Measuring Covariation between Preference Parameters:

A Simulation Study

Frank Vossmann[a], Martin Weber[b]

March, 2005

[a] Lehrstuhl für Bankbetriebslehre, Universität Mannheim, 68131 Mannheim – Germany, e-Mail: vossmann@bank.bwl.uni-mannheim.de; [b] Lehrstuhl für Bankbetriebslehre, Universität Mannheim, 68131 Mannheim – Germany, e-Mail: weber@bank.bwl.uni-mannheim.de

Abstract

Much of the empirical success of Rank-Dependent Expected Utility Theory and Cumulative Prospect Theory is due to the fact that they allow for nonlinearity towards both outcomes (through the utility function) and probabilities (through the probability weighting function). Since risk attitude is jointly determined by the shapes of the two functions, it would be instructive to measure how the degree of risk aversion incorporated in the utility function empirically covaries with its counterpart from the probability weighting function. We conduct a large-scale simulation to assess whether an elicitation procedure based on the trade-off method, which essentially equals that used in recent empirical studies, allows to reliably measure the quantity of interest. We find a strong systematic distortion of measurement, which points at the limitations of the presently available elicitation techniques.

Keywords

Decision under Risk and Uncertainty, Rank-Dependent Expected Utility Theory, Cumulative Prospect Theory, Probabilistic Risk Attitude, Simulation.

JEL classification: D81

Introduction

The classical normative model for decision under risk is Expected Utility (EU) theory developed by von Neumann and Morgenstern (1944, 1947). The famous contribution by Allais (1953) and numerous experimental studies thereafter (for recent surveys, see e.g. Starmer 2000 and Wu et al. 2004), however, challenged the descriptive validity of EU. As a response, descriptive theories of choice were proposed that could account for many of the systematic deviations from EU. Among the most influential of these are Rank-Dependent Expected Utility (RDEU) theory (Quiggin 1982, Yaari 1987) and Cumulative Prospect Theory (CPT) (Tversky and Kahneman 1992, Chateauneuf and Wakker 1999).

Contrary to EU, RDEU and CPT are able to model that most decision makers do not treat probabilities linearly. Probability distortion enters into the respective evaluation functional in the form of the probability weighting function. Together with its counterpart for decision under uncertainty (Schmeidler 1989), Wakker (2004) considers “this development the main step forward for decision under incomplete information of the last decades”. Under EU, the only subjective component of the evaluation functional is attitude towards outcomes, captured by the utility function. Under RDEU and CPT, the probability weighting function constitutes another subjective component that reflects individual attitude towards probabilities, also called probabilistic risk attitude (Wakker 1994).

Theoretical studies, particularly in the framework of RDEU, often involve joint conditions on the utility function and the probability weighting function, for instance in characterizing risk aversion. Empirical studies that estimate (components of) the RDEU / CPT evaluation functional at an individual level usually report summary statistics of the parameters in question (e.g., Tversky and Kahneman 1992) or display the distribution of parameter estimates, separately for each parameter (e.g., Abdellaoui 2000). The issue of interrelation between the different components of the RDEU / CPT evaluation functional has not received much attention yet. This is in contrast to the issue's practical relevance. As will be shown in detail later, the sheer variation of the degree of interrelation between components of preference functionals can alter the risk behavior of decision makers, holding all other relevant characteristics constant

More specifically, we are interested in the interrelation between the curvature of the utility function and the elevation of the probability weighting function. This naturally leads to the topic of a suitable methodology allowing for an unbiased assessment of the quantity of interest. We therefore adopt a preference elicitation procedure based on the trade-off method (Wakker and Deneffe 1996), which is both theoretically appealing (Wakker and Deneffe 1996, p. 1145f.) and has been used in several recent empirical studies (e.g., Abdellaoui 2000, Bleichrodt and Pinto 2000, Etchart-Vincent 2004). We will investigate through an extensive simulation study whether this methodology is able to produce unbiased estimates of the degree of interrelation in the presence of decision error.[1]

Simulation presents itself as a suitable methodology because it permits to fix the quantity of interest, generate "quasi-experimental" data, obtain the measure of the quantity of interest and finally contrast the latter value with the true value.

The paper is structured as follows. Section 1 briefly introduces the decision-theoretical framework. Section 2 first offers additional motivation for our primary research objective and then presents an overview of the related literature. Section 3 covers all details of the specification of the simulation, including a description of the two-stage elicitation procedure. The results of the simulation are given in Section 4. Section 5 summarizes and discusses our findings.

