Further Comments and Elucidations for

Wakker (2010) "Prospect Theory: for Risk and Ambiguity"

July, 2016

P. xiii [Adding acknowledgment in Preface]: Salvatore Modica used my manuscript since its first version in 1996, suggested improvements, and encouraged me throughout to turn the manuscript into a book.

P. 15: teachers working on decision under uncertainty all their life may overlook that the decision under uncertainty model takes time to grasp. Exercise 1.1.1 is useful for this purpose.

P. 26, 3rd para [Selling x^j in financial market]: we assume free trade in financial markets, where you can sell any x you like. Selling x means buying –x. That is, you can sell x without possessing x in some sense. □

P. 28, Exercise 1.6.1 [maxmin]: maxmin means that you go by the worst-case scenario. □

P. 31 ℓ.5. Eq. 1.6.3 can indeed be considered to be a weakened version of additivity as described there, but this takes some thinking. The weakened version is that we only impose the implication $[x \sim y \Rightarrow x+z \sim y+z]$ if x and y are either constant or a prospect that has all but one of its outcomes equal to 0 (this is the case in Eq. 1.6.2

and its special case Eq. 1.6.1). We need the implication [$p \sim 1_{E_1}0 \Rightarrow p\alpha \sim \alpha_{E_1}0$]. This implication can be derived from the weakened version of additivity by repeating the proof of Lemma 1.11.2 on p. 42, and then never needing more than the mentioned weakened version. It involves the other conditions such as monotonicity and existence of CEs. □

P. 50, ℓ. −6 [Subjective probability ≠ objective probability]: the discussion of $p = 0.54$ following Exercise 4.10.6 also illustrates such a discrepancy. □

P. 63, Comment 2.6.5 [Variance of utility]. The mistake criticized there, of not realizing that the utility unit already comprises risk attitude, and that speculating on risk attitudes wrt utils is double counting, was also made by

Baumol, William J. (1951) “The von Neumann-Morgenstern Utility Index—An Ordinalist View,” *Journal of Political Economy* 59, 61–66.

P. 64 there argues that, with utils as unit of payment, $600_{1/6}420 \succ 600_{5/6}60$ is a reasonable preference because of the security of 420, but that it violates EU because the EUs are 450 and 510, respectively.

P. 65 Exercise 2.7.2. That independence and the sure-thing principle for risk are equivalent under usual continuity assumptions was claimed informally, without elaboration given, in Observation 6.2 of

Fishburn, Peter C. & Peter P. Wakker (1995) “The Invention of the Independence Condition for Preferences,” *Management Science* 41, 1130–1144.

P. 69 ff. (§3.1). Many analyses of this kind are in the journal Medical Decision Making; see for instance

van den Akker–van Marle, M. Elske, Mascha Kamphuis, Helma B. M. van Gameren–Oosterom, Frank H. Pierik, Job Kievit, & NST Expert Group (2013)

"Management of Undescended Testis: A Decision Analysis," *Medical Decision Making* 33, 906–919.

P. 77 following Theorem 3.4.1 on the Pratt-Arrow measure:

On p. 700/701, the following paper introduced, before Pratt/Arrow, the Pratt/Arrow measure –u''/u' and its elementary properties such as:

- it being a measure of concavity;
- the 50/50 gamble for gaining or losing h being equivalent to losing h^2 divided by the measure (P.s.: that's the special case of risk premium when expected value is zero);
- the measure also being related to an excess probability for gaining
- it entirely comprising all of u that's relevant.

de Finetti, Bruno (1952) "Sulla Preferibilitá," *Giornale degli Economistii e Annali di Economia* 11, 685–709.

P. 92, bottom:

PROOF OF THEOREM 3.4.1 USING FORMULAS.

For every prospect $f = p_1x_1 \cdots p_nx_n$ we define the associated prospect $\bar{f}$ as $p_1U_1(x_1) \cdots p_nU_1(x_n)$. Because U_1 strictly increases, for every associated prospect there is a unique prospect with which it is associated. Define $\bar{\succcurlyeq}$ by $\bar{f}\ \bar{\succcurlyeq}\ \bar{g}$ if and only if $f \succcurlyeq_2 g$, where $\bar{f}$ is associated with f and $\bar{g}$ with g. Note that $\alpha \sim_1 f$ means that δ $(= U_1(\alpha))$ is the expected value of $\bar{f}$, i.e., $\sum_{j=1}^{n} p_jU_1(x_j)$). $\succcurlyeq_2$ being more risk averse than $\succcurlyeq_1$ means exactly that $\bar{\succcurlyeq}$ is risk averse. $\bar{\succcurlyeq}$ maximizes EU with respect to the utility function $\varphi(\cdot) = U_2(U_1^{-1}(\cdot))$. By Theorem 3.2.1a, risk aversion of $\bar{\succcurlyeq}$ is equivalent to concavity of φ. Obviously, $U_2(\cdot) = \varphi(U_1(\cdot))$. □

