadogan is gratefully acknowledged.

Abstract

We introduce a monotone class theory of value functions, which shows that they can be replaced almost surely by a topological lifting comprised of a class of compact isomorphic maps that embed weakly co-monotonic probability measures, attached to state space, in outcome space. Thus, agents solve a signal extraction problem to obtain estimates of empirical probability weights for prospects under risk and uncertainty. By virtue of the topological lifting, we prove the following almost sure isomorphism theorem between stochastic choice and well defined outcome:

$$\lim_{n \rightarrow \infty} \Pr \left\{ \sup_{A_n \in \mathfrak{B}(X)} \left| J(w \circ P)(A_n, \omega) - (V^{-1}TV)(A_n, \omega) \right| > \varepsilon \right\} = 0$$

where S is a states space, \mathcal{P} is a partition of S , J is a stochastic choice operator, w is a probability weight function, P is a probability measure on the sample space Ω for states of nature–attached to S , $\{A_n\}_{n=1}^{\infty}$ is a monotone sequence of subsets in $\mathfrak{B}(\mathcal{P})$ –the σ -field of Borel subsets of \mathcal{P} , T is a risk operator in the Hoffman-Jorgensen class of lifting operators, and V is a *value function operator* with respect to Radon measure m . In particular, V is an *averaging operator*, and $T = T_g \oplus T_\ell$ where \oplus is a binary operation on the space of value functions; T_g is an isometric gain operator; T_ℓ is a skewed loss operator; and $\Lambda_T = \{1\} \cup \Lambda_{T_\ell}$ is the [point] spectrum of T when Λ_{T_ℓ} is the spectrum for T_ℓ .

JEL Codes: C3, C5

AMS 2010 Subject Classification: 60G20, 60J25, 60J60, 60J65

Contents

1	INTRODUCTION	3
2	OVERVIEW OF PROSPECT THEORY	12
2.1	Risk, Uncertainty, Rank Dependent Utility, and Choquet Capacity	15
3	A SPECTRAL THEORY OF VALUE FUNCTIONAL DECOMPOSITION	17
3.1	Binary conditions for value operator	20
4	PROSPECT TOPOLOGY	23
4.1	Topological lifting of prospect space	23
4.1.1	Preliminaries	23
4.1.2	Commutative map of Prospect Theory's Liftings	25
4.1.3	Hoffman-Jorgensen lifting on Lebesgue spaces	26
4.1.4	Kernel function generation of probability and choice	27
4.2	Existence of mapping from state space to reals	28
4.3	Weak comonotone probability embedding	29
4.4	Almost sure isomorphisms of lifted choice functions	35
5	Conclusion	42
6	APPENDIX	43
A	Rank dependent utility and Schmeidler's Choquet representation	43
B	Topological attachment and embedding of probability spaces	44
C	Elements of operator theory used in this paper	46
C.1	Definitions and theorems	46
D	Some Limit Theorems	47
E	Proofs of theorems, lemmas, and propositions	48

List of Figures

1	Commutative Map of Prospect Theory's Liftings	25
---	---	----

1 INTRODUCTION

At issue is the following excerpt from (Tversky and Khaneman, 1992, pg. 300):

Let S be a finite set of states of nature; subsets of S are called events. It is assumed that exactly one state obtains, which is unknown to the decision maker. Let X be a set of consequences also called outcomes.

* * * * *

An uncertain prospect f is a function from S into X that *assigns* to each state $s \in S$ a consequence $f(s) = x$ in X . To define the cumulative functional, we arrange the outcomes of each prospect in increasing order. A prospect f is then represented as a sequence of pairs (x_i, A_i) which yields x_i if A_i occurs

* * * * *

Cumulative prospect theory asserts that *there exists* a strictly increasing value function $v : X \rightarrow Re$, satisfying $v(x_0) = v(0) = 0$, . . . [Emphasis added].

Why do prospect theory's agents map from state space to outcome space, and then map a value function from outcome space to the reals? Why can't they simply map directly from state space to the reals? In other words, are value functions irrelevant? (Luce and Narens, 2008, pg. 1) characterized problems of this type thusly:

Most mathematical sciences rest upon quantitative models, and the theory of measurement is devoted to making explicit the qualitative assumptions that underlie them. This is accomplished by first stating the qualitative assumptions empirical laws of the most elementary sort

in axiomatic form, and then showing that there are *structure preserving mappings*, often but not always *isomorphisms*, from the qualitative structure into a quantitative one. The set of such mappings forms what is called a scale of measurement. [Emphasis added].

To answer the research questions posed, this paper introduces several “scale[s] of measurement”–derived from an isomorphic topological lifting of the imputed direct map from “qualitative” state space to the reals. Along the way, we prove (1) that the direct map is measurable with respect to stochastic choice functions¹ if and only if a value function defined on outcome space exists; (2) Radon integral representation of value functional with respect to stochastic choice; and (3) spectral theory of value function operators.

First, extant literature on discrete choice typically apply logit and probit type models to estimate binary choice probabilities². Debreu (1958) provided axiomatic foundations for the existence of [cardinal] utility functions on state space consistent with stochastic choice on that space. That foundation was extended by McFadden (1974) to include a panoply of discrete choice models. Recently, (Narens, 2007, pg. 76) explained Luce (1959) choice axiom, which provides a “ratio scale” for discrete choice parametrizations of conditional probability of choice sets—a key element of McFadden’s work. See (McFadden, 1980, pg, S16).

¹(Debreu, 1958, pg. 440) defines a stochastic choice function as follows:

Definition 1.1. Let Y be a choice set, and f be a function from $Y \times Y$ to $[0, 1]$ such that $f(a, b) + f(b, a) = 1$ for every $(a, b) \in Y \times Y$. Then f is a stochastic choice function.

²The literature on stochastic choice is huge. Dating at least back to Fechner’s psychophysics experiments in the 1850’s concerning stimulus intensity and a subject’s detection of the stimulus. See e.g., (Hunt, 2007, pp. 47-53). See also, Train (2003) and Stott (2003) for pertinent reviews of discrete choice models.

However, recent advances in prospect theory, implicating the efficacy of von Neuman Morgenstern utility models, suggest that at least a cursory re-examination of stochastic choice in the context of probability weight functions and comonotonic probabilities is apropos. Our spectral theory of value functions, and matrix representation of choice functions is founded in prospect theory’s function space, so it accounts for the impact of probability weight functions–phenomenon lacking in extant discrete choice models³.

Second, our signal extraction approach to estimating probability weight functions is distinguished from extant literature in which probability weight functions are parametrized with an inverted S-shape⁴ and empirically tested. See e.g., Tversky and Khaneman (1992). For instance, Prelec (1998), introduced a generalized inverted S-shape probability weight function which he derived from axiomatic foundations based on the concept of *compound invariance*⁵. Luce (2001)

³Acerbi (2001) introduced a risk based spectral theory that is fundamentally different from ours. de Palma et al. (2008) included probability weighting in a discrete choice model parenthetically but did not provide an explicit parametrization. Also, Cížek (2007) proposed a trimmed estimation scheme to address the sensitivity of maximum likelihood estimators of binary choice models to tail observations. However, the “trimmed” observations are precisely those that are overweighted or underweighted, accordingly, to induce the inverted S-shape phenomenon in probability weight functions. Additionally, despite Train (2003) encyclopedic review of discrete choice models, he is silent on matrix representation of choice functions.

⁴The inverted S-shape phenomenon derives from experiments which found that people tend to over weigh small probabilities (concavity) and under weigh large probabilities (convexity). See e.g. Hsu et al. (2009)(fMRI analysis). Thereby inducing concave and convex regions around a fixed point on a probability weight function that should otherwise be inclined at 45% to the horizontal. These phenomena are described as *lower and upper sub-additivity*, respectively, of the probability weight function. See (Gonzalez and Wu, 2003, pg. 7, note 2)

⁵Let (x, p) be a simple gamble in which the consequence x occurs with probability p and nothing otherwise. A gamble is compound if $((x, p), q) \sim (x, pq)$. (Prelec, 1998, pg. 503) defined compound invariance thusly, *mutatis mutandis*:

Definition 1.2. A preference order \succeq exhibits compound invariance if for any outcomes $x, y, x', y' \in X$, probabilities $q, p, r, s \in [0, 1]$, and compounding integer $N \geq 1$: If $(x, p) \sim (y, q)$ and $(x, r) \sim (y, s)$, then $(x', p^N) \sim (y', q^N)$ implies $(x', r^N) \sim (y', s^N)$.

extended Prelec (1998) by using the concept of *reduction invariance*⁶. More recent, al Nowaihi and Dhami (2005) extended Luce (2001) to include *power invariance*⁷. Wu and Gonzalez (1996, 1999) conducted some of the earlier empirical tests of shape parameters for probability weight functions and confirmed the inverted S-shape. Additionally, Gonzalez and Wu (1999) used a switching regression approach to estimating probability weight functions based on a “generalized” log-odds parametrization, i.e. a slight variation of Tversky and Khaneman (1992). An important paper by Bliechrodt and Pinto (2000) reported that in experiments they conducted on parameter free representation, the modal probability weight function was found to be inverted S-shape, and that probability over [under] weighting was especially prevalent in the lower [upper] tails⁸.

The *invariance* concept, above, is based on the notion that the *expected value* of a value function with respect to probability weights over a lottery rep-

⁶(Luce, 2001, pg. 170) defined *reduction invariance* as:

Definition 1.3. Let N be a natural number. Then N -*reduction invariance* is said to hold if and only if, for any consequence x and probabilities $p, q, r \in]0, 1[$, $((x, p), q) \sim (x, r)$ implies $((x, p^N), q^N) \sim (x, r^N)$.
When N -*reduction invariance* holds for all natural numbers N , we say *reduction invariance* holds.

⁷(al Nowaihi and Dhami, 2005, pg. 4) defined *power invariance* as:

Definition 1.4. The probability weighting function w satisfies *power invariance* (PI) if, for all $p, q \in [0, 1]$, $\lambda \in (0, \infty)$ and $m \in \mathbb{N}$, $(w(p))^m = w(q) \Rightarrow (w(p^\lambda))^m = w(q^\lambda)$

⁸By contrast, (Birchby et al., 2008, pp. 9-10) reported that in experiments they conducted the probability weight function was unstable. Specifically, they analogized Cumulative Prospect Theory’s probability weighting function (“PWF”) to the Phillips curve debate in Economics where aggregation bias led to analysts misperception of the shape of the Phillips curve in the short and long run. It should be noted in passing that their experiment provided a “baseline” lottery (Option A) and asked subjects to compare it with a variable lottery (Option B), and asked them to identify which lottery among Option B was equivalent to the baseline Option A. They included kernel function plots from a purported nonparametric estimation scheme, but did not include the parametrized model which generated the plots. Without more, it is not clear whether their experiment is conclusive.

resents the invariant value of the lottery to a subject. Much like the Birkhoff-Khinchin mean in an ergodic theory is invariant over time and space parameters. A *compound lottery* is based on a monotone increasing set of lotteries constructed by recursive union of other simpler lotteries. Thus, it is based on a monotone class theory. *Reduction invariance* implies that a compound lottery $((x, p), q)$ is reducible to an equivalent simple lottery (x, pq) . The equivalence is maintained even when the probabilities in each lottery are raised to a λ power. The idea is motivated by separable utility over gambles, i.e. $U(x, p) = U(x)w(p)$. *Power invariance* implies that raising the probability in a probability weight function by a [positive and finite] power, and then raising that probability weight function by a positive power, does not change its equivalence to the probability weight function of a different probability raised by the same power. See al Nowaihi and Dhami (2005) for an excellent set theoretic summary of these issues. See also, (Narens, 2007, pg. 51) for a probabilistic construct of compound invariance.

Third, our behavioral probability phase function appears to be new⁹. It provides many of the properties of a probability weight function, including but not limited to inverted S-shape. However, it accommodates cyclical behavior that arise in choice under risk and uncertainty¹⁰. Thereby providing a natural mechanism

⁹Stott (2003) conducted a meta-analysis of prospect theory's functional forms that include but is not limited to probability weight functions. However, he did not mention probability phase functions. Subsequently, Takahashi (2006) introduced an entropy based probability weighting scheme in which the expected value of a value function is $V = A(p^\alpha - T\{-(p \log_2 p + (1-p) \log_2 (1-p))\})$ where A is outcome, T is a degree of aversion, and $0 < \alpha < 1$ is psychophysical effect on small probabilities. Cf. Burns et al. (2010).