1. Decision-Theoretical Framework

Let X be a set of monetary *outcomes*. Outcomes are expressed as changes with respect to the status quo, i.e. gains or losses. In decision under risk, the objects of choice are probability distributions over X called *prospects*. In the present paper, we restrict attention to binary prospects $(x_1, p; x_2)$ yielding a monetary outcome x_1 with probability p and a monetary outcome x_2 with probability $1 - p$. A prospect that involves both a gain and a loss outcome is called *mixed*. Other prospects are called *non-mixed*.

If the prospect $(x_1, p; x_2)$ is evaluated according to RDEU, we assume without loss of generality that it is (re-) arranged in the direction of decreasing preference, i.e. $x_1 \geq x_2$. The RDEU value is then given by

$$\pi_1 \cdot u(x_1) + \pi_2 \cdot u(x_2) \tag{1.}$$

where π_1 and π_2 are *decision weights* depending on the ranking of outcomes and $u(.)$ is a strictly increasing *utility function* from X to $\Re$. The decision weights π_1 and π_2 are defined by

$$\pi_1 = w(p) \text{ and } \pi_2 = 1 - w(p) \;, \tag{2.}$$

where $w(.)$ is a *probability weighting function*, i.e. a strictly increasing function from [0, 1] to [0, 1] satisfying $w(0) = 0$ and $w(1) = 1$. In the absence of probability distortion, i.e. $w(p) = p$ for all $p \in (0, 1)$, RDEU reduces to EU.

Under CPT, the utility function satisfies $u(0) = 0$ and decision weights are determined through two probability weighting functions: $w^+(.)$ for gains and $w^-(.)$ for losses. The CPT value of the prospect $(x_1, p; x_2)$ is still given by Equation (1). If it involves only gains [losses] with $x_1 \geq x_2 \geq 0$ [$x_1 \leq x_2 \leq 0$], $w(.)$ is replaced by $w^+(.)$ [$w^-(.)$] in the decision weights π_1 and π_2 defined by

Equations (2). For mixed prospects with $x_1 > 0 > x_2$, decision weights are defined by $\pi_1 = w^+(p)$ and $\pi_2 = w^-(1-p)$. If the utility function satisfies $u(0) = 0$ and the *duality condition* $w^-(p) = 1 - w^+(1-p)$ holds for all $p \in (0, 1)$, CPT and RDEU coincide.

2. Motivation and Related Literature

2.1. Motivation

In EU theory, an individual's attitude towards risk is fully captured by the curvature of the utility function. As pointed out before, RDEU and CPT are generalizations of EU that allow for distortions of probabilities as reflected by the probability weighting function in the representation of the preference relation. As a consequence, an individual's preference over risky alternatives is determined jointly by the utility function and the probability weighting function. Wakker (1994, p. 6) states that "[t]he popularity of RD[E]U is probably explained because it is the first well-developed and axiomatized theory to permit a separate attitude towards marginal utility and probabilistic risk".

The separation of risk attitude into two components proves to be fruitful in empirical research. Some real-world economic phenomena like the purchase of full insurance coverage in the presence of positive marginal loading can hardly be accommodated under EU, but are readily explained in an RDEU framework (Mossin 1968, Segal and Spivak 1990). Moreover, numerous experimental studies suggest that RDEU / CPT are descriptively superior to EU (see e.g. Wu et al. 2004 for an overview).

To further motivate the primary research issue underlying the present paper, we offer a simple example illustrating the impact of different forms of the utility function and the probability weighting function on the evaluation of risky alternatives under RDEU / CPT. For notational simplicity, we restrict ourselves to the RDEU case. Consider the prospect l = (100, 0.5; 0). Assume that decision maker A has EU preferences with utility function $u_1(x) = x^{1/2}$ or, equivalently, RDEU preferences with probability weighting function $w_0(p) = p$ and utility function $u_1(x) = x^{1/2}$. It is easily verified that her certainty equivalent for prospect l amounts to $CE^A(l) = 25$. Alternatively, assume that decision maker B has RDEU preferences with probability weighting function $w_1(p) = p^2$ and utility function $u_0(x) = x$. Analogous computations yield $CE^B(l) = 25$. Her certainty equivalent is therefore smaller than the expected value of the prospect. This is noteworthy because her utility function is linear, which under EU corresponds with risk neutrality and implies equality between the certainty equivalent and the expected value of the prospect.

