J Risk Uncertain (2014) 48:1–17
DOI 10.1007/s11166-014-9185-0

An experimental test of prospect theory for predicting choice under ambiguity

Amit Kothiyal · Vitalie Spinu · Peter P. Wakker

Published online: 1 March 2014

Abstract Prospect theory is the most popular theory for predicting decisions under risk. This paper investigates its predictive power for decisions under ambiguity, using its specification through the source method. We find that it outperforms its most popular alternatives, including subjective expected utility, Choquet expected utility, and three multiple priors theories: maxmin expected utility, maxmax expected utility, and α-maxmin expected utility.

Prospect theory was originally introduced for decisions under risk (Kahneman and Tversky 1979). There were two problems with the original 1979 version (*original prospect theory* or *OPT* henceforth). First, it could not handle more than two nonzero outcomes.[1] Second, it generated violations of stochastic dominance that were not descriptively realistic. These problems were resolved in 1992, when Tversky and

[1] All extensions to more outcomes proposed in the literature have problems (Wakker 2010, Section 9.8).

Electronic supplementary material The online version of this article (doi:10.1007/s11166-014-9185-0) contains supplementary material, which is available to authorized users.

A. Kothiyal
Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany

V. Spinu
Anderson School of Management, UCLA, 110 Westwood Plaza, Los Angeles, CA 90095, USA

P. P. Wakker (✉)
Erasmus School of Economics, Erasmus University, P.O. Box 1738, Rotterdam, 3000 DR, the Netherlands
e-mail: wakker@ese.eur.nl

Kahneman (1992) used Quiggin's (1982) rank dependence in their new version of prospect theory (called PT),[2] as is well known.

Another more important advance of PT has, as yet, received less attention. Tversky and Kahneman (1992) did not only use Quiggin's (1982) rank dependence to solve the two aforementioned problems. They also used Schmeidler's (1989) rank dependence to extend PT to ambiguity. Ambiguity refers to situations of uncertainty with no probabilities given (risk refers to situations with probabilities given). Unfortunately, there is still a widespread misunderstanding that prospect theory could only be applied to risk. The popularity of prospect theory for risk may have contributed to this misunderstanding. Another reason for this misunderstanding is that prospect theory is purely descriptive, with no normative claims made, whereas most researchers on ambiguity have so far focused on developing models for normative purposes, without seeking to describe the phenomena under ambiguity found empirically.

Hey et al. (2010), HLM henceforth, presented a predictive test to compare various theories on decision under ambiguity. These theories include subjective expected utility (Savage 1954), Choquet expected utility (CEU; Gilboa 1987; Schmeidler 1989), maxmin expected utility (Chateauneuf 1991; Cozman 2012; Gilboa and Schmeidler 1989; Wald 1950), maxmax expected utility (Drèze 1961, 1987, Chs. 2 and 3), and the α-maxmin model (Ghirardato et al. 2004; Jaffray 1994; Luce and Raiffa 1957, Ch. 13). These theories, used descriptively here, were primarily developed for normative purposes, assuming expected utility for risk and an Anscombe and Aumann (1963) model. The descriptive prospect theory of Tversky and Kahneman (1992) was not included by HLM.[3] We add the latter, using its special case of Abdellaoui et al.'s (2011) source method. We find that it outperforms all theories considered by HLM. Thus, prospect theory seems to be the best descriptive theory of ambiguity presently available.

This paper proceeds as follows. Section 1 describes the history and the scarcity of quantitative descriptive studies of ambiguity. Section 2 presents the general theories of ambiguity considered in this paper, mostly multiple priors and rank-dependent models (the latter include PT). Here we also explain where we deviate from HLM's theoretical analysis. Section 3 explains how we reduced the number of parameters for each theory. We used Abdellaoui et al.'s (2011) source method for the rank-dependent theories, and we used HLM's method for multiple priors. Sections 2 and 3 provide guidelines for empirical researchers interested in tractable tests and applications of ambiguity models. We describe the basic experimental design in Section 4, present our results in Section 5, and discuss and conclude in Sections 6 and 7. A Web Appendix provides further details.

---

[2]Although prospect theory is often used to refer to OPT, the new version of 1992 has now replaced OPT and deserves this nontechnical and short name, rather than the technical and long "cumulative prospect theory" or CPT. Our terminology was Tversky's preference.

[3]HLM (p. 87 bottom) explain that the theories that they call prospect theory are not Tversky and Kahneman's (1992) version or any other common version.

1 Ambiguity and its empirical tests

Until the 1990s, quantitative descriptive studies focused on risk, in the absence of quantitative models for ambiguity.[4] These descriptive studies were often carried out in the probability triangle (probability distributions over three fixed outcomes). Since 1990, quantitative ambiguity models have become available (Gilboa 1987; Gilboa and Schmeidler 1989; Schmeidler 1989), and many extensions have been proposed after (reviewed by Etner et al. 2012). These models were all normatively motivated (Gilboa et al. 2012, p. 18 3rd para), assuming expected utility for risk, backward induction in an Anscombe and Aumann (1963) model, and (with α-maxmin excepted) universal ambiguity aversion. These assumptions, even if accepted normatively, fail descriptively.[5]

Some quantitative descriptive studies fitted data for simple situations with no more than two events.[6] Until recently, comparative studies of ambiguity models were qualitative, testing particular predictions. Ahn et al. (2013) compared kinked models (mostly multiple priors and rank-dependent models) with smooth models and found that the former performed better. Hayashi and Wada (2010) falsified multiple priors models by showing that not only the priors with maximal and minimal expected utility play a role in decisions, but also intermediate priors.

HLM achieved a recent breakthrough: They extended the comparative studies of different nonexpected utility models in the probability triangle from risk to ambiguity. One key step was the introduction of a valuable device to generate ambiguity: a bingo blower. Another step was that they brought together different ambiguity models, making them quantitatively testable for empirically tractable one-stage prospects. HLM did include two models in their competition that they called prospect theory, but these were not Tversky and Kahneman's (1992) model, or any other common model.[7] Hence, HLM did not test Tversky and Kahneman's (1992) PT for ambiguity.

Although HLM succeeded in implementing CEU in full generality, CEU has too many parameters to perform well. We will use Abdellaoui et al.'s (2011) source method to reduce the number of parameters of CEU. We will extend it to PT, the reference-dependent generalization of CEU.

[4]See Bernasconi (1994), Camerer (1989), Harless and Camerer (1994), Hey and Di Cagno (1990), Hey and Orme (1994), Sopher and Gigliotti (1993) and Starmer (1992).