¹⁰See (Davidson and Marschak, 1958, pg. 1) who stated:

Common experience suggests, and experiment confirms, that a person does not always make the same choice when faced with the same options, even when the circumstances of choice seem in all relevant respects to be the same. However, the bulk of economic theory neglects the existence of such incon-

for explaining behavior over time¹¹.

This paper comprises topological analysis of prospect theory's function space under risk and uncertainty. It does not address axiomatic foundations of stochastic choice¹². Specifically, we show that an isomorphic topological lifting of the imputed direct map from state space to the reals, defined on probability weights attached to the state space, and embedded in outcome space, renders the value function superfluous¹³. So we must solve a signal extraction problem to obtain estimates of probability weights when outcomes are known but concomittant probabilities are not¹⁴. For instance, instead of the strong comonotonic prospect

$$(x_{(1)}, P_{(1)}; x_{(2)}, P_{(2)}; \dots; x_{(n)}, P_{(n)})$$

in increasing order (best to worst) of preference for outcomes $\{x_i\}_{i=1}^n$, we propose a weak comonotonic ranking of outcomes that depend on probabilities¹⁵. In other words instead of the ranked pair (x, p) we embed probability p in x and rank the functions $\{x(p)\}_{x \in X}$ ¹⁶. Conceptually, our procedure exploits the equivalence class

sistencies and the best known theories for decision making, for example, those of von Neumann and Morgenstern [12] or Savage [15], base the existence of a measurable utility upon a pattern of invariant two-place relations, sometimes called 'preference' and 'indifference'. This raises a difficulty for any attempt to use such theories to describe and predict actual behavior

See also, (Dagsvik, 2006, pg. 1) and <http://moodviews.com/>.

¹¹(Fabiyi, 2008, pg. 113) introduced an heuristic hyperbolic tangent function (tanh) to parametrize Tversky and Khaneman (1992) type value functions. However, he did not use that function for probability weighting.

¹²See Dagsvik (2006) for a recent survey and extension

¹³We are assuming, without deciding, that prospect theory's "choice function" from state space to outcome space is regular.

¹⁴ Wu et al. (2004) reviewed how decisions are made under risky conditions when *objective* probabilities are known. However, they did not broach decision under uncertainty

¹⁵See (Debreu, 1953, Thm. II, pg. 6) for justification of this type of embedding

¹⁶ This implies nonseparability of gambles. At best, it implies weakly separable utility for a simple lottery or gamble (x, p) because $U(x(p))$ may not be equal to $U(x)w(p)$. Cf. (Luce, 2001, pp. 167-168). Even though one could

of comonotone ranking of outcomes and probabilities. To see this, let \mathcal{M} be the space of probability measures, and X be a separable outcome space. For open $U \subset \mathcal{M}$ define the equivalence class \mathcal{R} on $\mathcal{M} \times X$ by

$$\mathcal{R} = \{(x, p) | p \in U, x : U \rightarrow X\} \subset \mathcal{M} \times X$$

We are interested in the quotient space¹⁷ $(\mathcal{M} \times X)/\mathcal{R}$ that include ranked outcomes

$$\left(x_{(1)}(p_{k_1^{(n)}}); x_{(2)}(p_{k_2^{(n)}}); \dots; x_{(n)}(p_{k_n^{(n)}}) \right)$$

There, embedded unknown probabilities $\{p_{k_j^{(n)}}\}$ are weakly comonotonic by virtue of being embedded in ranked outcomes $\{x_{(j)}(p_{k_j^{(n)}})\}$ —for the index $k_j^{(n)}, j = 1, 2, \dots, n$ which belongs to the index set $\{1, 2, \dots, n\}$. Thus, we have for the subsequence of indexes $\{k_j^{(n)}\}$ the weak limit

$$\lim_{n \rightarrow \infty} x_{(j)}(p_{k_j^{(n)}}) = x_{(j)}(p_{k_j})$$

In that setup, we prove that outcomes are related to choice functions and probability weights by the *assignment* relation

$$f \circ w \circ p + o_p(1) = x$$

conceptually argue for pseudo-separability where $x(p) \sim (y, p)$ for some y and that $U(x(p)) = U(y)w(p)$

¹⁷See (Jacobson, 1951, pp. 5-6) for definition of equivalence class, and quotient space. See also, (Debreu, 1953, pg. 5).

where $f \circ w \circ p$ is an unspecified composite choice function, i.e. operator, from state space to outcome space, and w is a probability weight function that transforms p to $w(p)$. Intuitively, if f is a choice function, then it is *an operation* which maps probability measures into outcome space. After performing this operation agents *assign* a given outcome to it modulo measurement error¹⁸. This “operator” aspect of f permits a matrix representation in function space.

In order to be consistent with the literature on choice functions, we let

$$f = \beta g, \text{ where } 0 \leq g \leq 1$$

and β is a scale factor. Thus f is really a scaling of the “true” choice function g .

We adopt the rule that

$$x_i \mapsto \frac{x_i}{\max_i x_i \in f(\mathcal{P})},$$

where \mathcal{P} is a partition of state space. In that way, we retain the use of f as a probabilistic choice function. We make slight modifications to a monotone class theorem in (Blumenthal and Geetor, 1968, pg. 7, Prop. 2.7) which, in the context of our model, essentially states that our direct map is measurable with respect to KT92 choice functions iff there is a monotone sequence of functions on outcome space X , i.e. strictly increasing value functions. Our weak comonotonic embedding procedure applies to the proof.

¹⁸For instance, the canonical logit parametrization $\log \frac{P}{1-P} = \mathbf{x}^T \beta + \varepsilon$ where P is the probability associated with a given choice, ε is measurement error, and \mathbf{x} is a vector of attributes, i.e., outcomes, *assigns* the logit of P to outcomes \mathbf{x} .

To estimate $w(p)$, we must either parametrize the composite function or proceed nonparametrically. Either way, some estimation scheme on the measure space $(X, \mathfrak{B}(X))$ is needed to extract $w(p)$ from the [ranked] functional $x(w(p))$. To formulate a scheme, we endow outcome space with a suitable metric induced norm, and transform it into a Banach space. There, we introduce a spectral theory of loss aversion and the solution space for prospect theory is extended. In a companion paper, we use well known ordinary least squares¹⁹ and factor analysis²⁰ approaches to devise estimation schemes for probability weight functions (“PWFs”). In order not to overload this paper, we did not address the critical issue of estimation of loss aversion index. That issue is addressed in a another paper where a semiparametric two stage regression approach is used to show that the asymptotic distribution of Tversky and Khaneman (1992) loss aversion index is a Brownian bridge functional.

To make this paper as self contained as possible we make liberal use of definitions of esoteric terms throughout, and reproduce some key theorems in an appendix. This paper proceeds as follows. In section 2 we provide an overview of prospect theory. In section 3 we use elementary operator theory to establish a nexus between stochastic choice functions and value functionals in Theorem 3.1. An independently important spectral theory of loss aversion is introduced—based on binary operations in subsection 3.1 for gain-loss averaging operators motivated

¹⁹Gonzalez and Wu (1999) used a nonparametric alternating regression for ”switching” δ to estimate probability weight functions $\frac{\delta p^\gamma}{\delta p^\gamma + (1-p)^\gamma}$. Our approach estimates the *entire* probability weight function by a reverse regression approach.

²⁰See (Härdle and Simar, 2003, Chapter 10) for details on this procedure.

by the *invariance principles* above. Thus indicating in Theorem 3.4 that a given value function for losses supports a spectrum of admissible loss aversion indexes. In Theorem 3.3, on page 21, we provide mathematical proof that the gain operator is isometric, whereas the loss operator is skewed. In section 4 the topology of prospect space is introduced. A commutative map, produced in Figure 1 in subsection 4.1.1, provides a pictorial overview of the salient characteristics of Prospect Theory's function space. Theorem 4.2 provides a proof for the existence of the direct map from state space to the reals. The main results on the lifted space and embedding of weakly ranked comonotonic weighted probabilities are addressed in Propositions 4.5 and 4.7. Theorem 4.9 on page 35 is the main result in this paper. It proves almost sure convergence between the lifting of stochastic choice functions and an isomorphic mapping of value functions. We conclude with perspectives in section 5.

2 OVERVIEW OF PROSPECT THEORY

We begin with a concise overview of prospect theory-taking all axioms as given. Suppose we are given a separable outcome space X , and an array of consequences or outcomes $\{x_1, x_2, \dots, x_n\}$. Associated with each outcome x_i is a gamble or measurable event E_i drawn from an abstract state space S and its associated σ -algebra \mathfrak{G} , where $\mathcal{P} = \{E_1, E_2, \dots, E_n\}$ is a partition of S . So that $S = \bigcup_{j=1}^n E_j$ for measurable subsets $E_j \in S$. A prospect is defined on the product space $X \times S$,

i.e., $\{x_1, E_1; x_2, E_2; \dots; x_n, E_n\} \subset X \times S^{21}$. Let P be a probability measure on the measurable space (S, \mathfrak{S}) so that $P(E_i) = p_i$ is the probability assigned to the i -th event. If the probability distribution is known, then the prospect is risky. If the probability distribution is unknown, then it (the prospect) is uncertain²². Rank order the choices monotonically, i.e., in ascending order from most important $x_{(1)}$ to least important $x_{(n)}$ to get the permutation²³

$$\sigma\{x_1, x_2, \dots, x_n\} \rightarrow \{x_{(1)}, x_{(2)}, \dots, x_{(n)}\} \quad (2.1)$$

Notationally, $E_{(i)}$ is the event attached to $x_{(i)}$ —it is not an independent ranking of the partition. Define a “cumulative event” $E(i) = \bigcup_{j=1}^i E_j$. Presumably, there is some (probability) measure m such that if $x_i > x_j$, then $m(E_i) > m(E_j)$ ²⁴. The rank dependant utility (RDU) procedure introduced by Quiggin (1982) is as follows²⁵:

The class of RDU representations involve proving from the axioms the existence of an order-preserving, utility function U over pure conse-

²¹Blackwell (1951) characterized a similar space.

²²See (Horton, 2004, pg. 20)

²³ (Loewenstein and Prelec, 1991, pg. 351) report that experiments show that “to most persons, a deteriorating series of utility levels is a rather close approximation to the least attractive of all possible patterns, regardless of the nature of events that are being ordered”. Recently, Ingersoll (2008) also derived a non-monotonicity result, via “comparative statics” of Tversky and Khaneman (1992) probability weighting function, when he showed that TK92 PWF is not increasing for all parameter values. So that it can assign negative weights in some cases. Nonetheless, the selected permutation is an artifact of the comonotonic probability distribution of prospects. In this paper, we embed the probabilities in outcome space. So that when agents rank outcomes they ineluctably rank probabilities *a fortiori*, and invariably produce a weak monotonic ranking of probabilities. While our theoretical procedure does not negate the Loewenstein-Prelec-Ingersoll critique, it does imply that in this paper ranked outcomes is a sufficient condition for weak comonotonicity. Inasmuch as the probability space is incomplete, i.e., subprobability measures are incomplete, violations of Tversky and Khaneman (1992) comonotonicity are accomodated by our embedding scheme.

²⁴Ironically, this definition could be internally inconsistent if any of the sets contain a compact set. Thus, E is presumably compact.

²⁵See Komarova, N. L, Jameson, K. A., & Narens, L., Theory of Measurement, *mimeo, mutatis mutandis*. In Blume, L. & Durlauf, S. N. (Eds.) Palgrave Dictionary of Economics.

quences and gambles and, in general, non-additive weighting function w over the chance events such that

$$U(x_1, E_1; x_2, E_2; \dots; x_n, E_n) = \sum_1^n U(x_i)[w(E(i) \cup E(i-1)) - w(E(i-1))]. \quad (*)$$

Note that the weighting function is essentially the incremental impact of adding E_i to $E(i-1)$. When w is finitely additive, i.e., for disjoint A and B , $w(A \cup B) = w(A) + w(B)$, then Equation (*) reduces to subjective expected utility (SEU). If there is a unique consequence e , sometimes called a reference level and sometimes taken to be no change from the status quo, then the consequences and gambles can be partitioned into gains, where $x_i \succeq e$, and the remainder, losses. In such cases, usually it follows from the assumptions made that $U(e) = 0$ and, usually, the weighting functions are sign dependent (i.e., their form depends on whether their consequences are positive with respect to e or negative). Also, the RSDU representation includes cumulative prospect theory (Tversky and Khaneman (1992)) as a special case having added restrictions on both U and w .