The result is due to the shape of her probability weighting function which reflects "probabilistic risk aversion". It can be seen that $w_1(p) < p$ for all $p \in (0, 1)$. This inequality implies that the preferred outcome of a binary prospect receives a decision weight smaller than its objective probability. In RDEU, a relative dislike of risky prospects can be modeled either through the utility function or the probability weighting function.

Let (w_i, u_j) denote a decision maker with probability weighting function $w_i(.)$ and utility function $u_j(.)$, where i, j $\in$ {0, 1}. The missing permutations are (w_0, u_0), corresponding to an expected value maximizer with $CE(l) = 50$, and (w_1, u_1), representing a highly risk averse person with $CE(l) = 6.25$.

We have now prepared the ground for illustrating the economic relevance of our primary research question. Assume a population of RDEU decision makers of size N, among which N/2 have utility function $u_0(.)$ [$u_1(.)$] and likewise N/2 have probability weighting function $w_0(.)$ [$w_1(.)$].[2] Given the parametric families $u(x) = x^{\alpha}$, $\alpha > 0$, and $w(p) = p^{\beta}$, $\beta > 0$, this information amounts to the marginal probability distributions over the preference parameters α and β respectively. This information is compatible with quite different scenarios: Each of the combinations (w_0, u_0), (w_0, u_1), (w_1, u_0), and (w_1, u_1) might be equally likely, N/2 of the decision makers might be characterized by (w_0, u_0) and (w_1, u_1) respectively, etc. The distinction between such scenarios is economically meaningful. In the framework of the simple example above, think of deriving the supply schedule of risky claims if each of the N decision makers is endowed with one unit of prospect *l*. As individual willingness-to-accept is given by the respective certainty equivalent computed above, it can be seen that the resulting supply schedule depends upon the proportion of the four (w_i, u_j)-combinations. Put more generally, our interest is in the empirical covariation between risk attitude as captured by the curvature of the utility function and risk attitude as implied by the shape of the probability weighting function.

2.2. Related Theoretical Literature

The interrelation between utility function and probability weighting function is of considerable importance also from a theoretical point of view, in particular for characterizing risk aversion under the various theories of risky decision making. A decision maker is called *risk averse* if she prefers the sure receipt of the expected value of a non-degenerate lottery over the lottery itself. It is well known that under EU this condition is equivalent to concavity of the utility function. With the advent of more general preference models, a distinction between different forms of risk aversion has proved useful. In this framework, a decision maker with the above property is called *weakly risk averse*. In contrast, a decision maker is called *strongly risk averse* if she dislikes mean-preserving spreads, i.e. – loosely spoken – shifts of probability mass from the center of the distribution of outcomes to its tails holding the expected value constant (Rothschild and Stiglitz 1970). Under EU, this property corresponds to concavity of the utility function. Thus both weak and strong risk aversion lead to the same characterization of the utility function under EU. Under the more general preference models, this conclusion is no longer true.

Chew et al. (1987) study the general RDEU model, imposing only a number of "technical" conditions, above all a differentiability requirement. They show that strong risk aversion is satisfied if and only if the utility function is concave and the probability weighting function is convex. Schmidt and Zank (2002) investigate whether this result carries over to CPT. If only gain [loss] prospects are considered, Schmidt and Zank obtain the same conditions for strong risk aversion as Chew et al. (1987), i.e. concavity of the utility function in the gain [loss] domain and convexity of the probability weighting function for gains [losses].[3] Interestingly, if mixed prospects are considered, strong risk aversion does not imply concavity of the utility function over the whole domain. More specifically, non-concavity at the status quo is permitted.

Contrary to the case of EU, weak risk aversion and strong risk aversion lead to different characterizations of the components of the preference functional. If concavity of the utility function is assumed, a sufficient condition for weak risk aversion is $w(p) \leq p$ for all $p \in (0, 1)$ (see e.g. Segal 1987), which is termed "pessimism" in Quiggin (1993). Chateauneuf and Cohen (1994) extend this analysis by dispensing with the requirement of a concave utility function. They demonstrate that weak risk aversion and a convex utility function can coexist, provided that the probability weighting function is "sufficiently pessimistic". As Quiggin (1993, p. 79) remarks, this requirement cannot be met by some families of utility functions, notably the power family largely employed in the literature.