[5]For descriptive violations of backward induction, see Barkan and Busemeyer (1999), Budescu and Fischer (2001), Carbone and Hey (2001), Cubitt et al. (1998), Dominiak et al. (2012), Hey and Knoll (2011), Hey and Lotito (2009), Hey and Panaccione (2011), Rachlin and Green (1972) and Yechiam et al. (2005).

[6]See Andersen et al. (2012), Choi et al. (2007), Hsu et al. (2005), and Huettel et al. (2006). Only for rank dependence and PT have there been some quantitative studies considering more than two events: Abdellaoui et al. (2011), Abdellaoui et al. (2005), Baillon and Bleichrodt (2012) and Diecidue et al. (2007). We focus here on revealed-preference based studies. Hogarth and Einhorn (1990) present and cite influential work using introspective inputs.

[7]Their first version was an extension of OPT to ambiguity. However, OPT has the aforementioned problems, making it ill suited for the three-outcome domain considered here. The second version was Schmeidler's (1989) CEU with a level parameter of utility added to capture loss aversion. However, this second version did not incorporate the sign-dependent and reflected weighting of losses that is typical of prospect theory, and that is needed to make the added level parameter of utility identifiable.

We next discuss multiple priors models. Sets of priors are often used as a tool of communication, such as when describing stimuli in experiments, but they rarely arise exogenously from naturally generated ambiguity. Thus, they do not arise exogenously with the bingo blower. They are, accordingly, hard to implement in incentivized experiments with guaranteed absence of deception. The focus of this paper is, hence, on endogenous sets of priors. These sets comprise even more parameters than PT or CEU (Eron and Schmeidler 2012). They have an even greater need for a specification of tractable submodels before they can become implementable. An important contribution of HLM was the introduction of such a tractable submodel. To the best of our knowledge, HLM were the first to reveal endogenous sets of priors from preferences when there are more than two events. We will use their submodel in our analysis (Subsection 3.3).

Because our stimuli do not have exogenously given separate stages, we will, following HLM, not consider two-stage models of ambiguity such as those in Ergin and Gul (2009), Nau (2006) or Neilson (2010). Taking the stages endogenous, as in Klibanoff et al.'s (2005) smooth model, would lead to even more free parameters: the collection of all second-order probability distributions over the (first-order) probability measures over S is of higher dimensionality than, for instance, the collection of all sets of priors. Developing tractable subcases of the smooth model is a topic for future research.

2 Definitions of the general deterministic decision theories tested

HLM used a bingo blower containing pink, blue, and yellow balls in proportions unknown to the subjects. Further experimental details are in Section 3. A ball was drawn at random and its color determined the outcome. Only three possible outcomes were considered, namely, £100, £10, and −£10. Further experimental details are in Section 4.

We write (E_1, E_2, E_3) for the *prospect* yielding £100 under event E_1, £10 under event E_2, and −£10 under E_3. The *events* E_1, E_2, and E_3 partition the set of possible colors, denoted $S = \{pink, blue, yellow\}$. For example, if $E_1 =$ yellow, $E_2 = \emptyset$, and $E_3 =$ {pink, blue}, then the prospect yields £100 for yellow, −£10 for pink and blue, and outcome £10 never occurs. Subjects made choices between pairs of prospects. We discuss the main theories in HLM, and examine in detail which of these theories performs best. In all these theories, subjects maximize

$$\pi_1 U(£100) + \pi_2 U(£10) + \pi_3 U(-£10), \quad (1)$$

with U a *utility function*, and the πs *decision weights*. The πs are nonnegative and will sum to 1 except under PT. In all theories considered, the unit of utility is free to choose, and so is the level except under PT.

Under (*subjective*) *expected utility* (*EU*), there exists a (*subjective*) *probability measure* P on S such that

$$\pi_j = P(E_j). \quad (2)$$

Under EU, not only the unit but also the level of utility is free to choose. We can always choose $U(£100) = 1$ and $U(-£10) = 0$. $U(£10)$ is the only parameter of utility to be determined. (*Subjective*) *expected value* (*EV*) is the special case where U is linear.

Under *Choquet expected utility* (*CEU*), there exists a *weighting function* W ($W(\emptyset) = 0$, $W(S) = 1$, and $A \supset B \Rightarrow W(A) \geq W(B)$) such that

$$\pi_1 = W(E_1), \pi_2 = W(E_1 \cup E_2) - W(E_1), \pi_3 = 1 - W(E_1 \cup E_2). \quad (3)$$

The decision weights π_j add to 1, implying that U is unique up to unit and level as it is under EU. Again, only one parameter of utility remains to be determined.

Under *prospect theory* (*PT*), there exist weighting functions W^+ and W^- such that

$$\pi_1 = W^+(E_1), \pi_2 = W^+(E_1 \cup E_2) - W^+(E_1), \pi_3 = W^-(E_3). \quad (4)$$

Relative to CEU, the loss is weighted using W^- instead of W^+, and the weighting of losses is done in a dual manner. Now the decision weights need no longer add to 1 or to another constant. Hence, the level of utility U is not free to choose, and is uniquely determined in sufficiently rich models.[8] For Eq. 4, the scaling $U(0) = 0$ is imposed and it is essential.[9] Any other scaling leads to incorrect formulas. If we set $U(£100) = 1$ which is always possible, then both $U(£10)$ and $U(-£10)$ are empirically meaningful parameters. Utility for losses is often decomposed into basic utility u and loss aversion λ, with $U(-£10) = \lambda u(-£10)$. For gains, we then set $u = U$. Prospect theory and CEU, and several specifications introduced later, are called *rank-dependent* theories.

Multiple priors theories consider a set of *priors* Π—a set of probability distributions over S. Π may be assumed to be closed and convex. *α-maxmin EU* assumes a number $0 \leq \alpha \leq 1$, and maximization of the functional

$$\alpha \min_{P \in \Pi} EU + (1 - \alpha) \max_{P \in \Pi} EU. \quad (5)$$

The *maxmin EU* model is the special case of $\alpha = 1$, and the *maxmax EU* model is the special case of $\alpha = 0$.[10] The decision weights in Eq. 1 correspond with α times the minimizing probability plus $(1 - \alpha)$ times the maximizing probability from Π and, hence, always sum to 1. Utility is, therefore, unique up to level and unit as it is under EU and CEU.

HLM further consider a *maxmin* (*maxmax*) theory that evaluates a prospect lexicographically, starting with the minimal (maximal) outcome, and a *minreg* (minimax regret) theory. We refer to HLM for details.

---

[8] Theoretical claims of uniqueness and identifiability are usually derived under the assumption of continuums of domains. For discrete observations as always obtained in practice, those claims provide a lower bound for uniqueness and an upper bound for identifiability.