To evaluate the sum in (*) (Schmeidler, 1989, pg. 580) used a Choquet (1955) integral²⁶ subsequently adopted by Tversky and Khaneman (1992). However, a cursory inspection of that approach shows that it is averaging over (weighted) jumps in the probability weight function. Therefore, any “integral devised for [that special] purpose”²⁷; i.e., for averaging over utility or value functions to get an expected value, should yield satisfactory estimates of the sum²⁸.

²⁶see Appendix A.

²⁷See McShane (1963).

²⁸In fact, Schmeidler states “[i]ntuitively, uncertainty aversion means that “smoothing” or averaging utility distribution makes the decision maker better off”, *id* at 582.

2.1 Risk, Uncertainty, Rank Dependent Utility, and Choquet Capacity

To address uncertainty about probabilities, TK92 defined a "state choice" function $f : S \rightarrow X$ where S is a state space, and $X \subset \mathbb{R}$ is a space of outcomes or consequences; let A_i be a measurable event in S ; defined a prospect f by a sequence of pairs (x_i, A_i) ; assigned a number $V(f)$ to the prospect f ; and modified Quiggin's (1982) rank dependent utility for risk assessment as follows

$$\pi_n^+ = w^+(A_n), \pi_{-m}^- = w^-(A_{-m}) \quad (2.2)$$

$$\pi_i^+ = w^+(A_i \cup \dots \cup A_n) - w^+(A_{i+1} \cup \dots \cup A_n), 0 \leq i \leq n-1 \quad (2.3)$$

$$\pi_i^- = w^-(A_{-m} \cup \dots \cup A_i) - w^-(A_{-m} \cup \dots \cup A_{i+1}), 1-m \leq i \leq 0 \quad (2.4)$$

Here π_i is an unobservable probability weight induced by the nonlinear weight function w . The quantities w^+ and w^- are the probability weights assigned to gains and losses, respectively. Technically, they are generalized measures of the commensurate set which is presumed compact. By assigning

$$V(f^+) = \sum_{i=0}^n \pi_i^+ v(x_i) \quad \text{and} \quad V(f^-) = \sum_{i=-m}^0 \pi_i^- v(x_i) \quad (2.5)$$

they defined

$$V(f) = \sum_{i=-m}^n \pi_i v(x_i) \quad (2.6)$$

They concluded that if the prospect $f = (x_i, A_i)$ is qualified by a "probability distribution" $p(A_i) = p_i$, then for an aggregate of measurable events $\{A_i\}_{i \in \mathcal{I}}$, for

some index \mathcal{I} , it can be viewed as a probabilistic or risky prospect (x_i, p_i) . In particular,

$$\pi_i = w^+ \left(\sum_{i=0}^{i=n} p_i \right) - w^+ \left(\sum_{i=1}^{i=n-1} p_i \right) \quad (2.7)$$

for gains, and vice versa for losses with w^+ replaced by w^- and the commensurate probability sums defined over $-m \leq i \leq 0$. Further, (Tversky and Khaneman, 1992, pg. 301) states that “the decision weight associated with an outcome can be interpreted as the marginal contribution of the respective event, defined in terms of capacities W^+ and W^- . That definition is equivalent to one in which a probability distribution function F , estimated by F_n , has discontinuities of the second kind. So that if $F_n^+ = \sum_{i=0}^{i=n} p_i$ and $F_n^- = \sum_{i=-1}^{i=n} p_i$, then

$$\pi_n = w^+(F_n^+) - w^+(F_n^-) \quad (2.8)$$

This is the *sui generis* of Quiggin (1982) “anticipated” or “rank dependent” utility theory where prospects are ranked and then the “expected” value is derived by summing over the product of ranked utility and probability weights²⁹.

²⁹See Appendix A for Choquet integral representation.

3 A SPECTRAL THEORY OF VALUE FUNCTIONAL DECOMPOSITION

Of particular importance to this paper is TK92 dichotomous value function³⁰

$$v(x) = \begin{cases} x^\alpha & \text{if } x \geq 0 \text{ for gains,} \\ (-\lambda)(-x)^\beta & \text{if } x < 0 \text{ for losses} \end{cases} \quad (3.1)$$

where λ is a loss aversion index, and α and β are shape parameters. The *invariant* value functional $V(f)$ is represented by the integral

$$V(f) = \int_{x \in X} v(x) w(dF)(x) \quad (3.2)$$

Technically, $V(f)$ is a functional of the probability weight function $w(dF)$, and it should probably be written as $V_f(w)$ for reasons that will be apparent in the sequel.

Definition 3.1 (Radon measure). (Hewitt and Stromberg, 1965, pg. 114).

Let $\mathcal{B}(X)$ be the space of all bounded functions defined on a nonempty set X . Let $\mathcal{C}_0(X) \subset \mathcal{B}(X)$ be the subspace of continuous functions with compact support in X , i.e. if $K \subset X$ is compact, then for some measure μ we have $\mu(X \setminus K) < \varepsilon$ for some $\varepsilon > 0$. A Radon measure is a nonnegative linear functional I defined on $\mathcal{C}_0(X)$ if for $f, g \in \mathcal{C}_0(X)$ we have:

- i. $I(f + g) = I(f) + I(g)$;
- ii. $I(\alpha f) = \alpha I(f)$, for some scalar α ;

³⁰This section is excerpted from Cadogan (2010)

iii. $I(f) \geq 0$, if $f \in \mathcal{C}_0^+(X)$.

Furthermore, instead of the Choquet integral representation, we posit that $w(dF)$ is a Radon measure³¹ ψ . We present that claim in a formal proof in the sequel. Specifically, we have the following standard definition for essentially bounded functions, and averaging operators that involve them.

Definition 3.2 (Essentially bounded value functions). An essentially bounded value function is defined on a subset of the reals with Radon measure zero as follows

$$\text{ess sup } v = \inf\{a \in \mathbb{R}_+ \mid \psi(\{x : v(x) > a\}) = 0\}$$

Definition 3.3 (Averaging Operator). (Rota, 1960, pg. 53). Let A be a linear operator on the space of Lebesgue integrable functions $L^p(X, \mathfrak{B}(X), m)$, for some measure m , $1 \leq p \leq \infty$. Then A is an averaging operator if it has the following three properties:

i. A is a contraction mapping for $v \in L^p(X, \mathfrak{B}(X), m)$

$$\int_X |(Av)(x)|^p m(dx) \leq \int_X |v(x)|^p m(dx) \quad (3.3)$$

ii. Let $v, h \in L^p(X, \mathfrak{B}(X), m)$ and $y = \text{ess sup } h$. Then

$$(Av)(x)(Ay)(x) \in L^p(X, \mathfrak{B}(X), m), \text{ and } A(yv) = (Ay)(Av)$$

iii. Let $1 \in X$, then $A1 = 1$

³¹It can be shown that Choquet integrals are functionally equivalent to Stone-Daniell integrals which in turn are based on local Radon integrals. See Richards (1959) and Thomas and Volčič (1989).

Theorem 3.1 (Radon Integral of Value Function). *Let S be a locally compact state space, with σ -field of Borel subsets \mathfrak{S} , and Ω be a sample space, i.e. state of nature, attached to S , with σ -field of Borel subsets \mathfrak{D} ; and μ be a probability measure on Ω , and $(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)$ be the associated probability measure space. Let $L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)^*$ be the dual space of linear functionals defined on the space of Lebesgue integrable functions $L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)$. For any set $B \in \mathfrak{S}$ and $\omega \in \Omega$ let $\chi_B(s, \omega) \in L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)$ be a stochastic choice function, i.e. characteristic function, defined by $\chi_B(s, \omega) = 1$, if $s \in B$ and $\chi_B(s, \omega) = 0$, otherwise. Let V be a value function operator so that $V : L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu) \mapsto L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)^*$. Let $\psi(B, \omega)$ be a random set function and assign $V(\chi_B(s, \omega)) = \psi(B, \omega)$. For any function $f \in L^\infty(S \times \Omega, \mathfrak{S} \times \mathfrak{D}, \mu)$*

$$V(f) = \int_S f(s, \omega) \psi(ds, \omega) \quad (3.4)$$

$$\|V\| = \sup_{\text{ess sup } |f(s, \omega)| \leq 1} \left| \int_S f(s, \omega) \psi(ds, \omega) \right| \quad (3.5)$$

where $\psi(ds, \omega)$ is a Radon measure on $S \times \Omega$.

Proof. See Appendix E. □

Remark 3.1. It is clear that $\chi_B(s, \omega)$ is a discrete stochastic choice function of the choice set of “attributes” B . Implicit in the theorem is the “attached” probability p , say, when $\chi_B(s, \omega) = 1$ and $1 - p$ when $\chi_B(s, \omega) = 0$. The logit function $f(s, \omega) = \log\left(\frac{p}{1-p}\right)$ is admissible under Radon-Nikodym Theorem. So the class of logit and probit models is subsumed by this theorem.

Remark 3.2. The measure $w(dF)$ in equation (3.2) is a special case of the more abstract measure $\psi(ds, \omega)$. To avoid technical pedantics, we assumed that S is locally compact so that the Radon measure is well defined.

Accordingly, let T_ν be an averaging operator³² for a value function ν defined on the domain D_{T_ν} . This operator is consistent with the various “invariance principles” for prospects enunciated earlier by Luce, Prelec and al-Nowaihi, et al. Let ν_g and ν_ℓ be the parts of ν that constitute the value function for gains and losses, respectively, and T_g and T_ℓ be their corresponding averaging operators. So that $\nu_g \in D_{T_g}$ and $\nu_\ell \in D_{T_\ell}$. Undeniably, TK92 suggests the existence of a *binary operation* \oplus on gains and losses based on the following axiomatic foundations.

3.1 Binary conditions for value operator

A binary operator \oplus is regular if it is:

Assumption 3.2. (1) associative, i.e., $(a \oplus b) \oplus c = a \oplus (b \oplus c)$;

(2) distributive, i.e., $m(a \oplus b) = m.a \oplus m.b$; and the

(3) identity exists for $a \oplus -a = 0$. We might also add,

(4) reflexivity $\forall a \in \mathfrak{M}_\nu, a \oplus a \in \mathfrak{M}_\nu$,

(5) completion for $a, b \in \mathfrak{M}_\nu$ either $a \oplus b \in \mathfrak{M}_\nu$ or $b \oplus a \in \mathfrak{M}_\nu$,

(6) transitivity $a, b, c \in \mathfrak{M}_\nu$ and $a \oplus b \in \mathfrak{M}_\nu$ and $b \oplus c \in \mathfrak{M}_\nu$ implies $a \oplus c \in \mathfrak{M}_\nu$,

and

³²See Moy (1954) and Rota (1960).

(7) antisymmetric $a \oplus b \in \mathfrak{M}_v$ and $b \oplus a \in \mathfrak{M}_v \Rightarrow a = b$. Cf. Lahiri (2004).

Remark 3.3. Banach algebra for the averaging operators require a different inverse relationship than that in 3.2 (3) above.

According to TK92 theory, $v = v_g \oplus v_\ell$, from which it follows that the averaging operation $T_v = T_g \oplus T_\ell$. Technically, T_g and T_ℓ are reductions of T_v . In which case

$$(T_v v)(x) = (T_g v \oplus T_\ell v)(x) = V(f) = \int_{x \in X} v(x) w(dF)(x) = T_v(v_g \oplus v_\ell)(x). \quad (3.6)$$

By definition, T_v is a bounded linear operator on X since the space of outcomes is bounded. So the binary operation is well defined. Thus, by the distributive property of \oplus , $T_g v = T_v v_g$ and $T_\ell v = T_v v_\ell$. However, introduction of the loss aversion parameter λ in the TK92 model, together with the invariance axioms above, imply that $T_g v = T_v v_g = v_g$ and $T_\ell v = T_v v_\ell = -\lambda v_\ell$. Specifically, TK92 implies that the operator for gains is isometric while the operator for losses is skewed. These results are summarized by the following theorems.