2.3. Related Empirical Literature

The interrelation between elements of preference functionals in individual decision making has also been the object of study in empirical research. While we are not aware of any study sharing the focus of the present paper, several are directed at related topics and are therefore relevant for us with view to both methodology and results. We shall briefly review these below.

Anderhub et al. (2001) conducted an experiment to assess the interaction between subjects' attitude towards risk and their time preference, which jointly govern for instance an individual's optimal choice in an intertemporal consumption allocation problem under conditions of risk. In their experiment, participants had to state their valuations for three risky alternatives that differed in the date at which the (risky) payment was due. From these three valuations, Anderhub et al. derive measures of the discount factor and of the degree of risk aversion. They find a significant negative correlation between the degree of risk aversion and the discount factor, i.e. more risk averse subjects tend to discount future utility more heavily.

Anderhub et al. (2001, p. 246) point out a caveat concerning their methodology, however, in which the relevant measures are not derived from a formal theory of choice. They argue that a more rigorous approach might consist of adopting some parametric family of utility functions, estimating the respective parameter at an individual level and using these estimates in computing correlations instead. As will be shown in detail in Section 3, the approach investigated in the present paper is exactly in this vein.

Cohen et al. (1985) present a large-sample study designed to characterize individual decision making under risk and under uncertainty. They obtained participants' certainty equivalents for ten different binary prospects, comprising four prospects with objective probabilities and one with unknown (Ellsberg (1961)-type) probabilities, for both gain and loss consequences. This set-up is comparable to a factorial design in which two factors with two “levels” each (known vs. unknown probabilities, gains vs. losses) yield four distinct conditions.

One of the issues of the research agenda of Cohen et al. is to find out whether subjects’ behavior is related across these four conditions. Remarkably, this does not seem to be the case. Neither are certainty equivalents nor is ambiguity attitude[4] significantly correlated across domains. Moreover, Cohen et al. conclude that risk attitude and ambiguity attitude are not related in an economically meaningful way, for both gain and loss domains. In interpreting these results, it should be kept in mind that what enters into the calculation of correlation coefficients are certainty equivalents or differences thereof, in a sense “raw data”. Possibly due to the early stage of the development of formal descriptive theories for behavior under risk and uncertainty, Cohen et al. did not make the attempt to fit a model to their data and reconsider the issue of correlation from this perspective.

Finally, a paper which only marginally touches on the issue of interrelation between preference parameters, but which extensively covers the issue of individual estimation of utility function and probability weighting function in an RDEU / CPT framework, is Gonzalez and Wu (1999). In order to be able to reliably assess preferences at an individual level, Gonzalez and Wu employed what they call “a traditional psychophysical paradigm” (relatively few subjects, many trials per subject). They obtained participants’ certainty equivalents for 165 different gain prospects. Based on these data, they employ both non-parametric and parametric techniques to estimate the utility function and the probability weighting function, individually for each subject.

Among other things, Gonzalez and Wu find convergence between the two different techniques and demonstrate that heterogeneity in preferences necessitates choice of a functional form for the probability weighting function that allows for an independent variation of elevation and curvature. Gonzalez and Wu state that "[t]hese properties [i.e., elevation and curvature] were observed to be somewhat independent across the 10 participants" (p. 159), which might be interpreted as a non-formal statement concerning the interrelation between the preference parameters in question. The authors also offer a psychological interpretation of elevation and curvature of the probability weighting function as attractiveness and discriminability respectively. In the discussion section of their paper, Gonzalez and Wu hypothesize that these concepts can also vary intrapersonally in decision making under uncertainty, i.e. differ for different sources of uncertainty. They argue that "even though discriminability and attractiveness are logically independent …, in the real world the two concepts most likely covary across contexts" (p. 161).

The above citation can serve as point of departure for the present investigation. Though we are interested in interpersonal instead of intrapersonal comparison and our main focus is on the interrelation between curvature of the utility function and elevation of the probability weighting function instead of the interrelation between elevation and curvature of the probability weighting function, the idea is essentially the same: How do particular components of a preference functional covary "in the real world"?

3. Specification of the Simulation

3.1. Design of the Experiment

Our approach for assessing the components of the RDEU / CPT preference functional will be a two-step procedure. The first step, which is based on the trade-off method (Wakker and Deneffe 1996), elicits the utility function by determining a standard sequence of outcomes, i.e. a sequence of outcomes equally spaced in utility units. In the second step, the probability weighting function is elicited using the utility values obtained in the first step as inputs. Our experimental design approximates to the one chosen by Abdellaoui (2000) and Bleichrodt and Pinto (2000), as will be explained in detail below.