[9] Although the outcome 0 does not occur in our domain, it is essential for the decision weights that $U(x)$ reflects the distance from x to 0 in utility units, for all x.

[10] We use the terms maxmin EU and maxmax EU instead of HLM's terms G&S maxmin and G&S maxmax because our terms are common in the literature.

3 Reducing the numbers of parameters

This section describes how we reduced the number of parameters of the main general theories defined in the preceding section, so as to achieve the tractability and parsimony required for empirical applications.

3.1 SCEU: a parsimonious version of CEU

Under general CEU, the weighting function W comprises six parameters, being the weights of the eight events except Ø and S. Data fitting is still theoretically possible for CEU in full generality, and HLM did so, but the results were not very good. Many authors have indeed pointed out that nonadditive measures are too general to be tractable, and that special subcases must be developed.[11] Abdellaoui et al.'s (2011) *source method* achieves this as follows: For a given subject in a given treatment, there exist subjective probabilities P over S, and a *source function* w mapping [0, 1] to [0, 1], strictly increasing and continuous, with $w(0) = 0$ and $w(1) = 1$. Then $W(E) = w(P(E))$. By allowing w to be different for different sources of uncertainty (such as for the known Ellsberg urn versus the unknown Ellsberg urn), the source method generalizes Machina and Schmeidler's (1992) probabilistic sophistication. By deviating from the probability weighting function for risk, w captures ambiguity attitudes.

So as to limit the number of parameters considered, we only use one parameter for the source function w, by using Prelec's (1998) family:

$$w(p) = exp(-(-ln(p))^{\alpha}). \tag{6}$$

The function is linear for $\alpha = 1$, is more inverse-S shaped as α is more below 1, and is more S-shaped as α is more above 1. There is one popular alternative one-parameter family, by Tversky and Kahneman (1992). In Section 5.3, we discuss this alternative, and other two-parameter alternatives.

We emphasize that the source function w and its parameter α comprise ambiguity attitudes. The probabilities p are subjective, are derived from behavior, and may not be known to subjects in any conscious manner. In our experiment they will deviate from the objective probabilities that we know but that the subjects did not know. In general, we can expect an ambiguous source function to deviate more from linearity than the weighting function for objective probabilities, where the extra curvature of the source function reflects ambiguity. In general, ambiguity amplifies the phenomena under risk (see Wakker's 2010 review (p. 292) and Maafi 2011).

Because U is unique up to level and unit, we set $U(£100) = 1$ and $U(-£10) = 0$, and $U(£10)$ is the only utility parameter. Under linearity, $U(£10)) = 2/11$. The preference family resulting from these restrictions on CEU preferences is called the *source-CEU (SCEU)* model. It has reduced the number of parameters for W from six to three: two for P (P(pink) and P(blue)) and one for w (parameter α). SCEU has

[11] See Chateauneuf et al. (2007), Grabisch and Labreuche (2008, Sections 2.7 and 7), Ivanov (2011, p. 367), Narukawa and Torra (2011) and Wakker (2010, Section 11.2)

five parameters in total: besides the three for W, there is one for U (U(£10)) and one for the error variance (σ) that will be explained later.

3.2 SPT: a parsimonious version of PT

PT in its full generality generalizes CEU by allowing a separate treatment of losses, with separate parameters. To limit the number of parameters, SPT, our version of PT, will have five parameters as did SCEU. It will modify rather than generalize SCEU.

We again use Abdellaoui et al.'s (2011) source method for the decision weights π_j. Thus, we assume subjective probabilities P. Now

$$W^+ = w^+(P(E)) \; and \; W^- = w^-(P(E)). \tag{7}$$

To limit the number of parameters, we assume $w^- = w^+$, implying $W^- = W^+$. We again use Prelec's one-parameter family of Eq. 6.

Given the dual weighting of losses of PT in Eq. 4, the equality $W^- = W^+$ leads to a reflected weighting of losses. It is the most common special case of PT, deviating from CEU. Now π_3 of Eq. 1 is $W(E_3)$ ($= w(P(E_3))$) rather than $1 - W(E_1 \cup E_2)$. Uncertainty attitudes for losses are reflected relative to gains rather than being identical.

Because the decision weights do not add to 1 or another constant, utility U is a ratio scale. That is, given the requirement $U(0) = 0$, we can still choose $U(£100) > 0$ freely. We set it equal to 1. But now both $U(£10)$ and $U(-£10)$ are meaningful parameters, at least in theory. It is, however, empirically plausible that utility is approximately linear for small stakes of the same sign. The main curvature of utility is generated by loss aversion, giving a kink at 0. Hence, we assume u linear, with $U(£10) = u(£10) = 1/10$ and $U(-£10) = \lambda u(-£10) = -\lambda/10$. Thus, we only use one utility parameter, being loss aversion λ. Such a piecewise linear utility U was recommended by Köbberling and Wakker (2003, end of Section 5), and was used, for instance, by Barberis et al. (2001, p. 17), Fehr and Schmidt (1999), Gächter et al. (2007), Rosenblatt-Wisch (2008) and Wakker (2010, Example 9.4.2).

The resulting special case of PT is called the *source-PT (SPT)* model. It has five parameters: two for P (P(pink) and P(blue)), one for w (α), one for U (λ), and one for the error variance (σ).

3.3 HLM's parsimonious version of multiple priors

In their full generality, the multiple priors models have too many parameters to be tractable, even more than CEU or PT have. With the set of probability measures (priors) over S two-dimensional, the collection of all its closed convex subsets Π is infinite-dimensional. Recall that the set of priors is not exogenously given in our case, but must be derived from preference. Further, even with a set Π given, the minimization and maximization in Eq. 5 can be difficult to implement for general sets Π. For multiple priors models, even more than for the CEU or PT models, finding tractable subfamilies is necessary for the purpose of obtaining quantitative measurements.[12]

To our best knowledge, HLM are the first to come up with a tractable subfamily of multiple priors, making it possible to do a quantitative empirical measurement. They consider sets of priors

$$\Pi = \{P : P(\text{pink}) \geq \epsilon_1,\ P(\text{blue}) \geq \epsilon_2,\ P(\text{yellow}) \geq \epsilon_3\}, \tag{8}$$

with the ϵ_js nonnegative and summing to less than 1.[13] This turns the set of all sets of priors into a three-parameter space, which is tractable enough for fitting and prediction.

The resulting versions of maxmin and maxmax EU, called *MxEU* and *MnEU*, have five parameters: three for Π, one for U, and one for the error variance (σ). The resulting version of the α-maxmin EU model, called αMM, has six parameters, including α.