Theorem 3.3 (Gain Loss Risk Operators). *Assume that the invariance principle holds. Let v be a Tversky and Khaneman (1992) value function, and v_g, v_ℓ be the gain and loss components of v , respectively. Let \mathfrak{M}_v be the space of value functions endowed with a binary operation \oplus such that $v = v_g \oplus v_\ell$. Let T_v be an averaging operator with domain D_{T_v} and $v \in D_{T_v}$. Let λ be a loss aversion index. Then the following holds:*

(i) $T_v = T_g \oplus T_\ell$

(ii) T_g is an isometric operator

(iii) T_ℓ is a skewed operator

Proof. See Appendix E. □

Theorem 3.4 (Spectral Theorem). *Let \mathcal{T} be the class of Tversky and Khaneman (1992) value functions and T_v be a linear averaging operator absolutely continuous with respect to a Radon measure $w(dF)$ on the space of distribution functions $F \in \mathfrak{M}_F$. Let \oplus be a regular binary relation on the domain D_{T_v} of T_v such that if T_g is an averaging operator for gains, and T_ℓ an averaging operator for losses, each absolutely continuous with respect to the Radon measure $w(dF)$, then $T_v = T_g \oplus T_\ell$. So for any value function $v \in \mathcal{T}$ the spectrum Λ_{T_g} for gains is comprised of the eigenvalue $\lambda = 1$; and the spectrum Λ_{T_ℓ} for losses is characterized by the set*

$$\Lambda_{T_\ell} = \{\lambda | (T_\ell + \lambda E)v = 0, v \neq 0\} \quad (3.7)$$

where E is the identity operator. Moreover, the spectrum of T_v is given by

$$\Lambda_{T_v} = \{1\} \cup \Lambda_{T_\ell} \quad (3.8)$$

Specifically, T_g is isometric and T_ℓ is skewed.

4 PROSPECT TOPOLOGY

4.1 Topological lifting of prospect space

4.1.1 Preliminaries

The proverbial elephant in prospect theory space is metrizability³³ of outcomes. For implicit in an agent's ranking of outcomes, based on mapping a choice function f from state space S into outcome space X , is her ability to measure and hence rank outcomes in X via some metric. Before we embark on the journey towards topological lifting of value functions, we include a few definitions to clarify terminology.

Definition 4.1 (Hausdorff space). (Dugundji, 1966, pg. 137) .

A space Y is Hausdorff (or separated) if each of two distinct points have non-intersecting neighbourhoods. That is, whenever $p \neq q$ there are neighbourhoods $U(p)$, $V(q)$ such that $U(p) \cap V(q) = \emptyset$.

Definition 4.2 (Homeomorphism). (Dugundji, 1966, pg. 87).

A continuous bijective map, i.e., 1-1 and onto, $f : S \rightarrow X$, such that $f^{-1}X : \rightarrow S$ is called a homeomorphism (or bicontinuous bijection) and denoted by $f : S \cong X$. Two spaces are homeomorphic, written $S \cong X$ if there is a homeomorphism $f : S \cong X$.

Definition 4.3 (Embedding). (Dugundji, 1966, pg. 89).

If Z is any space, and if $f : X \rightarrow Z$ is a map establishing that X is homeomorphic

³³A topological space (Y, τ) is metrizable if its topology τ is one induced by a metric on the space Y . See (Dugundji, 1966, pg. 183).

to $f(X) \subset Z$, i.e. $X \cong f(X) \subset Z$, then f is called an embedding map of X into Z .

Definition 4.4 (Covering). (Munkres, 2000, pg. 164)

A collection A of subsets of a space X is said to cover X , or to be a covering of X , if the union of elements of A is equal to X . It is called an open covering of X if its elements are open subsets of X .

Definition 4.5 (Partition). .

Let \mathcal{J} be an index set. If $\{A_\alpha \mid \alpha \in \mathcal{J}\}$ is a covering of S , and if $A_\alpha \cap A_\beta = \emptyset$ whenever $\alpha \neq \beta$, then the family $\mathcal{P} = \{A_\alpha \mid \alpha \in \mathcal{J}\}$ is called a partition of S .

Definition 4.6 (Covering spaces). (Munkres, 2000, pg. 336).

Let $\{A_\alpha\}$ be a partition of S , $\alpha = 1, 2, \dots, n$ and $f : S \rightarrow X$ be a continuous onto mapping of S . Let $U \subset X$ be an open set in X . If $f^{-1}(U) = \bigcup_{\alpha_k} A_{\alpha_k}$ for some indexes $\alpha_k \in 1, 2, \dots, n$ such that for each α_k the restriction of f to A_{α_k} , i.e., $f|_{A_{\alpha_k}}$, is a homeomorphism of A_{α_k} onto U , then U is said to be evenly covered by f . The collection $\{A_{\alpha_k}\}$ will be called a partition of $f^{-1}(U)$ into slices.

Definition 4.7 (Covering map). (Munkres, 2000, pg. 337).

Let $f : S \rightarrow X$ be continuous and onto. If every outcome $x \in X$ has a neighbourhood U that is evenly covered by f , then f is called a covering map, and S is called a covering space.

Definition 4.8 (Lifting). (Munkres, 2000, pg. 342).

Let $v : X \rightarrow R$ be a mapping of real valued value functions (v) on outcome space.

If $g : S \rightarrow \mathbb{R}$ is a continuous real valued mapping defined on the state space S , a *lifting* of g is a map $f : S \rightarrow X$ such that $v \circ f = g$.

4.1.2 Commutative map of Prospect Theory's Liftings

To keep track of the myriad liftings and composite maps in the sequel we modify the old adage to a commutative map is worth a thousand words. The dia-

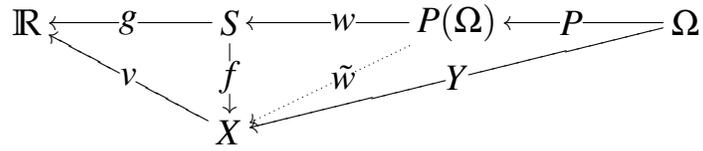


Figure 1: Commutative Map of Prospect Theory's Liftings

gram in Figure 1 plainly shows that the stochastic choice map f is a lifting of the imputed direct map $g = v \circ f$ from state space S to the reals \mathbb{R} . Compare Tversky and Khaneman (1992) mapping scheme in the introduction section 1 of this paper. Additionally, the composite direct map $g \circ (w \circ P) = v \circ Y$, from sample space Ω to the reals \mathbb{R} , is a lifting of Y . In that case, for a given outcome $x \in X$, the map $v(Y(x))$ is a functional. Thus, averaging over that quantity gives rise to an averaging operator. Further, the probability weight function w is a lifting of the direct map $\tilde{w} = f \circ w$ from $P(\Omega)$ to S . Perhaps most important, the composite map $w \circ P$ is a lifting of the direct map $Y = f \circ (w \circ P)$ from sample space Ω to outcome space X . The stochastic choice functions in extant literature, see e.g., Debreu (1958) and McFadden (1974), considers a mapping $P : \Omega \rightarrow S$. But not the intermittent composite mapping $w : P(\Omega) \rightarrow S$ which embeds probability weights in state space S ,

and indirectly in X through f or directly through \tilde{w} . The commutative map plainly shows that the introduction of probability weighting map w should be incorporated in any stochastic choice map $f : S \rightarrow X$. That is the *sui generis* of this paper.

The mapping Y has the following interpretation. Since $Y : \Omega \rightarrow X \Rightarrow Y(\Omega) \subseteq X$, there exists a lottery $\{(x_1, p_1), (x_2, p_2), \dots, (x_n, p_n)\}$ such that $Y(\omega)$ takes the values (x_1, x_2, \dots, x_n) with corresponding joint probability distribution (p_1, p_2, \dots, p_n) . So that for a given realization of outcomes, $Y_i(\omega) = f \circ (w \circ p_i) = f(w(p_i)) = (x_i, p_i) \cong x_i(p_i)$, for some index i . Additionally, let F_Y be the probability distribution function of Y . So that for rank ordered Y we have the relation $\pi_y = w(F_Y^+(y)) - w(F_Y^-(y))$ as the probability weight assigned to the marginal lottery at the jump of F . In any event, the commutative diagram plainly shows how probabilities and or probability weights are embedded in outcome space X . The rest of this paper constitutes analytic proofs of these facts, with ensuing applications in abstract Hilbert spaces.

4.1.3 Hoffman-Jorgensen lifting on Lebesgue spaces

Hoffman-Jorgenson (1971) described a lifting by defining a measure space (S, Σ, μ) with Σ_μ a Lebesgue extension of Σ ; a function space $L_\infty(\mu) \subset X$ of essentially bounded functions with norm $\|f\|_\infty^\mu = \mu - \text{ess sup}_{s \in S} |f(s)|, \forall f \in L_\infty(\mu)$. He also, introduced the set $B(S, \Sigma_\mu)$ of all Σ_μ -measurable bounded real functions on S , endowed with the norm $\|f\|_\infty$. What is more, he introduced an ordering \leq on $L_\infty(\mu)$ and $B(S, \Sigma_\mu)$ so that for $s \in S$, we have $f(s) \leq g(s) \mu$ a.e., and he

defined a map $\mathfrak{L} : L_\infty(\mu) \rightarrow B(S, \Sigma_\mu)$ so that (\mathfrak{L}, \leq) is a vector lattice. In particular, Hoffman-Jorgenson defined a lifting as a linear operator $\ell : L_\infty(\mu) \rightarrow B(S, \Sigma_\mu)$ with the additional properties that $h = fg \Rightarrow \ell(h) = \ell(f)\ell(g)$ and $\ell(f) = f$. That definition is consistent with (Quiggin, 1982, pg. 326) and functionally equivalent to our probabilistic approach *a fortiori*. Specifically, in the composite function h Hoffman-Jorgensen's ℓ -operator is isometric with respect to the lifting element f . A result consistent with our value function operator for gains.

4.1.4 Kernel function generation of probability and choice function sequences

An alternative approach to the embedding procedure used here is to define a kernel $k : S \times \Omega \rightarrow X$ and let $k_n(s, \omega) = n^2 k(ns, n\omega)$ where $k(s, \omega)$ has the following properties with respect to measures ν on S and μ on Ω :

- (i) $\int_S \int_\Omega k_n(s, \omega) \nu(ds) \mu(d\omega) = 1$
- (ii) $\int_S \int_\Omega k_n(s, \omega) \nu(ds) \mu(d\omega) \leq K_0$ for some constant K_0 ;
- (iii) $\lim_{n \rightarrow \infty} \int_{S - \delta S} \int_{\Omega - \delta \Omega} k_n(s, \omega) \nu(ds) \mu(d\omega) = 0$

Property (iii) implies that the tail of the kernel function vanishes. For a given “choice function” $f : S \rightarrow X$, and probability measure P on Ω , and an operator K we can generate a sequence of probability measures

$$P_n(E) = (Kf)(\omega) = \int_E k_n(s, \omega) f(s) \nu(ds), \quad E \in \mathfrak{S} \quad (4.1)$$

or a sequence of choice (set) functions

$$f_n(A) = (KP)(s) = \int_A k_n(s, \omega) P(\omega) \mu(d\omega), \quad A \in \mathfrak{D} \quad (4.2)$$

See (Bochner, 1955, pp. 1-2). Equation (4.1) shows that a sequence of probability measures can be generated by transformation of the choice function with an appropriate kernel. Similarly, equation (4.2) shows that a sequence of choice functions can be generated by transforming probability measures via kernel estimation. In fact, the stochastic choice function is an operator K on probability measures P . Of course, the whole process could be simplified by using Lebesgue measure instead of the general measures ν and μ . By truncating k_n in (i), in equation 4.1 P_n becomes a sub-probability measure on Ω . Similarly, we can generate estimates of choice functions in equation 4.2 for pre-specified probability distributions. This kernel approach provides an independently important mechanism for generating sequences of choice functions and probability measures. Nonetheless, in this paper we emphasize the *topological lifting, embedding, and concomitant weak comonotonic feature of our proposed topology*.

4.2 Existence of mapping from state space to reals as value function composite

In this sub-section, we provide justification for the research question concerning the imputed map from state space to the reals, and our lifting hypothesis. The

lemma and existence theorem that follow are modifications of (Blumenthal and Geertor, 1968, pp. 6-7).

Lemma 4.1. *Let $\mathfrak{M}_{v \circ f}$ be a vector space of real valued value functions on $S \times \Omega$, and $\mathfrak{S} \times \mathfrak{D}$ be the corresponding product σ -field of Borel subsets. Assume that*

i. $1 \in \mathfrak{M}_{v \circ f}$

ii. *if $\{v_n\}$ is an increasing sequence of value functions in $\mathfrak{M}_{v \circ f}$ such that $v = \sup_n v_n$ is bounded, then $v \in \mathfrak{M}_{v \circ f}$*

iii. *The product $\prod_{i < \infty} (\chi_{A_i} \circ f)(s, \omega)$ is in $\mathfrak{M}_{v \circ f}$ for $A_i \in \cap f_i^{-1}(A_i)$*

Then $\mathfrak{M}_{v \circ f}$ contains all real valued bounded functions in the product sub- σ -field $\mathfrak{S}_i \times \mathfrak{D}_i$.