Elicitation of the utility function. A standard sequence of (gain) outcomes is obtained as follows. Let $0 \leq r < R < x_0$ denote three fixed outcomes and $p \in (0, 1)$ a fixed probability.[5] As a first step, the outcome x_1 is determined such that the decision maker is indifferent between the prospects $(x_0, p; R)$ and $(x_1, p; r)$. As a second step, the decision maker is called to state the outcome x_2 such that indifference between the prospects $(x_1, p; R)$ and $(x_2, p; r)$ holds. Assuming that CPT is an adequate descriptive theory of choice, the combination of the equations resulting from the above two indifference statements implies the equality of $u(x_2) - u(x_1)$ and $u(x_1) - u(x_0)$.

The next steps follow the general principle that once outcome x_i has been elicited, outcome x_{i+1} leading to indifference between $(x_i, p; R)$ and $(x_{i+1}, p; r)$ has to be determined. The elicitation procedure results in an increasing sequence of outcomes $x_0, x_1, ..., x_n$[6] such that

$$u(x_{i+1}) - u(x_i) = u(x_i) - u(x_{i-1}), \text{ i} = 1, \dots, \text{n} - 1. \tag{3.}$$

Given the uniqueness property of the utility function, we choose the convenient normalization $u(x_0) = 0$ and $u(x_n) = 1$.

Viewed in isolation, Equation (4) evokes a strength of preference interpretation. It should be observed, however, that "the elicited utilities have been derived solely from "ordinal" indifferences and are not susceptible to the methodological criticisms of approaches that take strength of preference as a directly observable primitive" (Wakker and Deneffe 1996, p. 1147).

Elicitation of the probability weighting function. The determination of the probability weighting function (for the gain domain), which builds upon the standard sequence of (gain) outcomes $x_0, x_1, \dots, x_n$, proceeds as follows. For each of the interior elements of the standard sequence x_j, j = 1, …, n – 1, the decision maker is asked to state the probability p_j such that she finds the prospect (x_n, p_j; 0) and the certain receipt of x_j equally attractive. Under CPT and the above normalization convention for u(.), this indifference statement translates into:

$$w(p_j) = u(x_j) = j/n, \; j = 1, \ldots, n-1. \tag{4.}$$

We will refer to $p_1, p_2, \ldots, p_{n-1}$ as the standard sequence of probabilities.

Our methodology to estimate the probability weighting function coincides with that in Abdellaoui (2000). An alternative approach consists of asking the decision maker to state the certainty equivalent for the prospect (x_n, q_k; 0), for a set of probabilities $\{q_k\}$. The adjustment of an *outcome* of an incomplete prospect such that indifference to a fully specified prospect holds constitutes the commonality between the latter approach and the methodology used by Bleichrodt and Pinto (2000). They thereby share the slight disadvantage that the utility value of the outcome stated by the participant – given the normal case that the stated outcome is not an element of the standard sequence of outcomes – is not immediately available but needs to be computed otherwise, for example by linear interpolation or parametric fitting. A potential advantage of the method as suggested by Bleichrodt and Pinto (2000), p. 1489, plays no role in the present framework, which induced us to adopt the approach of Abdellaoui (2000).

3.2. Generation of Hypothetical Preference Statements

The white noise model in general. Whereas it is widely accepted that RDEU and above all CPT are descriptively superior to EU, it must be recognized that regularly even the generalized theories do not exactly fit to a subject's choice behavior. The common way to deal with this problem is to assume that subjects make their decisions with error.[7] Three different specifications of the error component have received most attention: the white noise model (Hey and Orme

1994), the constant error probability model (Harless and Camerer 1994), and the random preference model (Loomes and Sugden 1995), see Loomes et al. (2002) for details. Our simulations will be based on the white noise model for two reasons. First, it has received favorable results in empirical assessment (Carbone 1997). Second, it satisfies the criterion of parsimony by introducing only one additional parameter.

If the stochastic component is taken into account, a subject's actual evaluation of a prospect l, henceforth denoted by $V(l)$, is given by the equation $V(l) = V^{core}(l) + \varepsilon_l$, where $V^{core}(l)$ denotes the evaluation of the prospect according to the core theory (sometimes also paraphrased as "true" evaluation) and ε_l denotes the associated error term. In the present paper, the core theory is CPT, i.e. $V^{core}(l) = CPT(l)$. The error term is assumed to have an expected value of zero and a variance of σ_ε^2. If more than one prospect is evaluated, the respective error terms are assumed to be independently distributed. Additional details about the distributional assumption for ε_l will be spelled out below.