P. 92, bottom:

P. 103, 3 lines above Assignment 4.2.8 [Continuity versus differentiability]. This book uses continuity assumptions rather than differentiability assumptions for the following reason. Binary choices typically generate discrete data. These will be combined with richness assumptions implying that the domain of all conceivable choice prospects is a continuum. Continuity can, unlike differentiability (to the best of my knowledge), easily be defined in terms of preferences, which is why we often use it, and why techniques from differential calculus will not be used frequently in this book.

P. 103, Exercise 4.3.1.
This exercise can be simplified by assuming $U(\alpha^0) = 0$ and $U(\alpha^1) = ¼$, and then letting the students determine $U(\alpha^2)$ from the indifferences in Figures 4.1.1a and b. $U(\alpha^3)$ and $U(\alpha^4)$ are then obtained similarly. Whereas this approach is somewhat simpler, it gives fewer insights. Typically, for b- or c-students (economics) I will do the exercise as in the book, and for a-students (business or psychology) I will use the variation suggested here. Giving the elaboration also for $U(\alpha^3)$, and showing it is the same as for $U(\alpha^2)$ but only the superscripts of α increased by one, gives further insights.

P. 105 ff. When elaborating, it is more convenient for the students to write the notation α^1, α^2, and so on, rather than to write the particular numbers that they chose. Using the recommended notation, they can readily compare their elaborations with each other and with the elaborations of the teacher/book.

P. 112, §4.7, beginning: the extra Assignment 4.7.2, added on internet, illustrates how inconsistencies of probability lead to inconsistencies in utility measurement.

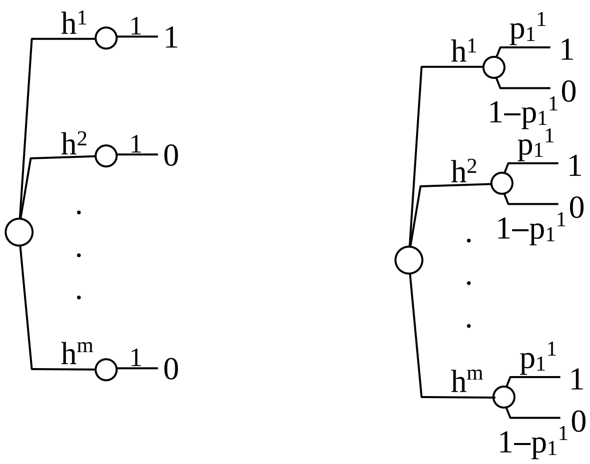
Pp. 125 ℓ 7 [linearity of payment in probability for nonexpected utility]. Selten, Sadrieh, & Abbink (1999) showed, more generally, that payment in probability can be used to generate linear utility for every model that has probabilities in the first stage (either objective or subjective, the latter under probabilistic sophistication), reduction of compound lotteries, weak ordering, and stochastic dominance.

P. 167 [Exercise 5.6.2]. Exercise 5.6.2 does NOT serve to illustrate the preceding text. Students sometimes get confused thinking it does.

P. 184 [TO method robust to intransitivities]. Section 6.5.1 has demonstrated that the tradeoff method is robust to violations of expected utility due to rank-dependent probability weighting. The following paper shows that the tradeoff method is also robust to violations of transitivity. That is, it also measures utility under regret theory.

Bleichrodt, Han, Alessandra Cillo, & Enrico Diecidue (2010) "A Quantitative Measurement of Regret Theory," *Management Science* 56, 161–175.

P. 125, 2nd para of §4.9.4 [Implementing $1_h0 \sim 1_p0$ in a multistage setup]. Let us consider the indifference $1_h0 \sim 1_p0$ in the multistage setup in Fig. 4.9.4b. Here 1_h0 can be obtained for instance by taking $n = 1$, $x_1^1 = 1$, and $x_1^j = 0$ for all $j \geq 2$, as in the left figure below. 1_p0 can be obtained for instance by taking $n = 2$, $p_1 = p^1$, $p_2 = 1-p^1$, and $x_1^j = 1$ and $x_2^j = 0$ for all j, as in the right figure below.