Theorem 4.2 (Existence of Mapping from State space to Reals). *Let $S \times \Omega$ be the product space of state space and states of nature, and $\mathfrak{S} \times \mathfrak{D}$ the corresponding product σ -field. Let $f : S \times \Omega \rightarrow X$ where X is outcome space, and $\mathfrak{B}(X)$ is the σ -field of Borel measurable subsets of X . Then there exist a $\mathfrak{S} \times \mathfrak{D}$ -measurable mapping $\phi : S \times \Omega \rightarrow \mathbb{R}$ if and only if there exist a real valued $\mathfrak{B}(X)$ -measurable positive value function v such that $\phi = v \circ f$.*

Proof. See Appendix E. □

4.3 Weak comonotone probability embedding

Let \mathfrak{S} be a σ -field of Borel subsets of the state space S , and endow the measure space (S, \mathfrak{S}) with a sub-probability measure P such that $0 < P(S) \leq 1$. Thus,

(S, P) is a metric space. If \mathcal{P} is a partition of S , then each [nonempty] subset $A_i \in \mathcal{P}$, $i = 1, 2, \dots, n$ is comprised of elementary events, i.e., states, $s \in S$. However, (Tversky and Khaneman, 1992, pg. 300) defines a mapping $f : S \rightarrow X$ such that $f(s) = x \in X$; and then state that “[a] prospect f is then represented as a sequence of pairs (x_i, A_i) , which yields x_i if A_i occurs, where $x_i > x_j$ iff $i > j$ ”. In that definition, it is not clear whether the sequence of x ’s are elements of a vector and or an array. However, taken literally, it implies that if event A_i occurs, then the other $n - 1$ events in \mathcal{P} did not occur, i.e., events are mutually exclusive and do not occur jointly. Nonetheless, in the sequel we propose a scheme to ensure metrizable of X and its co-monotonicity with P .

Lemma 4.3 (Attaching map). ·

Let f be a homeomorphic mapping, i.e., it preserves topological properties in its domain and range space. Let \mathfrak{D} be the σ -field of Borel measurable subsets of the sample space Ω . Similarly, let \mathfrak{S} be the σ -field of Borel measurable subsets of the state space S . Let P be a subprobability measure defined on Ω . Define the composite map

$$f \circ P : (S \times \Omega, \mathfrak{S} \times \mathfrak{D}) \rightarrow X$$

so that for $A_i \in \mathfrak{S}$, and $x_i \in X$ we write $(f \circ P)(A_i, \omega) \equiv (f \circ P)(A_i)(\omega) = x_i$. Then P is an attaching map, i.e., it attaches Ω to S .

Proof. See Theorem B.3 on attaching spaces in the appendix for justification of this extended composite map with attachment. See also, Musial and Macheras

(2002); Musial et al. (2009) for proof of lifting on extended topological probability product spaces. □

Remark 4.1. The quantity $(f \circ P)(A_i)(\omega)$ is a Borel measurable random variable assigned the outcome x_i . Vizly, it is a random operator. Specifically, by virtue of the discrete partition $S = \bigcup_{i=1}^n A_i$ the composite map is a “point function” defined on S into X , and the probability law of $f \circ P$ is a probability measure on the measurable outcome space $(X, \mathfrak{B}(X))$.

To proceed with the weak embedding proposition we need the following important

Lemma 4.4 (Homeomorphic extension). *Let \mathcal{J} be an index set. If $\{A_\alpha | \alpha \in \mathcal{J}\}$ is a partition of S , and if for each $\alpha \in \mathcal{J}$ there is a given map $f_\alpha : A_\alpha \rightarrow X$, then there exist a unique $f : S \rightarrow X$ which is an extension of f_α .*

Proof. See (Dugundji, 1966, Corollary 6.8, pg. 13). □

Thus, the extended composite map in Lemmas 4.3 and 4.4 leads to the following

Proposition 4.5 (Weak embedding). *Let f be a homeomorphic mapping on the state space S , and P be an attaching map for the sample space Σ . Then the probability of event A_i is weakly embedded in X under the mapping $f \circ P : S \times \Omega \rightarrow X$.*

Proof. The homeomorphism hypothesis $S \cong f(S) \subset X$ implies that f is an embedding map of S into X . Let $\mathcal{J} = \{1, 2, 3, \dots, n\}$ be an index set. Since $A_i, i \in \mathcal{J}$ is an element of the partition of S , it follows from Lemma 4.4 that $A_i \cong f(A_i)$ is a

local homeomorphism of the extended map f . Let $P(A_i) = p_i$. Since every point p_i is put in 1-1 correspondence with the set A_i in the partition \mathcal{P} , the map P is homeomorphic. So that $P^{-1}(p_i) = A_i \cong f(A_i)$. Since both P , and f are homeomorphic, every point p_i can also be put in 1-1 correspondence with the composite $(f \circ P)(A_i(\omega))$. So by Lemma 4.4 the composite map is also homeomorphic. Specifically,

$$p_i \leftrightarrow f(P(A_i)(\omega)) \Rightarrow p_i \cong f(P(A_i)) = (f \circ P)(A_i(\omega)) \subset (f \circ P)(S)$$

Therefore, by 4.3 p_i is embedded in $(f \circ P)(S)$. Let $V(p_i)$ be an open neighbourhood of p_i . By definition, S is a Hausdorff space, and P induces a discrete topology. Thus, we have $P(A_i(\omega)) \cong V(p_i)$. By Lemma 4.4 P is a unique extension of this local homeomorphism. Additionally, each of the neighbourhoods $V(p_i), i = 1, 2, \dots, n$ are in 1-1 correspondence with the A_i 's through P . So that for disjoint A_i we have $\bigcup_{i=1}^n V(p_i) \cong P(\bigcup_{i=1}^n A_i) = P(S)$. Thus, by 4.6 $\{V(p_i), i \in \mathcal{J}\}$ constitute the covering spaces, induced by P , in correspondence with the partition \mathcal{P} of S . In other words, $\{V(p_i), i \in \mathcal{J}\}$ is a partition of unity³⁴ weakly homeomorphic to S through P . Since $S \cong f(S)$, we get $P(S) \cong (f \circ P)(S) \subset X$, where the latter subset relation exist by Lemma 4.4 and definition of the composite map. Thus, by 4.3 the probability $P(A_i)$ is weakly embedded in X through the neighbourhoods $V(p_i)$. □

³⁴For instance, $\sum_{\alpha \in \mathcal{A}} k_\alpha(p) = 1$ for some Hausdorff space Y covered by $\{V(p_i), i \in \mathcal{J}\}$ where $p \in V(p_i)$, and $k_\alpha : Y \rightarrow [0, 1]$. See (Dugundji, 1966, pg. 170).

Corollary 4.6. *The probability of event A_i is weakly embedded in X under the mapping $f \circ w \circ P : S \times \Omega \rightarrow X$*

Proof. Since w is an isomorphic map, substitute $w(P)$ for P in Proposition 4.5 and the proof follows through as before. \square

As new information becomes available agents presumably re-rank their prospects. This process puts restrictions on the function f . In particular, f must preserve monotonicity in P . These restrictions are weakened by embedding a probability weight function $w(P)$ in outcome space X . Thereby inducing a comonotonic ranking on X *a fortiori* through the attached embedding. That is, if the composite map in Proposition 4.5 is transformed further, by replacing P with $w(P)$, we get $(f \circ w \circ P)(A_i, \omega) = f \circ w(P(A_i, \omega))$. Whereupon by 4.6 the mapping $f \circ w \circ P : S \times \Omega \rightarrow X$ embeds $w(P)$ in X since w is isomorphic—by definition. Thus, X is weakly metrizable, i.e., ranked, because the composite map includes a probability weight metric for ranked probabilities. Furthermore, the probability law of the composite $f \circ w \circ P$ is a probability measure on $(X, \mathfrak{B}(X))$. This argument is summarized by the following

Proposition 4.7 (Weak comonotonic ranking of outcomes). *Let X be outcome space; Ω be sample space; w be a probability weight function; P be a probability measure on Ω , and f be a stochastic choice function. The composite map $f \circ w \circ P : S \times \Omega \rightarrow X$ embeds probability weights $w(P)$ into outcome space X . Thus, the pair $(X(w(P(A, \omega))), \succeq)$, $A \in \mathcal{P}$ is a weak comonotonic ranking of X based on some*

ordering \succeq on \mathcal{P} .

Probablistically, $f(w(P(A_i, \omega)))$, $A_i \in \mathcal{P} \cap \mathfrak{S}$ is functionally equivalent to an isomorphism of the value functional $V(f)$ obtained by summing over the partition \mathcal{P} of S . To see this, consider the discrete topology in outcome space X . Let T be an averaging operator over the space of probability weights. Let \mathfrak{M}_v be a sufficiently dense class of value functions. So that for a value function v defined on X , the operation T performed on v results in an average value function in \mathfrak{M}_v . By virtue of TK92 reference point hypothesis, the point zero is in the spectrum of T , i.e. $0 \in \Lambda_T$. So the spectrum contains a convergent sequence $\lim_{n \rightarrow \infty} \lambda_n = 0$ where $\{\lambda_n\}_{n \geq 1} \in \Lambda_T$. This implies the existence of a convergent sequence $\{v_n\} \in \mathfrak{M}_v$ since $Tv_n = \lambda_n v_n$. Specifically, for every open neighbourhood $N(v_0) \subset \mathfrak{M}_v$; metric ρ defined on \mathfrak{M}_v ; and $\varepsilon > 0$, there exist $\delta > 0$, such that whenever $v \in N(v_0)$ then $\rho(v, v_0) < \delta$, and $\|Tv - Tv_0\| < \varepsilon$. Thus, the family of maps $\mathcal{A}(T) \subset \mathfrak{M}_v^{\mathfrak{M}_v}$ are equicontinuous. We summarize this in the following

Lemma 4.8. *Let $\mathfrak{M}_v^{\mathfrak{M}_v}$ be the space of all maps from \mathfrak{M}_v into itself, and $\mathcal{A}(T)$ be the family of averaging maps T in $\mathfrak{M}_v^{\mathfrak{M}_v}$. Let Λ_T be the spectrum of T . If $0 \in \lambda_T$, then $\mathcal{A}(T)$ is equicontinuous. Furthermore, $\{v_n\}$ is a weakly convergent sequence induced by Λ_T .*

4.4 Almost sure isomorphisms of lifted choice functions in outcome space

This subsection is based on a fairly lengthy proof of almost sure convergence between lifted stochastic choice functions with embedded probability weights, and isomorphic mappings of value functions. The proof is based on topological considerations in outcome space. It is broken up in three parts motivated by (1) Prokhorov's Theorem for weak compactness of probability measures in function space; (2) properties of compact operators; and (3) Skorokhod's Theorem for convergence of random functions. We begin with the observation that since $T : \mathfrak{M}_v \rightarrow \mathfrak{M}_v$, the result $v^{-1}Tv$ is an isomorphic mapping of outcomes that induce a set $E_v = \{x | v^{-1}Tv = x\} \subset X$. However, by Corollary 4.6 there exist a set $E_f = \{x | f \circ w \circ P : S \times \Omega \rightarrow X\} \subset X$. Since f is homeomorphic, it is a closed map, see (Dugundji, 1966, Thm 12.2 pg. 89). Therefore, the set $E_v \cap E_f \neq \emptyset$. In particular, the closure $\overline{f(S)}$ in X subsumes $\bigcup_{v \in \mathfrak{M}_v} E_v$. That is, $\bigcup_{v \in \mathfrak{M}_v} E_v$ is dense in $f(S)$. So that $\bigcup_{v \in \mathfrak{M}_v} E_v = \overline{f(S)}$. Essentially, the proof is based on the fact that different sequences of outcomes in a closed dense subspace of X must converge to the same limit almost surely. This result is summarized as