4 Experimental design

The bingo blower used by HLM contained pink, blue, and yellow balls in 0.2, 0.3, and 0.5 proportions. There were three treatments (between subjects), with a total number of 10, 20, and 40 balls in the bingo blower, respectively. Subjects were not informed about these numbers but could watch the bingo blower. The more balls in total, the harder it was for subjects to guess the proportions, and the more ambiguity they perceived. Given that only three possible outcomes were considered (£100, £10, and −£10), the domain considered is an extension of the probability triangle from risk to ambiguity.

$N = 48$ subjects participated, 15 in treatment 1 (10 balls in total), 17 in treatment 2 (20 balls in total), and 16 in treatment 3 (40 balls in total). The subjects made 162 choices between pairs of prospects, which basically involved all nontrivial choices, with one choice randomly selected and played for real. For each theory, 135 choices were used for fitting the models, that is, finding the parameters that optimized the likelihood of those 135 choices. The remaining 27 choices were used as a prediction (test) set.

The comparisons between the models are based on the predicted log-likelihood of the test set. The *predicted log-likelihood* of a model is the logarithm of the probability of the observed decisions on the test set, using the parameters of the model estimated by maximum likelihood on the training set. The higher the predicted log-likelihood,

[12]It is even more so for generalizations of multiple priors models, including Chateauneuf and Faro (2009), Gajdos et al. (2008), Maccheroni et al. (2006), Nascimento and Riella (2010), Strzalecki (2011) and Siniscalchi (2009). Ghirardato et al. (2004) and Siniscalchi (2006) gave preference conditions, assuming an Anscombe-Aumann (1963) model, to determine whether or not a particular probability measure is contained in the set of priors. This requires a two-stage setup and then needs infinitely many observations (one for each possible prior) to determine the set of priors.

[13]Although HLM do not state it explicitly (footnote 16 and p. 109), the boundaries ϵ_j are lower bounds and are not upper bounds.

the better the theory. This measure of performance is widely used in statistics. It is well-known that increasing the number of parameters of a model beyond the optimal number will worsen its prediction, because the added parameters start picking up more noise than systematic factors from the training set (overfitting).

HLM combine the derministic decision theories with two error theories, one Fechnerian and one contextual. As explained by HLM, the two are virtually identical for the stimuli considered here. HLM put the most common one, the Fechner error theory, central in their discussions, and we will use it too. It involves one extra parameter (σ), for the error variance. HLM consider one extra decision theory, a version of *decision field theory* (*DFT*), which agrees with EU for its deterministic part but adds its own specific error theory. For details see HLM (p. 92, 108–109). Wilcox (2008) surveys alternative error theories.

5 Results

5.1 The analyses of HLM with a modified version of prospect theory

Table 1 displays the average predicted log-likelihoods of the various theories considered. It reproduces Table 1 of HLM, with only their Fechner error theory, and with their two versions of prospect theory replaced by SPT.[14] SPT gives the best result, and MxEU, the winner in HML, is now second.

Table 2 similarly reproduces Table 2 of HLM. It indicates, for each theory, for how many subjects this theory gives the best result.[15] Relative to HLM, we again replaced HLM's two versions of prospect theory by SPT. SPT again has the best result. For HLM, αMM was best, winning for 11 subjects, and MxEU was second, winning for 8 subjects. Especially the former loses many subjects to SPT. DFT now outperforms both theories.

A drawback of Table 2 is that the performance of a theory depends largely on how many similar theories are competing. Similar theories will fight for the same subjects, and an outside theory can run away with the victory. This happened during the 2000 US presidential elections when Nader caused Gore to lose to Bush. The mentioned drawback can be avoided by comparing theories pairwise, and examining how each single theory gains (being better for the majority of subjects) or loses to each other single theory—that is, by considering a tournament between theories. This is the topic of the next subsection.

A drawback of both Tables 1 and 2 is that no significance levels are given, so that it is not clear whether the inequalities found are systematic or due to chance. The analysis in the next subsection will provide significance levels.

---

[14]HLM (footnote 18) point out that subject 35 was an outlier, performing very poorly for αMM and MnEU, and greatly influencing the average likelihoods of those models. Hence, HLM left this subject out. Although this means favoring the models affected, we follow HLM and also leave this subject out from Table 1. With subject 35 incorporated, SPT would still win, αMM and MnEU would be among the worst theories, and the other theories would not be affected seriously; see Table WA1 in the Web Appendix.

[15]Now subject 35 is included, again following HLM.

Table 1 Mean predicted log-likelihoods for the three treatments, and overall

	SPT	MxEU	MnEU	αMM	EU	DFT	CEU	EV	MaxMin	MaxMax	MinReg
All	−3.83[b]	−3.92	−4.31	−4.32	−4.43	−4.57	−5.53	−11.50	−13.31	−14.06	−14.58
Tr 1	−2.60[b]	−2.82	−3.13	−2.79	−3.44	−3.50	−3.22	−12.15	−12.57	−14.52	−14.48
Tr 2	−5.16	−5.14[b]	−5.70	−5.58	−5.61	−5.93	−6.35	−11.03	−14.17	−14.45	−14.86
Tr 3	−3.90[b]	−3.96	−28.73	−29.19	−4.39	−4.53	−8.68	−11.34	−13.14	−12.94	−14.29

The biggest (least negative) log-likelihood, indicated by a superscript b, is the best in each row; SPT: prospect theory with the source method, using Eqs. 6 and 7 and the other assumptions in Section 3.2; MxEU [MnEU; αMM]: maxmax [maxmin; αMM] expected utility using Eq. 8; EU: subjective expected utility; DFT: decision field theory, defined in HLM; CEU: Choquet expected utility (Eq. 3); EV: subjective expected value; MaxMin [maxMax]: evaluation by minimal [maximal] outcome; MinReg: minimax regret (see HLM)

5.2 Statistical analyses

Table 3 presents the results of a tournament between theories. SPT wins, beating every other theory. MxEU is the best of the non-rank-dependent theories, beating every other non-rank-dependent theory. Table 3 gives significance levels. We did not use *t*-tests because there were outliers in predicted log-likelihoods that skew the distributions for virtually every model considered. Instead, we used a two-sided Wilcoxon rank-signed test, which is valid irrespective of the underlying distributions. SPT's victory over MxEU is not significant and may be interpreted as a draw. There are several other draws.

The many outliers of log-likelihoods that invalidate *t*-tests also complicate the interpretations of the averages in Table 1. Hence, medians and trimmed means provide useful alternative information. Table 4 adds such indicators of central tendency. The table confirms our findings based on Wilcoxon tests. This table (and other analyses not reported here) suggests that a mean trimmed at 10% represents the data well. It orders the theories as reported in Table 3, which closely follows the pairwise significance tests.