In other simulation studies of decision making under risk, the hypothetical subject either faces a series of pairwise choice situations or is asked to rank a given set of prospects according to her preference (Carbone and Hey 1994). Our experimental design differs in that the hypothetical subject's task consists of adjusting a value (outcome or probability) of an incomplete prospects so as to render it equally preferable to a fully specified prospect.[8] To incorporate the stochastic element, we let the evaluation of the fully specified prospect be disturbed by the error term ε_l.

Applying the white noise model. Applied to the utility function tasks, the outcome value provided by the subject in the i-th step (i = 1, …, n) of the elicitation process is therefore given by:

$$x_i = u^{-1}\left[\frac{w(p)\cdot u(x_{i-1}) + (1-w(p))\cdot(u(R)-u(r)) + \varepsilon_l}{w(p)}\right]. \qquad (5.)$$

We have now prepared ground for the presentation of the distributional assumption for ε_l. One natural assumption to make is that ε_l is normally distributed, $\varepsilon_l \sim N(0, \sigma_\varepsilon^2)$. The implied unboundedness of the support of ε_l would, however, lead to serious problems. First, it could not be ensured that the argument of $u^{-1}(.)$ in Equation (6) belongs to the domain of the inverse function at all. Second, it could not be ensured that the preference statements obey first-order stochastic dominance (FOSD). FOSD is generally regarded as a fundamental rationality criterion which, moreover, is rarely violated in practice (Carbone and Hey 1995, Carbone 1997, Loomes and Sugden 1998, Levy and Levy 2001) – at least if the relation is not particularly intransparent (Birnbaum and Navarrete 1998).

In a utility function task, FOSD is satisfied if and only if $x_i > x_{i-1}$, which is equivalent to the following restriction on the error term ε_l:

$$\varepsilon_l > -(1-w(p))\cdot(u(R)-u(r)). \qquad (6.)$$

To incorporate this restriction, we resort to the truncated normal distribution. To preserve the property of zero expected value of ε_l, we eventually assume that ε_l follows a doubly truncated normal distribution – derived from the normal distribution presented above – with lower [upper] truncation point $t_u^- = -(1 - w(p)) \cdot (u(R) - u(r))$ [$t_u^+ = (1 - w(p)) \cdot (u(R) - u(r))$].

The probability weighting function tasks are treated analogously. The probability value provided by the subject in the j-th step (j = 1, …, n – 1) of the elicitation process is therefore given by:

$$p_j = w^{-1}\left[\frac{u(x_j)-u(x_0)+\varepsilon_l}{u(x_n)-u(x_0)}\right]. \qquad (7.)$$

FOSD is satisfied if and only if $0 < p_j < 1$, which is equivalent to the following restriction on the error term ε_l:

$$-(u(x_j)-u(x_0)) < \varepsilon_l < u(x_n)-u(x_j)\,. \qquad (8.)$$

Consequently, we assume that ε_l follows a doubly truncated normal distribution with lower [upper] truncation point $t_w^- = -\min\{u(x_j) - u(x_0);\ u(x_n) - u(x_j)\}$ [$t_w^+ = \min\{u(x_j) - u(x_0);\ u(x_n) - u(x_j)\}$].

One distinction between the utility function tasks and the probability weighting function tasks deserves special mention. Whereas imposing satisfaction of the requirement of FOSD for each of the utility function tasks is sufficient to guarantee that the standard sequence of outcomes $x_0, x_1, \ldots, x_n$ is increasing, the analogous claim for the standard sequence of probabilities $p_1, p_2, \ldots, p_{n-1}$ cannot be established. This is due to the fact that the elicitation of $p_1, p_2, \ldots, p_{n-1}$ is not "chained", i.e. p_j does not enter into the assessment of p_{j+1}. The requirement of $p_1 < p_2 < \ldots < p_{n-1}$ could easily be incorporated into the present framework by strengthening Inequality (9). We refrain from doing so because empirical evidence suggests that violations of the requirement of FOSD "across decision situations" are not comparably infrequent as those "within decision situations" (see the results in Gonzalez and Wu 1999, p. 145).