□

Pp. 178 [Postponing Exercise 6.4.2]. When teaching I usually do Exercise 6.4.2 only after Observation 6.5.1 and Exercise 6.5.1. □

Pp. 182 [Postponing Exercise 6.5.2]. When teaching I usually do Exercise 6.5.2 only after Exercise 6.5.6 and §6.5.4. □

Pp. 264 [Fitting many parameters to limited data]. A difficulty with the data fitting in this book, in particular for the exercises in Section 9.5, is that we fit many parameters to limited data sets. Limited data sets is done only for simplifying the calculation work. The parameters may interact in unclear ways. It may happen that utility curvature for gains and losses is not influenced only by risk attitudes for gains and losses, respectively, but that it interacts with the fitting of loss aversion, for instance.

P. 268 [Interpretation of P under probabilistic sophistication].

P is not just a mathematical devise under the normative Bayesian assumption that for every person, in whatever state of information, there exists an underlying probability measure that best reflects the state of knowledge of this person. Then P (being the only probability measure that respects the required symmetry of n-fold uniform partitions of the state-space for all n,) is the most plausible approximation of that underlying probability measure and, thus, of the rational beliefs of the person given the information possessed. Descriptively, however, the cognitive beliefs actually held by the person then need not coincide with P and could, for instance, be a nonlinear transformation of P.

P. 319 [Footnote 3: neo-additive versus probabilistic sophistication]. Definition 11.2.1 gives the most common and simplest version of neo-additive weighting functions. Other versions are possible, and the definition in Chateauneuf, Eichberger, & Grant (2007) includes such other versions. One such version concerns probabilistic sophistication, where $W(\cdot) = w \circ P$ for a probability measure P and a neo-additive transformation w (the latter in the sense of Eq. 7.2.5); see Chateauneuf et al. Remark 3.1. It differs from the version in Definition 11.2.1 because nonempty *P-null* events (meaning their P value is 0) have weight 0 for all possible ranks in the probabilistically sophisticated case, but not by Definition 11.2.1. Under Definition 11.2.1, they have nonzero weight with best rank if $b > 0$ and nonzero weight with worst rank if $a+b < 1$.

The two cases of neo-additive weighting functions are two extremes. Under Definition 11.2.1, all nonempty P-null events are in the set that Chateauneuf et al. denote $\mathcal{N}$, call null events, and that I will call preferentially null events here[1]. Given that Chateauneuf et al. require that all preferentially null events are P-null (their congruency requirement for P), their set $\mathcal{N}$ then is maximal given P. Definition 11.2.1 has $\mathcal{N}$ minimal, containing only $\varnothing$. Chateauneuf et al. consider intermediate cases, where $\mathcal{N}$ contains some nonempty events but not all P-null events. □

P. 321 [Definition of ambiguity aversion]. The book gives no definition of ambiguity aversion. I hope that readers can guess what it must be: *ambiguity aversion* holds for a source $\mathcal{B}$ if there is source preference of the source of known probabilities over $\mathcal{B}$.

P. 322 [Forward contract for wheat]. *The Economist* of 11Sep2010, p.53, penultimate para in 3rd column, refers to forward contracts for farmers in wheat.

P. 324 [CORE].

The probability measure $(p_1,\ldots,p_n)$ indicates that player j receives a p_j share of the total earnings. No incentive to deviate means, for instance for the coalition $\{s_2,s_3\}$, that the $p_2 + p_3$ share of this coalition exceeds the amount they could make on their own, i.e.

$$(p_2 + p_3)W(S) \geq W\{s_2,s_3\}.$$

---

[1] It means that they are null with best rank, which in Chateauneuf et al.'s setup implies that they are also null with worst rank, and even null with all ranks.

P. 324 [Convex W].

In the proof of Theorem 11.4.1, here is a direct proof that convexity of W implies a nonempty CORE. The other direction is more difficult.

Assume that W is convex, and take any complete ranking of S. Let P be the probability measure assigning the decision weight of each state under that rankinh to that state. P is an element of the CORE. To see this point, consider an arbitrary event E; say, $\{s_{i_1},\ldots, s_{i_j}\}$. We may assume that, according to the complete ranking ρ from which we derived P, s_{i_1} has the best rank within this j-tuple, s_{i_2} the second-best, …, and s_{i_j} the worst. Now consider another complete ranking ρ´ that ranks i_1 as best of all states of nature, i_2 as 2nd best of all, and i_j as jth best of all. The total decision weight of these j states under ρ is P(E), and under ρ´ it is W(E). The latter weight has to be bigger because all ranks there are better and better ranks generate lower decision weights for a convex, pessimistic, W. Hence, P is a CORE element.

P. 326 [Crudeness of multiple priors]. The middle para on the page argues that multiple priors are too crude because they focus on best and worst EU priors, ignoring the ones in between. The following experimental test took multiple priors up on this point and found it violated:

Hayashi, Takashi & Ryoko Wada (2010) “Choice with Imprecise Information: An Experimental Approach,” *Theory and Decision* 69, 355–373.