Table 2 Number of subjects for whom a theory predicts best

	SPT	DFT	CEU	MxEU	αMM	EU	MnEU	MaxMax	EV	MaxMin	MinReg
All	10[b]	8	7	7	6	5	4	1	0	0	0
Tr 1	2	3	3	4[b]	1	0	2	0	0	0	0
Tr 2	5[b]	2	3	1	2	4	0	0	0	0	0
Tr 3	3[b]	3[b]	1	2	3[b]	1	2	1	0	0	0

The theories are as in Table 1. The biggest number, indicated by a superscript b, is the best in each row

Table 3 Pairwise comparison of theories

	SPT	MxEU	αMM	MnEU	EU	DFT	CEU	EV	MaxMin	MaxMax
MxEU	20^{28}	–	–	–	–	–	–	–	–	–
αMM	18^{30}*	21^{16}	–	–	–	–	–	–	–	–
MnEU	13^{35}**	18^{30}*	18^{27}	–	–	–	–	–	–	–
EU	15^{33}**	16^{31}*	16^{32}	15^{29}	–	–	–	–	–	–
DFT	16^{32}**	15^{33}**	17^{31}	18^{30}	19^{29}	–	–	–	–	–
CEU	14^{34}***	15^{32}**	12^{34}***	22^{25}	23^{24}	26^{22}	–	–	–	–
EV	2^{46}***	1^{47}***	5^{43}***	3^{45}***	2^{46}***	5^{43}***	8^{40}***	–	–	–
MaxMin	1^{47}***	1^{47}***	4^{44}***	1^{47}***	0^{48}***	1^{47}***	5^{43}***	16^{32}	–	–
MaxMax	1^{47}***	1^{47}***	3^{45}***	2^{46}***	2^{46}***	3^{45}***	4^{44}***	10^{37}***	18^{29}	–
MinReg	0^{48}***	0^{48}***	2^{46}***	1^{47}***	0^{48}***	0^{48}***	3^{45}***	13^{35}***	17^{31}*	20^{28}

The theories are as in Table 1 Counts m^n mean that the row model fits better for m subjects and the column model fits better for n subjects. For example, SPT beats MxEU by 28 to 20. The fit is measured by predicted log-likelihoods and significance levels are conventional (* $<$ .05, ** $<$ .01, *** $<$ .001)

5.3 Variations of prospect theory

We consider some variations of SPT to analyze the role of the various parameters used. Glöckner and Pachur (2012) present a similar analysis for decision under risk. Table 5 displays the resulting "within-PT" tournament. Dropping the probability weighting parameter ($\alpha = 1$) leads to EU and clearly worsens the prediction. Dropping the loss aversion parameter ($\lambda = 1$: $\text{SPT}_{\lambda=1}$) worsens the prediction even more, illustrating the importance of loss aversion. Hence, no parameters should be dropped from SPT.

Adding a utility parameter $U(£10)$ (SPT_u) leaves the overall predictive power almost unaffected. General sign dependence, by allowing a different probability weighting parameter for losses ($\alpha^- \neq \alpha^+$) than for gains ($\text{SPT}_\pm$), similarly leaves the overall predictive power almost unaffected. We prefer the parsimony of SPT in these cases, using a minimal number of parameters.

Relative to SPT, SCEU does not reflect the weighting of losses, but uses the regular rank-dependent weighting there.[16] This worsens the prediction somewhat, with SPT scoring better against the other theories (further illustrated in Table WA3 of the Web Appendix), showing that the reflected weighting of PT is preferable to the regular

[16]SCEU can be seen to be the special case of SPT with dual, rather than identical, weighting for losses, which is why we incorporate it in this subsection. Because utility then becomes unique up to unit and level, the one parameter that we used for utility under SPT (through $-£10$) is equivalent to the one parameter used for utility in CEU.

Table 4 Means, trimmed means and medians for all models (sorted on trimmed mean$_{.1}$)

	Mean$_{.1}$	Mean$_{.05}$	Mean	Median
SPT	−3.53[b]	−3.62[b]	−3.83[b]	−3.32
MxEU	−3.64	−3.75	−3.92	−3.30[b]
αMM	−3.84	−3.96	−4.32	−3.41
MnEU	−4.15	−4.20	−4.31	−4.10
EU	−4.34	−4.36	−4.43	−4.33
CEU	−4.47	−4.67	−5.53	−3.85
DFT	−4.51	−4.54	−4.57	−4.11
EV	−11.82	−11.73	−11.50	−11.80
MaxMin	−13.56	−13.43	−13.31	−14.26
MaxMax	−14.46	−14.27	−14.06	−14.44
MinReg	−14.65	−14.62	−14.58	−14.93

The theories are as in Table 1. The biggest (least negative), indicated by a superscript b, is the best in each column

rank-dependent weighting of CEU. Dropping the utility parameter from SCEU by taking utility linear (SCEV) considerably worsens the prediction.

We considered the following popular weighting functions as source functions instead of Eq. 6.

$$w(p) = (exp(-(-ln(p))^{\alpha}))^{\beta} \text{ [Prelec two-parameter; } SPT_2] \tag{9}$$

$$w(p) = \alpha + \beta p \ \ (0 < p < 1) \text{ [Neo-additive; } SPT_{NA}] \tag{10}$$

$$w(p) = \frac{p^{\alpha}}{(p^{\alpha} + (1-p)^{\alpha})^{1/\alpha}} \text{ [Tversky and Kahneman 1992; } SPT_{TK}] \tag{11}$$

$$w(p) = \frac{\beta p^{\alpha}}{(\beta p^{\alpha} + (1-p)^{\alpha})} \text{ [Goldstein and Einhorn 1987; } SPT_{GE}] \tag{12}$$

They all lead to somewhat worse predictions than SPT, but not by much. GPT (generalizated PT), finally, drops the restrictions of the source method and takes a general weighting function, adding three parameters to SPT. It performs poorly, which comes as no surprise given its many redundant parameters. The completely general case of PT with a sign-dependent W^- that can be chosen independently from W^+ and with utility also general (U(£10) as an extra parameter) has even more redundant parameters. Hey et al. (2011) showed that it performs poorly.

We compared the variations of SPT considered in this section with the other, non-rank-dependent models, in Table 6. It shows that most variations of SPT, although some worse than SPT, still outperform the other theories. Hence, our conclusion that prospect theory best predicts choice under ambiguity is not very sensitive to the particular parametrization chosen.