Theorem 4.9 (Almost sure isomorphism). *Let \mathfrak{M}_v be the class of Tversky and Khaneman (1992) value functions, and T be an averaging operator over \mathfrak{M}_v with respect to a probability weight functional $w(dF)$, where F is a probability distribution function, and $T : \mathfrak{M}_v \rightarrow \mathfrak{M}_v$. Let S be a state space; Ω be a sample space attached to S and endowed with probability measure P ; X be the*

separable space of outcomes characterized by the composite homeomorphic map $f \circ w \circ P : S \times \Omega \rightarrow X$, and $\mathfrak{B}(X)$ be the Borel subsets of X . Then for $A_n \in \mathfrak{B}(X)$, we have for some stochastic operator $J = f$

$$\lim_{n \rightarrow \infty} \Pr \left\{ \sup_{A_n \in \mathfrak{B}(X)} \left| (J(w \circ P)(A_n, \omega) - (v^{-1}Tv)(A_n)(\omega)) \right| > \varepsilon \right\} = 0 \quad (4.3)$$

Proof. The hypotheses of the proposition created an ambient space which facilitates attacking the proof in three parts. Notationally, we use f instead of J for convenience. The idea behind the proof is as follows. First, we will use Prokhorov's Theorem, see (Ash, 2000, pp. 37-40), for weak convergence of probability measures, i.e., $P_n \Rightarrow P$, to show that $f \circ w \circ P_n \Rightarrow f \circ w \circ P$ in X . Technically, because P is a subprobability measure, we have vague convergence. See (Athreya and Lahiri, 2006, Ch. 9.2). This means that a convergent sequence of subprobability measures need not converge in the set containing the sequence. That fact implicates our proof. To get around that problem, we make the following strong

Assumption 4.10. *The space of probability measures is locally compact.*

This assumption creates an ambient space for *weak convergence of conditional probability measures*. See (Billingsley, 1995, pg. 371). Second, we will use (Akheizer and Glazman, 1961, pp. 46-47) weak compactness theorem for operators in Hilbert space to show that the sequence of averaging operators $Tv_n \Rightarrow Tv$ is compact in \mathfrak{M}_v . The sequence of mappings $\{v^{-1}Tv(A_n)\}_{n \geq 1}$, $A_n \in \mathfrak{B}(X)$ generate a sequence of outcomes in X . Third, according to Skorokhod's Theorem

in (Brieman, 1968, pg. 340) and Arzela-Ascoli Theorem in (Royden, 1988, pg. 169), since each sequence of composite maps and isomorphic averaging operators generate separate sequences of outcomes in a dense subspace of X (endowed with a sup-norm metric) they must do so in the same cluster set for consistency of choice for *well defined outcomes*. That cluster set is covered by an ε -net. And it is that ε -net that gives us the probability bounds because the closure of the cluster set is bounded by the base topology on the metricized space X .

Definition 4.9 (Tightness). (Billingsley, 1968, pg. 9)

Let Ω be a sample space, and \mathfrak{D} be the σ -field of Borel sets of Ω . A sequence $\{P_n\}$ of probability measures defined on Ω is *tight* if for every $\varepsilon > 0$ there exist a compact set $\mathcal{K} \in \mathfrak{D}$ such that $P(\mathcal{K}) > 1 - \varepsilon$. Moreover, for every $A \in \mathfrak{D}$ $P(A) = \sup_{\mathcal{K} \subset A} P(\mathcal{K})$.

Lemma 4.11 (Prokhorov's Theorem).

Let Ω be a sample space, \mathfrak{D} be a σ -field of Borel measurable subsets of Ω , and U be a compact subset of Ω . Let P_1, P_2, \dots be finite measures on (Ω, \mathfrak{D}) with corresponding distribution functions F_1, F_2, \dots . Further, let $\mathfrak{D}|_U$ be the restriction of \mathfrak{D} to U . Assume that $F_n(E) = P_n(\omega \in E)$ for $E \in \mathfrak{D}|_U$. Suppose that $0 \leq P_n(\Omega) < 1$ for all n . Then the sequence $\{P_n\}$ is tight iff it is relatively compact.

Proof. See (Ash, 2000, pg. 340). See also, (Billingsley, 1968, pp. 37-40). □

Lemma 4.12 (Akheizer and Glazman (1961)).

For the weak convergence of a sequence of value functions $\{v_n\}_{n \geq 1}$ it is necessary

and sufficient that:

(i) $Tv_n \Rightarrow Tv$ for $v \in \mathcal{K}$ dense in \mathfrak{M}_v ;

(ii) The sequence $\{v_n\}_{n \geq 1}$ is bounded.

Proof. See (Akheizer and Glazman, 1961, pg. 47). □

Lemma 4.13 (Skorokhod's Theorem).

Let $\{v_n(x)\}_{n \geq 1}$, $x \in X$, be a sequence of value functions such that $v_n \xrightarrow{\mathcal{D}} v$, i.e. v_n converges in distribution to v . Then for any countable set $\mathcal{K} \subset X$, there exist $\{\tilde{v}_n(x)\}_{n \geq 1}$, $x \in \mathcal{K}$, defined on a common space such that

(1) For each n , \tilde{v}_n and v_n have the same distribution

(2) For every x we have $\tilde{v}_n \xrightarrow{a.s.} \tilde{v}$.

Proof. See (Brieman, 1968, Thm 13.29, pg. 293). □

Part I.

Let Ω be a sample space that corresponds to the states of nature, S be a state space partitioned by $\mathcal{P} = \{A_i\}_{i=1}^n$, and P be a subprobability measure defined on S , so that $P_i = P(A_i)$ is the probability associated with the outcome x_i . Since P is a subprobability measure we can extend it to a a probability measure by

$$\tilde{P}_n = \begin{cases} P(A_n) & A_n \in \mathfrak{G} \\ 0 & A_n \notin \mathfrak{G} \end{cases}$$

But this extended probability function implies the existence of a truncated distribution or locally compact sets in the Skorokhod space $D[0,1]$ of functions with discontinuities of the second kind. Let $C(\{P\})$ be the cluster set of $\{P_n\}$, i.e. $C(\{P\}) = \{0\} \cup \{1\} \cup [a, b]$ where $a > 0$, $b < 1$. According to Bolzano-Wierstrass Theorem every point of $C(\{P\})$ is a limit point for the sequence $\{P_n\}$. See (Shilov, 1973, pg. 75). Thus, every Cauchy sequence $\{P_n\}$ has a convergent subsequence with limit point in $C(\{P\})$. Whereupon $C(\{P\})$ is closed and relatively compact. So that by Lemma 4.11, $P_n \Rightarrow P$. It follows from continuity of the probability weight function w and choice function f that the composite map $f \circ w \circ P_n \Rightarrow f \circ w \circ P$. See (Shilov, 1973, pg. 135).

Part II.

From Lemma 4.8 the family of operators $\mathcal{A}(T)$ on \mathfrak{M}_v is equicontinuous. Furthermore, by virtue of the convergent spectrum, T maps weakly convergent sequences of value functions into convergent sequences. Thus, by Lemma 4.12 T is a compact operator.

Part III.

Take the outcomes of a lottery as a countable set of points $\mathcal{K} \subset X$. Specifically, the set $\mathcal{K} \cap E_v$ is countable, where $E_v = \{x \mid v^{-1}Tv = x\} \subset X$. Similarly, for the partition $\bigcup_i A_i = S$, the probability $P(A_i)$ corresponds to a discrete outcome $x_i \in \mathcal{K} \cap E_v$. So for the set $E_f = \{x \mid f \circ w \circ P : S \rightarrow X\}$, the extended composite function $(f \circ w \circ P)(A_i) = x_i$ is in $\mathcal{K} \cap E_f$. Extend E_f to E_v by the relation $\{x \mid x \in$

E_f if $x \in E_v$, and $E_f = \emptyset$ otherwise} then we have a common space, call it E , restricted to \mathcal{K} , such that $E|_{\mathcal{K}} \subset X$. By construction of E_v and E_f , the set E is nonempty. It is clear that v_n converges to v in distribution for the spectrum Λ_T by virtue of compactness of T . We define an ε -net for \mathcal{K} by taking a lottery in $E|_{\mathcal{K}}$ and a point $x \in \mathcal{K}$ such that the metric $\rho_{E|_{\mathcal{K}}}(x, x_i) < \varepsilon 2^{-i}$. For any partition $\bigcup_i A_i = S$, let $\delta(A_i)$ be the diameter of A_i . The equicontinuity of $T \in \mathcal{A}(T)$ and continuity of v implies that $g = v^{-1}T v$ is at least a continuous isomorphism. By hypothesis, f is continuous. Therefore, the function $f - g$ is continuous. For $f, g \in C(X)$, define the metric $\rho(f, g) = \sup_{f, g \in C(X)} \|f - g\|$. So that by Lemma 4.13

$$\left| (f \circ w \circ P)(A_i) - (v^{-1}T v)(A_i) \right| = |f(w(P)(A_i) - g(A_i)| \quad (4.4)$$

$$\leq \sup_i |f(w(P)(A_i) - g(A_i)| \quad (4.5)$$

$$\text{Since } |w(P)(A) - \delta(A)| \leq \sup \delta(A) \quad (4.6)$$

$$\text{by continuity of } (f - g) \text{ we have } \left| (f \circ w \circ P)(A_i) - (v^{-1}T v)(A_i) \right| \quad (4.7)$$

$$\leq \sup_{f, g \in C(X)} \|f - g\| \sup_{1 \leq i \leq n} \delta(A_i) \quad (4.8)$$

By virtue of the ε -net \mathcal{K} , and $A_i \in E|_{\mathcal{K}}$, let $\sup_{1 \leq i \leq n} \delta(A_i) = \varepsilon 2^{-i}$. Thus

$$\lim_{n \rightarrow \infty} \Pr\left\{ \sup_{1 \leq i \leq n} \left| (f \circ w \circ P)(A_i) - (v^{-1}Tv)(A_i) \right| > \varepsilon \right\} \quad (4.9)$$

$$\leq 1 - \sum_{i=1}^{\infty} \sup_{f, g \in C(X)} \|f - g\| \varepsilon 2^{-i} \quad (4.10)$$

$$\leq 1 - \sup_{f, g \in C(X)} \|f - g\| \varepsilon \quad (4.11)$$

Since the function $f - g$ is continuous, and ε is arbitrary, choose

$$\varepsilon = \left(\sup_{f, g \in C(X)} \|f - g\| \right)^{-1} \quad (4.12)$$

Whereupon the right hand side of the inequality vanishes and we have $f \circ w \circ P = v^{-1}Tv$ almost surely. So we may write for $x \in E_v$

$$f \circ w \circ P + o_p(1) = v^{-1}Tv + o_p(1) = x \quad (4.13)$$

Thus, the proof is done. □

Corollary 4.14. *$J = f$ is a stochastic contraction operation that transforms probability weights into assigned outcomes.*

Proof. The proof follows from isomorphism with averaging operator T . □

Corollary 4.15. *$J = f$ is a self adjoint map.*

5 Conclusion

This paper provides analytic proofs for prospect theory's topological space. In particular, it employs a monotone class theory to derive results from weak comonotonic probability measures embedded in outcome space. The almost sure isomorphism it produced is an approximation result. By characterizing prospect theory with operators, we provide a nexus between theory and experiment because the [discrete] spectrum provides the range of admissible values one should observe in an experiment. This theory was applied elsewhere to assess agents' response to fair gambles, and it was found that they assign asymmetric probabilities for gains and losses for martingales. Further research in this area includes applications to conjoint analysis.

6 APPENDIX

A Rank dependent utility and Schmeidler's Choquet representation

Given a von Neuman Morgenstern utility function $u : X \rightarrow \mathbb{R}$, and a lottery

$$(x_1, \pi_1; x_2, \pi_2; \dots x_n, \pi_n)$$

rank order a vector $x = (x_1, x_2, \dots, x_n)$ of outcomes so that it is $x_{rank} = (x_{(1)}, x_{(2)}, \dots, x_{(n)})$.