P. 337 [Nondifferentiability of RDU in §6.4.2].

In the text:

Klibanoff, Marinacci, & Mukerji (2005) introduced a smooth model for ambiguity, avoiding the nondifferentiability of RDU (§6.4.2) and the other models discussed in

the reference to §6.4.2 concerns the para on pp. 179-180. (This para does not explicitly use the term (non)differentiability but describes it in words.)

P. 340 [Complex Mobius transform].

The following example could be added following Example 11.11.3. I thought it is too complex for what it gives and, hence, did not incorporate it in the book. But maybe some readers find it interesting.

EXAMPLE. Consider an urn that contains 100 balls of which m are blue (B), $50 - m$ are green (G), m are red (R), and $50 - m$ are yellow (Y). Here $0 \leq m \leq 50$ is unknown. We assume usual ambiguity aversion with $W(E) = 2/12$ for all singleton events, a special aversion through W(B,R) = W(G,Y) = 3/12, and further W(B,G) = W(R,Y) = W(B,Y) = W(G,R) = 6/12, and W(E) = 9/12 for all three-color events. Then $\varphi(E) =$ 2/12 for all singleton events, φ(B,R) = φ(G,Y) = –1/12, φ(B,G) = φ(R,Y) = φ(B,Y) = φ(G,R) = 2/12, $\varphi(E) = 0$ for all 3-color events, and φ(B,G,R,Y) = –2/12. There are some negative interactions, such as between B and R. The combination of these colors is, apparently, perceived as extra ambiguity in addition to the ambiguity of each event in isolation. W is 4-additive. □

P. 348 [Penultimate para].

The underlined sentence in the text:

> [$\alpha_E G x \succcurlyeq \alpha_E G y \Leftrightarrow \gamma_E L x \succcurlyeq \gamma_E L y$] for all gains $\alpha > 0$ and losses $\beta < 0$ whenever E has the same gain-rank in all four prospects.

E having the same gain-rank means that L = $(G \cup E)^c$. This follows immediately from equivalence with the preferences $0_E G x \succcurlyeq 0_E G y$ and $0_E L x \succcurlyeq 0_E L y$, where $0_E G x = 0_E L x$ and $0_E L y = 0_E G y$. Hence, we do not define a separate rank-sign sure-thing principle.

refers to:

$\alpha_E G x \succcurlyeq \alpha_E G y$ *is equivalent to* $0_E G x \succcurlyeq 0_E G y$

and

$\gamma_E L x \succcurlyeq \gamma_E L y$ *is equivalent to* $0_E L x \succcurlyeq 0_E L y$

P. 407 [1st para].

If utility is increased by τ, then so are the expected utilities of prospects: τ is added to the utility of each outcome, then these τ's are multiplied by the probabilities p_j, but those probabilities add to 1, so in total the EU of the prospect is increased by $\sum_{j=1}^{n} p_j \tau = \tau$.

P. 409 [3rd para].

Here is such a proof by induction:

In this derivation we replace all-but-one of the outcomes by one outcome. We use induction, and assume the result demonstrated for all prospects with n or fewer outcomes. Consider $p_1x_1 \ldots p_{n+1}x_{n+1}$, and its EU $\sum_{j=1}^{n+1} p_j U(x_j)$. It is $\sum_{j=1}^{n} p_j U(x_j) + p_{n+1}U(x_{n+1}) =$

$\sum_{j=1}^{n} p_j \times \frac{\sum_{j=1}^{n} p_j U(x_j)}{\sum_{j=1}^{n} p_j} + p_{n+1}U(x_{n+1}) \leq$ (by the induction hypothesis)

$\sum_{j=1}^{n} p_j \times U\left(\frac{\sum_{j=1}^{n} p_j x_j}{\sum_{j=1}^{n} p_j}\right) + p_{n+1}U(x_{n+1}) \leq$ (by the induction hypothesis for two outcomes, with $\sum_{j=1}^{n} p_j = 1 - p_{n+1}$)

$U\left(\sum_{j=1}^{n} p_j \times \left(\frac{\sum_{j=1}^{n} p_j x_j}{\sum_{j=1}^{n} p_j}\right) + p_{n+1}x_{n+1}\right) = U\left(\sum_{j=1}^{n+1} p_j x_j\right)$.

P. 440 Exercise 7.10.1: the values of a and b have been rounded. The exact values are $a = 0.725$ and $b = 0.0975$. The optimism and likelihood sensitivity indexes are exact.