Table 5 Pairwise comparison of the variations of Prospect Theory

	SPT	$SPT_{\pm}$	SPT_u	SCEU	SPT_{NA}	SPT_2	SPT_{GE}	SPT_{TK}	EU	GPT	SCEV
$SPT_{\pm}$	23^{25}	–	–	–	–	–	–	–	–	–	–
SPT_u	22^{26}	24^{24}	–	–	–	–	–	–	–	–	–
SCEU	24^{24}	25^{23}	25^{23}	–	–	–	–	–	–	–	–
SPT_{NA}	23^{25}	23^{25}	25^{23}	26^{22}	–	–	–	–	–	–	–
SPT_2	17^{31}*	17^{31}*	20^{28}	18^{30}	23^{25}	–	–	–	–	–	–
SPT_{GE}	14^{34}**	15^{33}**	18^{30}	17^{31}	21^{27}	21^{27}	–	–	–	–	–
SPT_{TK}	17^{31}**	18^{30}**	18^{30}*	17^{31}**	20^{28}	18^{30}	20^{28}	–	–	–	–
EU	15^{33}**	16^{32}*	18^{30}*	17^{31}*	16^{32}	18^{30}	19^{29}	19^{29}	–	–	–
PT	13^{35}***	13^{35}***	15^{33}***	14^{34}***	15^{33}***	14^{34}***	14^{34}***	17^{31}**	23^{25}	–	–
SCEV	1^{47}***	0^{48}***	1^{47}***	1^{47}***	5^{43}***	2^{46}***	1^{47}***	4^{44}***	4^{44}***	9^{39}***	–
$SPT_{\lambda=1}$	1^{47}***	0^{48}***	1^{47}***	1^{47}***	5^{43}***	2^{46}***	1^{47}***	3^{45}***	4^{44}***	9^{39}***	18^{30}

The interpretation of the counts and significance levels are as in Table 3. For example, SPT is better than $SPT_{\pm}$ for 25 subjects, and it is worse for 23 subjects. The PT theories are as follows: SPT: prospect theory with the source method, using Eqs. 6 and 7 and the other assumptions in Section 3.2. SCEU: Choquet expected utility with the source method; EU: expected utility; GPT: prospect theory with a general (but sign-independent) weighting function; SCEV: Choquet expected utility with the source method and linear utility; Subscripts to SPT: $\pm$: sign-dependence of weighting functions; u: nonlinear utility; NA: Eq. 10; 2: Eq. 9; GE: Eq. 12; TK: Eq. 11

6 Discussion

SPT performed better than SCEU, showing that the reflected weighting of outcomes at opposite sides of the reference point, as in Eq. 4 (with $W^- = W^+ = W$) outperforms the unreflected weighting of Eq. 3. Many studies have confirmed reflection not only for risk, but also for ambiguity. There is prevailing ambiguity seeking rather than ambiguity aversion for losses (Chesson and Viscusi 2003; for a review see Wakker 2010, Section 12.7). Wakker (2010, following Eq. 9.7.1) argued that the PT functional, which is, in fact, the Šipoš (1979) integral, also has a number of mathematical advantages over the CEU functional, which is the Choquet integral.

In our data fitting for SCEU, there is considerable nonlinearity of utility, with the parameter $U(£10)$ greatly improving the predictions over SCEV. The analysis of SPT suggests that this is mainly because the utility parameter substitutes for loss aversion. This suggestion supports Rabin's (2000) view that much of the utility curvature found in classical (reference-independent) empirical studies may, in fact, be due to a (mis)modeling of loss aversion (Wakker 2010, p. 244 and p. 267 top).

It is remarkable that HLM's DFT performs relatively well. It only deviates from EU by using a more sophisticated error theory that, apparently, works better than the other error theories. It will be interesting to combine this error theory with other theories.

Table 6 Variations of PT versus other models

	SPT	SPT$_{\pm}$	SPT$_u$	SCEU	SPT$_{NA}$	SPT$_2$	SPT$_{GE}$	SPT$_{TK}$	GPT
MxEU	20^{28}	20^{28}	24^{24}	23^{25}	18^{30}	23^{25}	22^{26}	25^{23}	$\overset{19}{29}{}^{**}$
αMM	$\overset{30}{18}{}^{**}$	$\overset{29}{19}{}^{**}$	21^{27}	22^{26}	23^{25}	25^{23}	26^{22}	28^{20}	$\overset{13}{35}{}^{***}$
MnEU	$\overset{35}{13}{}^{***}$	$\overset{34}{14}{}^{***}$	$\overset{30}{18}{}^{**}$	$\overset{31}{17}{}^{**}$	$\overset{31}{17}{}^{**}$	20^{28}	20^{28}	19^{29}	28^{20}
EU	$\overset{33}{15}{}^{***}$	$\overset{32}{16}{}^{**}$	$\overset{30}{18}{}^{**}$	$\overset{31}{17}{}^{**}$	16^{32}	18^{30}	19^{29}	19^{29}	25^{23}
DFT	$\overset{32}{16}{}^{***}$	$\overset{32}{16}{}^{***}$	$\overset{30}{18}{}^{***}$	$\overset{30}{18}{}^{**}$	19^{29}	18^{30}	18^{30}	20^{28}	23^{25}
CEU	$\overset{34}{14}{}^{***}$	$\overset{34}{14}{}^{***}$	$\overset{33}{15}{}^{***}$	$\overset{33}{15}{}^{***}$	$\overset{32}{16}{}^{***}$	$\overset{29}{19}{}^{**}$	$\overset{31}{17}{}^{**}$	20^{28}	$\overset{20}{28}{}^{*}$
EV	$\overset{46}{2}{}^{***}$	$\overset{46}{2}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{41}{7}{}^{***}$	$\overset{43}{5}{}^{***}$	$\overset{44}{4}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{39}{9}{}^{***}$
MaxMin	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{44}{4}{}^{***}$	$\overset{44}{4}{}^{***}$	$\overset{44}{4}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{41}{7}{}^{***}$
MaxMax	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{44}{4}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{46}{2}{}^{***}$	$\overset{46}{2}{}^{***}$	$\overset{43}{5}{}^{***}$
MinReg	$\overset{48}{0}{}^{***}$	$\overset{48}{0}{}^{***}$	$\overset{48}{0}{}^{***}$	$\overset{48}{0}{}^{***}$	$\overset{45}{3}{}^{***}$	$\overset{46}{2}{}^{***}$	$\overset{46}{2}{}^{***}$	$\overset{47}{1}{}^{***}$	$\overset{42}{6}{}^{***}$

The row-theories are as in Table 1, and the column-theories are as in Table 5 The entries are as in Table 3. For example, SPT is better than MxEU for 28 subjects, and it is worse for 20 subjects

7 Conclusion

This paper tested the predictive empirical power of some theories for decision under ambiguity. Most nonexpected utility theories outperform expected utility. Because prospect theory outperforms Choquet expected utility, its reflection is a desirable modification of rank dependence, not only for risk but also for ambiguity. The main component of utility is loss aversion, and curvature of utility found in other reference-independent theories to a large extent serves to capture loss aversion.