3.3. Specification of the Parameters

Choice of parametric families. The core theory in the present paper is CPT. As all decision situations involve gain prospects, we only need to specify the utility function over the gain domain and the probability weighting function (for gains). We use parametric families commonly found in the literature. For the utility function, we choose the power family

$$u(x) = x^{\alpha},\ x \geq 0, \tag{9.}$$

where $\alpha > 0$ can be interpreted as an anti-index of risk aversion (as captured by the utility function). For the probability weighting function, we choose the linear-in-log-odds family

$$w(p) = \frac{\delta \cdot p^{\gamma}}{\delta \cdot p^{\gamma} + (1-p)^{\gamma}} \tag{10.}$$

previously applied by for instance Goldstein and Einhorn (1987), Lattimore et al. (1992), Gonzalez and Wu (1999) and Abdellaoui et al. (2005). This parametric family is especially suited for our purposes because it permits an independent variation of and thus distinction between elevation and curvature. The parameter $\delta > 0$ mainly controls elevation, the parameter $\gamma > 0$ mainly controls curvature.

Curvature reflects the sensitivity of the decision maker with respect to changes in probability inside the interval (0, 1). Elevation can roughly be defined as the average level of transformed probabilities. Under a highly elevated probability weighting function, decision weights of the more favorable outcomes tend to exceed the corresponding objective probabilities and vice versa, which establishes the link between the degree of elevation and the concept of probabilistic risk aversion. By virtue of this relationship, δ can be interpreted as an anti-index of risk aversion (as captured by the probability weighting function).

Incorporating the correlation structure. Heterogeneity of preferences can be introduced by letting the parameter values vary across participants. To operationalize heterogeneity of preferences in the simulation framework, (multivariate) distributions are specified from which the (vectors of) preference parameters are drawn, individually for each hypothetical subject. Our focus being on the manifestation of risk aversion in the form of either the curvature of the utility function or the elevation of the probability weighting function, we vary the interrelation between the parameter α of *u*(.) and the parameter δ of *w*(.). Three special cases are considered: $\rho_{\alpha,\delta} = +1$, stochastic independence between α and δ, and $\rho_{\alpha,\delta} = -1$.[9] The parameter γ of *w*(.) is taken to be independently distributed from α and δ throughout.

To complete the description, the marginal distributions of the three preference parameters have to be specified. Carbone and Hey (1994) adequately argue that the choice of parameters in any Monte Carlo study is rather arbitrary. In an attempt to reduce arbitrariness as much as possible, our specification draws upon the empirical results obtained by Abdellaoui (2000). The marginal distributions are calibrated such that their respective median value coincides with the corresponding parameter estimate for median data in Abdellaoui (2000) (p. 1506 for *u*(.), Table 9 on p. 1509 for *w*(.)) and their respective variance is of approximately equal size as the dispersion inherent in the distribution of individual parameter estimates in Abdellaoui (2000) (Figure 6 on p. 1511). Two different families of distribution functions – uniform and triangular – are employed, primarily as a robustness check. Combined with the three different assumptions about the correlation structure, we arrive at six different scenarios summarized in Table 1 below.

[Insert Table 1 about here]

Setting the variance of the error term. It has already been pointed out that the variance of the error term is crucially important, yet at the same time difficult to determine in an empirically meaningful way. To deal with this problem, the simulations are conducted for a wide range of values for σ_ε^2.[10] This is not to be understood as meaning that we regard all values considered subsequently as empirically plausible but rather that we seek to cover the “true” value of the variance with a high degree of confidence. The magnitude of the error variance is the principal determinant of the degree of precision with which the preference parameters can be estimated. A measure of the degree of precision in dependence of σ_ε^2 will be presented in Subsection 4.2, thereby linking the magnitude of σ_ε^2 to a quantity which is accessible to interpretation more easily.

It is important to observe that σ_ε is expressed in units of utility. The utility function, however, is unique up to positive affine transformation, which implies that the magnitude of σ_ε must be viewed against the background of the normalization adopted for *u*(.). For this reason, a straightforward comparison of the magnitude of σ_ε with other simulation studies like Carbone and Hey (1994) or Carbone (1997) is not feasible.