Assign known probability $\pi_{(i)}$ in correspondence with $x_{(i)}$ for $i = 1, 2, \dots, n$ so that they are "comonotonic". Define a transformation function $p : [0, 1] \rightarrow [0, 1]$ where $p(0) = 0$, $p(1) = 1$ such that

$$w_{(i)} = p\left(\sum_{k=1}^{k=i} \pi_{(k)}\right) - p\left(\sum_{k=1}^{k=i-1} \pi_{(k)}\right) \quad (\text{A.1})$$

is nonlinear. Thus, the "expected" value of the utility associated with these risky choices is

$$W(x) = \sum_i^n w_{(i)} u(x_{(i)}) \quad (\text{A.2})$$

In the context of TK92, for $f \in R^S$, and measurable sets

$$A_t = \{s : f(s) \geq t\} \in \mathfrak{G} \quad (\text{A.3})$$

where \mathfrak{G} is the σ -field of Borel subsets of S . Define a measure ν on monotonic subsets of the measurable space (S, \mathfrak{G}) , so the "expected value" of f with respect

to the measure is defined by the Choquet integral:

$$E_{\nu}[f] = \int f d\nu = \int_{X \cap [-\infty, 0]} (\nu(f \geq t) - 1) dt + \int_{X \cap (0, \infty]} \nu(f \geq t) dt \quad (\text{A.4})$$

Under this setup if $f : [0, 1] \rightarrow [0, 1]$, and ν is a probability measure, the integral is decomposed by the signed measure $\nu = \nu^+ - \nu^-$ on $[0, 1]$. However, the no additivity assumption puts limitations on the signed measure.

B Topological attachment and embedding of probability spaces

In what follows Ω is a probability space, S is a state space, and X is outcome space.

Definition B.1 (Free union). (Dugundji 1966, p. 127) The free union $\Omega \uplus S$ of disjoint spaces Ω and S is the union $\Omega \cup S$ with the following topology. $U \subset \Omega \uplus S$ is open if and only if $U \cap \Omega$ is open in Ω and $U \cap S$ is open in S . Since $\Omega \cap S = \emptyset$, Ω and S keep their own topologies and are disjoint open sets in $\Omega \uplus S$. $B \subset \Omega \uplus S$ is closed if and only if $B \cap \Omega$ is closed in Ω and $U \cap S$ is closed in S .

Definition B.2 (Attaching maps). Let Ω, S be two disjoint spaces, $V \subset \Omega$ a closed subset, and $P : \Omega \rightarrow S$ be continuous. Construct an equivalence relation

$$\mathcal{R} = (\omega, s) | \omega \text{ in } V, P(\omega) = s \subset \Omega \times S$$

in the free union $\Omega \uplus S$ wherein for $\omega \in V$ we have $\omega \sim P(\omega)$. The quotient space $\Omega \uplus S / \mathcal{R}$ between the free union and equivalence relation is said to be "Ω attached

to S by P ,” and is written $\Omega \cup_P S$ and P is called the attaching map. Specifically, we identify each $\omega \in V$ with its image $P(\omega) \in S$.

Definition B.3 (Embedding). If X is any outcome space, S a state space, and $f : S \rightarrow X$ is a mapping establishing $S \cong f(S) \subset X$, then f is called an embedding of S into X .

Theorem B.1 (Embedding). *Let $\pi : \Omega \uplus S \rightarrow \Omega \cup_P S$ be a projection. Then:*

1. *S is embedded as a closed set, homeomorphic to S , and $\pi|_S$ is a homeomorphism. That is, $\bar{S} \cong S \subset S$ and $\pi|_S : S \rightarrow X$ is a homeomorphism.*
2. *Let \bar{V} be closed in Ω . Then $\Omega - \bar{V}$ is embedded homeomorphically as an open set, and $\pi|_{\Omega - \bar{V}}$ is a homeomorphism.*

Proof. See (Dugundji, 1966, pg. 128). □

Theorem B.2 (Embedded subspace). *Let Ω be attached to S by a closed subset $\bar{V} \subset \Omega$, i.e., $P : \bar{V} \rightarrow S$. Let $\Omega_1 \subset \Omega$ and $S_1 \subset S$ be closed subsets such that $P(\bar{V}_1 \cap \Omega_1) \subset S_1$, and attach $\Omega_1 \rightarrow S_1$ by the restricted map $P_1 = P|_{\bar{V}_1 \cap \Omega_1}$. Then $\Omega_1 \cup_{P_1} S_1$ is homeomorphic to a closed subset of $\Omega \cup_P S$.*

Proof. See (Dugundji, 1966, pp. 128-129). □

Theorem B.3 (Extension of continuous composite map on attached spaces). *Let Ω be attached to S by $P : \bar{V} \rightarrow S$, where $\bar{V} \subset \Omega$ is closed, and let $\pi : \Omega \uplus S \rightarrow \Omega \cup_P S$ be the identification map. Let $\varphi : \Omega \rightarrow X$ and $\psi : S \rightarrow X$ be a pair of continuous maps, and let $(\varphi, \psi) : \Omega \uplus S \rightarrow X$ be their unique common extension. If $\varphi(\omega) = \psi[P(\omega)]$ for each $\omega \in \bar{V}$, then $(\varphi, \psi)\pi^{-1} : \Omega \cup_P S \rightarrow X$ is continuous.*

Proof. See (Dugundji, 1966, pg. 129). □

C Elements of operator theory used in this paper

C.1 Definitions and theorems

The material that follow is drawn from Curtain and Pritchard (1977), Rudin (1973), Reed and Simon (1980), and Saxe (2002). In this section we provide definitions and outline a few theories pertaining to linear operators.

Definition C.1 (Banach space, Hilbert space). . A Banach space is a complete normed linear space. A Hilbert space is a complete inner product space. A Hilbert space endowed with a norm induced by its inner product is an example of a Banach space.

Definition C.2 (Linear operator). . Let \mathfrak{M}_v and Y be linear spaces-real or complex. A mapping $T : \mathfrak{M}_v \rightarrow Y$ is called a linear operator if its domain D_T is a linear subspace of \mathfrak{M}_v , and for α, β real or complex

$$T(\alpha v + \beta g) = \alpha T v + \beta T g$$

for $v, g \in D_T$. When T is a numerical values mapping, it is called a linear functional.

Definition C.3 (Topological isomorphism). . Let $(\mathfrak{M}_v, \|\cdot\|_{\mathfrak{M}_v})$ and $(Y, \|\cdot\|_Y)$ be two normed linear spaces. \mathfrak{M}_v is said to be topologically isomorphic to Y , i.e.,

$\mathfrak{M}_v \cong Y$, if there exist a bijective map $T : \mathfrak{M}_v \rightarrow Y$ such that $\alpha \|v\|_{\mathfrak{M}_v} \leq \|Tv\|_Y \leq \beta \|v\|_{\mathfrak{M}_v}$ for all $v \in \mathfrak{M}_v$.

Definition C.4 (Compact operator). . Let U be an open ball in \mathfrak{M}_v . A linear operator $T : \mathfrak{M}_v \rightarrow Y$ is compact if every sequence in the closure $\overline{T(U)} \subset Y$ contains a convergent subsequence with limit point in $\overline{T(U)}$.

Definition C.5 (Denseness). (Curtain and Pritchard 1977, p. 20). A subspace S of a normed linear space X is dense in X if its closure with respect to the norm is equal to X or contains X as a subset. This important property means that any element $x \in X$ may be approximated by some element $sinS$, as closely as we like i.e. for any $x \in X$ and $\varepsilon > 0$, there exists $s \in S$ such that $\|x - s\| < \varepsilon$.

D Some Limit Theorems

Proposition D.1 (Prokhorov-Varadarajan). *Let X be a space of outcomes, complete in norm $\|\cdot\|_X$, and $\mathfrak{B}(X)$ be the σ -field of Borel subsets of X . Let $\{\mu_n\}_{n \geq 1}$ be a sequence of probability measures on $(X, \mathfrak{B}(X))$. Then $\{\mu_n\}$ is tight if and only if given a subsequence $\{\mu_{n_k}\} \subset \{\mu_n\}$, there exist a nested subsequence $\{\mu_{n_k^{(j)}}\} \subset \{\mu_{n_k}\}$ and a probability measure μ on $(X, \mathfrak{B}(X))$ such that $\mu_{n_k^{(j)}} \xrightarrow{d} \mu$ as $j \rightarrow \infty$.*

Proof. See (Athreya and Lahiri, 2006, pg. 303). □

Definition D.1 (ε -net). (Billingsley, 1968, pg. 217). Let E be a compact set in X . An ε -net for E is a set of outcomes $\{x_k\}$, not necessarily in E , with the property

that for each outcome x in E there is an outcome x_k such that $\rho(x, x_k) < \varepsilon$. E is totally bounded if for every $\varepsilon > 0$ it has a finite ε -net.

E Proofs of theorems, lemmas, and propositions

Proof of Theorem 3.3

Proof. Let $L_w^2(\mathfrak{M}_F)$ be the space of square integrable functions with respect to TK92 probability weight functions. Define an inner product of this space by $(f, g) = \int_X f \bar{g} w(dF)$. Thus,

$$(T_g v_g, T_g v_g) = \|T_g v_g\|_{L_w^2}^2 = (v_g, v_g) = \|v_g\|_{L_w^2}^2 \text{ which implies } \|T_g\|^2 = 1. \quad (\text{E.1})$$

Thus, T_g is isometric. Therefore, it preserves measures of gains. By contrast,

$$(T_\ell v_\ell, T_\ell v_\ell) = \|T_\ell v_\ell\|_{L_w^2}^2 \quad (\text{E.2})$$

$$= (-\lambda v_\ell, -\lambda v_\ell) = |\lambda|^2 \|v_\ell\|_{L_w^2}^2 \quad (\text{E.3})$$

$$\Rightarrow \|T_\ell\|^2 = |\lambda|^2 \quad (\text{E.4})$$

$$\Rightarrow \sup_{v_\ell=1} \|T_\ell\| = \max_{\lambda \in \Lambda_\ell} \lambda. \quad (\text{E.5})$$

Which shows that the loss operator T_ℓ is skewed and it bounded by the maximal eigenvalue in the loss spectrum Λ_{T_ℓ} . □

Proof of Theorem 3.1.

The following proof is adapted from (Yosida, 1980, pp. 118-119)

Proof. Let $\mathcal{P} = \{A_1, A_2, \dots, A_n\}$ be a partition of S , and for $\varepsilon > 0$ let $\delta(A_i) = \text{diam}(A_i) < \varepsilon$. Define the set $B_i = \{s \in S \mid f(s) \in A_i\}$. So that for some state $\alpha_i \in A_i$ we have

$$\left\| f - \sum_{i=1}^n \alpha_i \chi_{B_i}(s, \omega) \right\|_{L^\infty(S \times \Omega, \mathfrak{G} \times \mathfrak{D}, \mu)} \leq \varepsilon \quad (\text{E.6})$$

$$\left\| V(f) - \sum_{i=1}^n \alpha_i \psi(B_i, \omega) \right\|_{L^\infty(S \times \Omega, \mathfrak{G} \times \mathfrak{D}, \mu)^*} \leq \|f\|_{L^\infty(S \times \Omega, \mathfrak{G} \times \mathfrak{D}, \mu)} |\varepsilon| \quad (\text{E.7})$$

where the last equation is due to linearity of the functional $V(f - \sum_{i=1}^n \alpha_i \chi_{B_i}(s, \omega))$. Since ε is arbitrary, let $\varepsilon = 2^{-n}$. So that as $n \rightarrow \infty$, $\varepsilon \downarrow 0$. By hypothesis $\psi(B_i, \omega)$ is a random set function. So that by Radon-Nikodym theorem it is finitely additive, and there exist a measure, call it dV , that is absolutely continuous with respect to $\psi(ds, \omega)$. That is, $\frac{dV}{d\psi} = f$. In particular, in the limit we have

$$V(f) = \int_S f(s, \omega) \psi(ds, \omega) \quad (\text{E.8})$$

$$\|V\|_{L^\infty(S \times \Omega, \mathfrak{G} \times \mathfrak{D}, \mu)^*} = \sup_{\text{ess sup } |f(s, \omega)| \leq 1} \left| \int_S f(s) \psi(ds, \omega) \right| \quad (\text{E.9})$$

Thus, the proof is done. □

Proof of Theorem 4.2

Proof. (i) Suppose that $v \in \mathfrak{B}(X)$. Then by definition $f \in S \times \Omega$ and $\phi(s, \omega) = (v \circ f)(s, \omega)$. Let

$$E_\alpha = \{(s, \omega) \mid \phi = (v \circ f)(s, \omega) > \alpha, \alpha \in \mathbb{R}\} \in \mathfrak{G} \times \mathfrak{D}$$

Then $\phi = \lim_{\alpha \rightarrow \infty} \bigcap_{\alpha} E_{\alpha}$ is measurable with respect to $\mathfrak{S} \times \mathfrak{D}$