We have demonstrated the importance of choosing an appropriate number of parameters, and have determined what such an appropriate number is. General non-additive measures clearly comprise too many parameters. Many general theories of ambiguity have been introduced in the recent literature besides the ones studied in this paper. Tractable subfamilies of those general theories should be developed, showing a good predictive performance, so that these theories then can be applied empirically.

Prospect theory is the most popular theory for predicting decisions under risk today. We find that, through its source-method specification, it also outperforms other theories for predicting decisions under ambiguity. It is convenient that the same functional can be used for the whole domain of uncertainty. This facilitates the study of ambiguity (the difference between uncertainty and risk), where prospect theory, unlike most other ambiguity theories today, need not commit to the descriptively failing expected utility for risk.

Acknowledgments We are grateful to John Hey, Gianna Lotito, and Anna Maffioletti for providing us with their data set, their analyses, and many explanations.

References

Abdellaoui, M., Baillon, A., Placido, L., Wakker, P.P. (2011). The rich domain of uncertainty: source functions and their experimental implementation. *American Economic Review*, *101*, 695–723.

Abdellaoui, M., Vossmann, F., Weber, M. (2005). Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty. *Management Science*, *51*, 1384–1399.

Ahn, D.S., Choi, S., Gale, D., Kariv, S. (2013). *Estimating ambiguity aversion in a portfolio choice experiment*. Quantitative Economics, forthcoming.

Andersen, S., Fountain, J., Harrison, G.W., Hole, A.R., Rutström, E.E. (2012). Inferring beliefs as subjectively imprecise probabilities. *Theory and Decision*, *73*, 161–184.

Anscombe, F.J., & Aumann, R.J. (1963). A definition of subjective probability. *Annals of Mathematical Statistics*, *34*, 199–205.

Baillon, A., & Bleichrodt, H. (2012). *Testing ambiguity models through the measurement of probabilities for gains and losses*. The Netherlands: Erasmus School of Economics, Erasmus University, Rotterdam.

Barberis, N., Huang, M., Santos, T. (2001). Prospect theory and asset prices. *Quarterly Journal of Economics*, *66*, 1–53.

Barkan, R., & Busemeyer, J.R. (1999). Changing plans: dynamic inconsistency and the effect of experience on the reference point. *Psychonomic Bulletin and Review*, *6*, 547–554.

Bernasconi, M. (1994). Nonlinear preference and two-stage lotteries: theories and evidence. *Economic Journal*, *104*, 54–70.

Budescu, D.V., & Fischer, I. (2001). The same but different: an empirical investigation of the reducibility principle. *Journal of Behavioral Decision Making*, *14*, 187–206.

Camerer, C.F. (1989). An experimental test of several generalized utility theories. *Journal of Risk and Uncertainty*, *2*, 61–104.

Carbone, E., & Hey, J.D. (2001). A test of the principle of optimality. *Theory and Decision*, *50*, 263–281.

Chateauneuf, A. (1991). On the use of capacities in modeling uncertainty aversion and risk aversion. *Journal of Mathematical Economics*, *20*, 343–369.

Chateauneuf, A., Eichberger, J., Grant, S. (2007). Choice under uncertainty with the best and worst in mind: NEO-additive capacities. *Journal of Economic Theory*, *137*, 538–567.

Chateauneuf, A., & Faro, J.H. (2009). Ambiguity through confidence functions. *Journal of Mathematical Economics*, *45*, 535–558.

Chesson, H.W., & Viscusi, W.K. (2003). Commonalities in time and ambiguity aversion for long-term risks. *Theory and Decision*, *54*, 57–71.

Choi, S., Fishman, R., Gale, D., Kariv, S. (2007). Consistency and heterogeneity of individual behavior under uncertainty. *American Economic Review*, *97*, 1921–1938.

Cozman, F.G. (2012). Sets of probability distributions, independence, and convexity. *Synthese*, *186*, 577–600.

Cubitt, P.R., Starmer, C., Sugden, R. (1998). Dynamic choice and the common ratio effect: an experimental investigation. *Economic Journal*, *108*, 1362–1380.

Diecidue, E., Wakker, P.P., Zeelenberg, M. (2007). Eliciting decision weights by adapting de Finetti's betting-odds method to prospect theory. *Journal of Risk and Uncertainty*, *34*, 179–199.

Dominiak, A., Duersch, P., Lefort, J.-P. (2012). A dynamic Ellsberg urn experiment. *Games and Economic Behavior*, *75*, 625–638.

Drèze, J.H. (1961). Les fondements logiques de l'utilité cardinale et de la probabilité subjective. *La Décision* (pp. 73–83). Paris: CNRS.

Drèze, J.H. (1987). *Essays on economic decision under uncertainty*. London: Cambridge University Press.

Ergin, H., & Gul, F. (2009). A theory of subjective compound lotteries. *Journal of Economic Theory*, *144*, 899–929.

Eron, S.A., & Schmeidler, D. (2012). Purely subjective maxmin expected utility. Mimeo.

Etner, J., Jeleva, M., Tallon, J.M. (2012). Decision theory under ambiguity. *Journal of Economic Surveys*, *26*, 234–270.

Fehr, E., & Schmidt, K. (1999). A theory of fairness, competition and cooperation. *Quarterly Journal of Economics*, *114*, 817–868.

Gächter, S., Herrmann, A., Johnson, E.J. (2007). Individual-level loss aversion in riskless and risky choice. Working Paper, University of Nottingham.

Gajdos, T., Hayashi, T., Tallon, J.M., Vergnaud, J.C. (2008). Attitude towards imprecise information. *Journal of Economic Theory*, *140*, 27–65.

Ghirardato, P., Maccheroni, F., Marinacci, M. (2004). Differentiating ambiguity and ambiguity attitude. *Journal of Economic Theory*, *118*, 133–173.

Gilboa, I. (1987). Expected utility with purely subjective non-additive probabilities. *Journal of Mathematical Economics*, *16*, 65–88.

Gilboa, I., & Schmeidler, D. (1989). Maxmin expected utility with a non-unique prior. *Journal of Mathematical Economics*, *18*, 141–153.

Gilboa, I., Postlewaite, A., Schmeidler, D. (2012). Rationality of belief or: Why Savage's axioms are neither necessary nor sufficient for rationality. *Synthese*, *187*, 11–31.