(ii) Suppose that ϕ is a measurable function on $S \times \Omega$. Let $\mathfrak{M}_{v \circ f}$ be the space of all real valued functions of the form $v \circ f$ defined on $S \times \Omega$ with v $\mathfrak{B}(X)$ -measurable. Clearly $1 \in \mathfrak{M}_{v \circ f}$ since for $B \in \mathfrak{S} \times \mathfrak{D}$ we have the stochastic choice function $\chi_B(s, \omega)$ defined on $S \times \Omega$ and $\chi_B(s, \omega) = 1, s \in B$ and 0 otherwise. Let $\{v_n \circ f\}$ be an increasing sequence in $\mathfrak{M}_{v \circ f}$ such that $g = \sup_n v_n \circ f < \infty$. Let $A = \{(s, \omega) \in S \times \Omega \mid \sup_n v_n < \infty\}$. Then $A \in \mathfrak{S} \times \mathfrak{D}$ and $f(S \times \Omega) \subset A$. To see this latter relation, suppose that $f(S \times \Omega) \not\subset A$. Then there exist f such that $\sup_n v_n \geq \infty$. This result is contrary to our hypothesis of bounded value functions. Therefore, $f(S \times \Omega) \subset A$. Define $v = \sup_n v_n$ on A and $v = 0$ on A^c . Then v is $\mathfrak{S} \times \mathfrak{D}$ -measurable, and consequently $g = v \circ f$. If $C \in \mathfrak{S} \times \mathfrak{D}$, then for some set $A \in \mathfrak{S} \times \mathfrak{D}$, we have $C = f^{-1}(A)$ so that $\chi_C(s, \omega) = (\chi_A \circ f)(s, \omega) \in \mathfrak{M}_{v \circ f}$. Thus, by Lemma 4.1 $\mathfrak{M}_{v \circ f}$ contains all real valued functions on $\mathfrak{S} \times \mathfrak{D}$. Since v is a real valued value function, it is numerical, i.e., a functional. Thus, we can write $y = \tan^{-1}(v) \in \mathfrak{S} \times \mathfrak{D}$. Thus, $y = \tilde{v} \circ f$ for some $\mathfrak{B}(X)$ -measurable function \tilde{v} . Assume that $\tilde{v}(X) \subset [-\frac{\pi}{2}, \frac{\pi}{2}]$ because by definition y takes values only in this interval. If we let $\phi = v \circ f$ with $v = \tan(\tilde{v})$, then the value function v is also measurable. So, there exist ϕ such that the value function v is $\mathfrak{B}(X)$ -measurable. Thus, the proof is done. \square

Remark E.1. Even though the choice of $v = \tan(y)$ is arbitrary in this proof, it nonetheless captures the salient characteristics of Tversky and Khaneman (1992) value functions—concave in gains and convex in losses, relative to the "reference point" $y=0$. That representation can also be derived from theoretical considera-

tions outside the scope of this paper.

References

- Acerbi, C. (2001). Risk aversion and coherent risk measures: A spectral representation theorem. <http://arxiv.org/abs/cond-mat/0107190>. Abaxbank Working Paper, Milano, Italy.
- Akheizer, N. and I. Glazman (1961). *Theory of Linear Operators in Hilbert Space*. Frederick Ungar Publ. Co.:New York. Dover reprint 1993.
- al Nowaihi, A. and S. Dhimi (2005). A simple derivation of Prelec's probability weighting functions. <http://ssrn.com/abstract=1130811>. Working Paper No. 05-20, Dept. Economics, Univ. Leicester, UK.
- Ash, R. B. (2000). *Probability and Measure Theory*. Harcourt-Academic Press. With contributions from Catherine Doléans-Dade.
- Athreya, K. B. and S. N. Lahiri (2006). *Measure Theory and Probability Theory*. New York, N.Y.: Springer-Verlag.
- Billingsley, P. (1968). *Convergence in Probability Measures*. New York, N.Y.: John Wiley & Sons, Inc.
- Billingsley, P. (1995). *Probability and Measure* (3rd ed.). New York: John Wiley & Sons, Inc.
- Birchby, J., G. Gigliotti, and B. Sopher (2008, July). Consistency and aggregation in individual choice under uncertainty. pp. 1–21. Foundations and Applications of Utility, Risk and Decision Theory–Barcelona. Dep't. Economics, Rutgers Univ. New Brunswick, New Jersey.
- Blichrodt, H. and J. L. Pinto (2000). A parameter free elicitation of the probability weighting function in medical decisions. *Management Science* 46(11), 1485–1496.
- Blumenthal, R. M. and R. K. Getto (1968). *Markov Processes and Potential*. New York: Academic Press.
- Bochner, S. (1955). *Harmonic Analysis and the Theory of Probability*. Berkeley, CA: Univ. California Press.
- Brieman, L. (1968). *Probability and Measure*. Reading, MA. Addison-Wesley. SIAM reprint 1992.

- Burns, Z., A. Chiu, and G. Wu (Forthcoming, Fall 2010). *Encyclopedia of Operations Research and Management*, Chapter *Overweighting of Small Probabilities*. John Wiley & Sons, Inc.
- Cadogan, G. (2010). On spectral analysis of risk operators with applications to prospect theory. Work-in-Progress.
- Cížek, P. (2007). Robust and efficient adaptive estimation of binary choice regression models. <http://arno.uvt.nl/show.cgi?fid=57456>. CentER Discussion Paper No. 2007-12, Tilburg University, Netherlands.
- Curtain, R. F. and A. J. Pritchard (1977). *Functional Analysis in Modern Applied Mathematics*, Volume 132 of *Mathematics in Science and Engineering*. Academic Press, New York.
- Dagsvik, J. (2006, July). Axiomatization of stochastic choice models under uncertainty. Discussion Paper No. 465, Statistics Norway, Research Department.
- Davidson, D. and J. Marschak (1958, July). Experimental tests of stochastic decision theory. Technical Report No. 17, Behavioral Sciences Division, Applied Math and Statistical Laboratory, Stanford Univ.
- de Palma, A., M. Ben-Akiva, D. Brownstone, C. Holt, T. Magnac, D. McFadden, P. Moffatt, N. Picard, K. Train, P. Wakker, and J. Walker (2008). Risk, Uncertainty and Discrete Choice. *Marketing Letters* 19, 269–285.
- Debreu, G. (1953, May). Representation of a preference ordering by a numerical function. *mimeo*, Cowles Comm. Discussion Paper No. 2083, Yale Univ.
- Debreu, G. (1958). Stochastic choice and cardinal utility. *Econometrica* 26, 440–444.
- Dugundji, J. (1966). *Topology*. Boston, MA: Allyn and Bacon, Inc.
- Fabiyi, M. E. (2008). *Advances in Decision Making Under Risk and Uncertainty*, Chapter II, pp. 109–118. Number 42 in Series C: Mathematical Programming and Operations Research. Berlin: Springer-Verlag.
- Gonzalez, R. and G. Wu (1999). On the shape of the probability weighting function. *Cognitive Psychology* 38, 129–166.

- Gonzalez, R. and G. Wu (2003, August). Composition rules in original and cumulative prospect theory. <http://faculty.chicagobooth.edu/george.wu/research/papers/CompositionRules.pdf>. Working Paper, Dep't. Psychology, Univ. Michigan.
- Härdle, W. and L. Simar (2003). *Applied Multivariate Statistical Analysis*. Springer-Verlag. Galley proof copy.
- Hewitt, E. and K. Stromberg (1965). *Real and Abstract Analysis*. Number 25 in Graduate Texts in Mathematics. Springer-Verlag.
- Hoffman-Jorgenson, J. (1971). Existence of conditional probability. *Math. Scand.* 28, 257–264.
- Horton, G. A. (2004). Defining risk. *Financial Analyst Journal* 60(6), 19–25.
- Hsu, M., I. Krajbich, and C. F. Camerer (2009, February). Neural response to reward anticipation under risk is nonlinear in probabilities. *Journal of Neuroscience* 29(7), 2231–2237.
- Hunt, E. (2007). *The Mathematics of Behavior*. New York: Cambridge University Press.
- Ingersoll, J. E. (2008, June). Non-monotonicity of the Khaneman-Tversky probability weighting function: A cautionary note. *European Financial Management* 14(3), 385–390.
- Jacobson, N. (1951). *Lectures in Abstract Algebra*. Springer-verlag, New York. Reprint 1975.
- Loewenstein, G. and D. Prelec (1991). Negative time preference. In *AEA Papers and Proceedings*, Volume 81, pp. 347–352. Amer. Econ. Rev.
- Luce, D. and L. Narens (2008). Theory of measurement. In L. Blume and S. N. Durlauf (Eds.), *Palgrave Dictionary of Economics* (2nd ed.). Palgrave Macmillan. Preprint.
- Luce, R. D. (1959). *Individual Choice Behavior*. New York,: John Wiley & Sons, Inc.
- Luce, R. D. (2001). Reduction invariance and Prelec's weighting functions. *Journal of Mathematical Psychology* 45, 167–179.

- McFadden, D. P. (1974). *Frontiers in Econometrics*, Chapter IV. *Conditional Logit Analysis of Qualitative Choice Behavior*, pp. 105–142. New York: Academic Press.
- McFadden, D. P. (1980, July). Econometric models for probabilistic choice among products. *Journal of Business* 53(3), S13–S29.
- McShane, E. (1963). Integrals devised for special purposes. *Bull. Amer. Math. Soc.* 69, 597–627.
- Moy, S.-T. C. (1954). Characterizations of conditional expectation as a transformation on function spaces. *Pacific Journal of Mathematics* 4, 47–63.
- Munkres, J. (2000). *General Topology*. New Jersey. Prentice-Hall, Inc.
- Musiał, K., W. Strauss, and N. D. Macheras (2009). Liftings for topological products of measures. <http://www.math.uni.wroc.pl/musial/mms22.pdf>. *mimeo*, Wrocław Univ., Poland.
- Musiał, K. Strauss, W. and N. D. Macheras (2002). *Handbook of Measure Theory*, Chapter 28. *Liftings*, pp. 1131–1184. Elsevier North-Holland.
- Narens, L. (2007). *Theories of Probability: An Examination of Logical and Qualitative Foundations*, Volume 2 of *Advanced Series on Mathematical Psychology*. Singapore: World Scientific.
- Prelec, D. (1998). The probability weighting function. *Econometrica* 60, 497–528.
- Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organization* 3, 323–343.
- Reed, M. and B. Simon (1980). *Modern Methods of Mathematical Physics* (Revised and Enlarged Edition ed.), Volume I: Functional Analysis. San Diego, CA: Academic Press, Inc.
- Richards, I. (1959). A note on the Daniell integral. *Rendicotti del Seminario Matematico della Univ. di Padova* 29, 401–410.
- Rota, G.-C. (1960). On the representation of averaging operators. *Rendicotti del Seminario Matematico della Univ. di Padova* 30, 52–64.

- Royden, H. L. (1988). *Real Analysis* (3rd ed.). New York: Macmillan Publishing Co.
- Rudin, W. (1973). *Functional Analysis*. New York, N.Y.: McGraw-Hill, Inc.
- Saxe, K. (2002). *Beginning Functional Analysis*. Springer-Verlag, New York, Inc.
- Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica* 57(3), 571–587.
- Shilov, G. (1973). *Elementary Real and Complex Analysis*. MA:MIT Press. Dover Publications, Inc. reprint 1996.
- Stott, H. P. (2003). Cumulative prospect theory's functional menagerie. mimeo, Dept. Psychology, Univ. Warwick, UK, *Journal of Risk and Uncertainty*, 32, 101-130 (2006).
- Takahashi, T. (2006). A mathematical framework for probabilistic choice based on information theory and psychophysics. *Med. Hypotheses* 67(1), 183–186.
- Thomas, E. G. F. and A. Volčič (1989). Daniell integral represented by Radon measures. *Rendicotti del Circolo Matematico di Palermo* 38, 39–59.
- Train, K. (2003). *Discrete Choice Methods With Simulations*. New York: Cambridge University Press.
- Tversky, A. and D. Kahneman (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5, 297–323.
- Wu, G. and R. Gonzalez (1996). Curvature of the probability weighting function. *Management Science* 42, 1676–90.
- Wu, G. and R. Gonzalez (1999). Nonlinear decision weights in choice under uncertainty. *Management Science* 45, 74–85.
- Wu, G., J. Zhang, and R. Gonzalez (2004). *Handbook of Judgment and Decision Making*, Chapter :Decision Under Risk, pp. 399–423. Oxford: Blackwell.
- Yosida, K. (1980). *Functional Analysis* (6th ed.). New York: Springer-verlag.