Glöckner, A., & Pachur, T. (2012). Cognitive models of risky choice: parameter stability and predictive accuracy of prospect theory. *Cognition*, *123*, 21–32.

Goldstein, W.M., & Einhorn, H.J. (1987). Expression theory and the preference reversal phenomena. *Psychological Review*, *94*, 236–254.

Grabisch, M., & Labreuche, C. (2008). A decade of application of the Choquet and Sugeno integrals in multi-criteria decision aid. *4OR*, *6*, 1–44.

Harless, D.W., & Camerer, C.F. (1994). The predictive utility of generalized expected utility theories. *Econometrica*, *62*, 1251–1289.

Hayashi, T., & Wada, R. (2010). Choice with imprecise information: an experimental approach. *Theory and Decision*, *69*, 355–373.

Hey, J.D., & Di Cagno, D. (1990). Circles and triangles: an experimental estimation of indifference lines in the Marschak-Machina triangle. *Journal of Behavioral Decision Making*, *3*, 279–306.

Hey, J.D., & Knoll, J.A. (2011). Strategies in dynamic decision making – an experimental investigation of the rationality of decision behaviour. *Journal of Economic Psychology*, *32*, 399–409.

Hey, J.D., & Lotito, G. (2009). Naïve, resolute or sophisticated? A study of dynamic decision making. *Journal of Risk and Uncertainty*, *38*, 1–25.

Hey, J.D., Lotito, G., Maffioletti, A. (2010). The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity. *Journal of Risk and Uncertainty*, *41*, 81–111.

Hey, J.D., Lotito, G., Maffioletti, A. (2011). *The descriptive and predictive adequacy of theories of decision making under uncertainty/ambiguity*. Corrigendum, Mimeo.

Hey, J.D., & Orme, C. (1994). Investigating generalizations of expected utility theory using experimental data. *Econometrica*, *62*, 1291–1326.

Hey, J.D., & Panaccione, L. (2011). Dynamic decision making: what do people do? *Journal of Risk and Uncertainty*, *42*, 85–123.

Hogarth, R.M., & Einhorn, H.J. (1990). Venture theory: a model of decision weights. *Management Science*, *36*, 780–803.

Hsu, M., Bhatt, M., Adolphs, R., Tranel, D., Camerer, C.F. (2005). Neural systems responding to degrees of uncertainty in human decision making. *Science*, *310*, 1680–1683.

Huettel, S.A., Stowe, C.J., Gordon, E.M., Warner, B.T., Platt, M.L. (2006). Neural signatures of economic preferences for risk and ambiguity. *Neuron*, *49*, 765–775.

Ivanov, A. (2011). Attitudes to ambiguity in one-shot normal-form games: an experimental study. *Games and Economic Behavior*, *71*, 366–394.

Jaffray, J.-Y. (1994). Dynamic decision making with belief functions. In R.R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.) *Advances in the Dempster-Shafer theory of evidence* (pp. 331–352). New York: Wiley.

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

Klibanoff, P., Marinacci, M., Mukerji, S. (2005). A smooth model of decision making under ambiguity. *Econometrica*, *73*, 1849–1892.

Köbberling, V., & Wakker, P.P. (2003). Preference foundations for nonexpected utility: a generalized and simplified technique. *Mathematics of Operations Research*, *28*, 395–423.

Luce, R.D., & Raiffa, H. (1957). *Games and decisions*. New York: Wiley.
Maafi, H. (2011). Preference reversals under ambiguity. *Management Science*, *57*, 2054–2066.
Maccheroni, F., Marinacci, M., Rustichini, A. (2006). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica*, *74*, 1447–1498.
Machina, M.J., & Schmeidler, D. (1992). A more robust definition of subjective probability. *Econometrica*, *60*, 745–780.
Narukawa, Y., & Torra, V. (2011). On distorted probabilities and m-separable fuzzy measures. *International Journal of Approximate Reasoning*, *52*, 1325–1336.
Nascimento, L., & Riella, G. (2010). On the uses of the monotonicity and independence axioms in models of ambiguity aversion. *Mathematical Social Sciences*, *59*, 326–329.
Nau, R.F. (2006). Uncertainty aversion with second-order utilities and probabilities. *Management Science*, *52*, 136–145.
Neilson, W.S. (2010). A simplified axiomatic approach to ambiguity aversion. *Journal of Risk and Uncertainty*, *41*, 113–124.
Prelec, D. (1998). The probability weighting function. *Econometrica*, *66*, 497–527.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behaviour and Organization*, *3*, 323–343.
Rabin, M. (2000). Risk aversion and expected-utility theory: a calibration theorem. *Econometrica*, *68*, 1281–1292.
Rachlin, H., & Green, L. (1972). Commitment, choice and self-control. *Journal of the Experimental Analysis of Behavior*, *17*, 15–22.
Rosenblatt-Wisch, R. (2008). Loss aversion in aggregate macroeconomic time series. *European Economic Review*, *52*, 1140–1159.
Savage, L.J. (1954). *The foundations of statistics*. New York: Wiley (Second edition 1972, New York: Dover Publications.)
Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica*, *57*, 571–587.
Siniscalchi, M. (2006). A behavioral characterization of plausible priors. *Journal of Economic Theory*, *128*, 91–135.
Siniscalchi, M. (2009). Vector expected utility and attitudes toward variation. *Econometrica*, *77*, 801–855.
Šipoš, J. (1979). Integral with respect to a pre-measure. *Math. Slovaca*, *29*, 141–155.
Sopher, B., & Gigliotti, G. (1993). A test of generalized expected utility theory. *Theory and Decision*, *35*, 75–106.
Starmer, C. (1992). Testing new theories of choice under uncertainty using the common consequence effect. *Review of Economic Studies*, *59*, 813–830.
Strzalecki, T. (2011). Axiomatic foundations of multiplier preferences. *Econometrica*, *79*, 47–73.
Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, *5*, 297–323.
Wakker, P.P. (2010). *Prospect theory for risk and ambiguity.* Cambridge: Cambridge University Press.
Wald, A. (1950). *Statistical decision functions.* New York: Wiley.
Wilcox, N.T. (2008). Stochastic models for binary discrete choice under risk: a critical primer and econometric comparison. In J.C. Cox, & G.W. Harrison (Eds.) *Risk aversion in experiments; Research in experimental economics* (Vol. 12, pp. 197–292). Emerald, Bingley.
Yechiam, E., Stout, J.C., Busemeyer, J.R., Rock, S.L., Finn, P.R. (2005). Individual differences in the response to forgone payoffs: an examination of high functioning drug abusers. *Journal of Behavioral Decision Making*, *18*, 97–110.