4. Results

The present section is structured as follows. First, we briefly describe how the parameters of the preference functional are estimated from the preference statements of the hypothetical subjects. We then study the impact of the magnitude of σ_ε^2 on the precision of the estimation. Subsequently, we turn to our main question, which is whether the correlation coefficient between estimates of preference parameters permits to assess its population counterpart in a distortion-free way. In this analysis, we distinguish between different sizes of the population of hypothetical subjects. We assume both N = 41, which exactly equals the number of participants in Abdellaoui et al. (2005) and should be fairly typical of related experimental investigations, and N = 2000, which is less a plausible sample size than the attempt to study the limit behavior of the procedure. Besides the number of hypothetical subjects, we distinguish between the different scenarios under which the preference parameters are determined as summarized in Table 1.

4.1. Estimation

Consistent with the two-step procedure of the experimental set-up, the parameter of the utility function is estimated separately from the parameters of the probability weighting function. The estimation of the utility function parameter is based upon a hypothetical subject's standard sequence of outcomes $x_0, x_1, ..., x_n$. Under the chosen normalization $u(x_0) = 0$ and $u(x_n) = 1$, it follows that $u(x_i) = i/n$, i = 1, …, n – 1. As $u(.)$ belongs to the power family, the normalization leads to the following expression:

$$u(x) = \frac{x^{\alpha} - x_0^{\alpha}}{x_n^{\alpha} - x_0^{\alpha}}, x \geq 0 \ . \qquad (11.)$$

The parameter α is estimated by nonlinear least squares regression with the outcome values x_i as independent variable and their associated utility values i/n as dependent variable. Clearly, the parameter estimate $\hat{\alpha}$ is invariant with respect to changes of the normalization of the utility function.

The estimation of the parameters of the probability weighting function is based upon a hypothetical subject's standard sequence of probabilities $p_1, p_2, \ldots, p_{n-1}$. As shown in Subsection 3.1, its elements satisfy $w(p_j) = u(x_j) = j/n$. The parameters δ and γ of the linear-in-log-odds function are estimated by nonlinear least squares regression with the probabilities p_j as independent variable and their respective decision weights j/n as dependent variable. It is noteworthy that the parameter estimate $\hat{\alpha}$ does not enter into the estimation of the probability weighting function.

4.2. Precision

The magnitude of σ_ε^2 is essential to the precision of the estimation. To highlight this fact, consider first the limiting case $\sigma_\varepsilon^2 = 0$. Three (appropriately selected) choice tasks would suffice to determine the three parameters of the preference functional without any estimation error, i.e. $\hat{\alpha} = \alpha$ and likewise for $\hat{\delta}$ and $\hat{\gamma}$. In the opposite case, $\sigma_\varepsilon^2 \to \infty$, a subject's actual evaluation of a prospect is almost exclusively governed by the error term, which in turn means that the informational content of a subject's preference statements with respect to inferences about the underlying preference functional vanishes.

The main purpose of this subsection is to establish a connection between the magnitude of σ_ε^2 and a measure of the precision of the estimation. In anticipation of the results to be displayed below, we also seek to demonstrate that we have covered a sufficiently wide range of possible values for σ_ε^2.

The measure of precision of the estimation is operationalized as the mean absolute value of the difference between the true parameter value and the corresponding estimate, i.e. $|\hat{\alpha} - \alpha|$ and likewise for $\hat{\delta}$ and $\hat{\gamma}$, where the mean is taken over all hypothetical subjects belonging to Scenario I. The results are displayed in Table 2. The respective figures for the other scenarios are highly similar and are omitted for lack of space.

[Insert Table 2 about here]

Firstly, Table 2 reflects the rather obvious fact that the (measure of) precision of the estimation deteriorates upon increasing the magnitude of σ_ε. The mean absolute deviation of the (individual) parameter estimate $\hat{\alpha}$ from its respective true value α, for example, increases monotonically from zero ($\sigma_\varepsilon = 0$) to 0.211 ($\sigma_\varepsilon = 10$). Due to the widespread use of the power family in utility measurement, the absolute level of the measure of precision for $\hat{\alpha}$ probably lends itself to interpretation most readily. In our view, the figures in Table 2 indicate that the range of possible

σ_ε values has been chosen sufficiently wide to embrace the large majority of potential decision makers.

In addition, it can be seen that the increase in the deviation measure induced by a given increase in the magnitude of σ_ε becomes smaller for higher values of σ_ε. In interpreting this fact, however, it should be kept in mind that σ_ε does not equal the standard deviation of the error term owing to the truncation of its distribution (see Footnote 9). For given truncation points, the standard deviation of ε_l increases in σ_ε, yet at a slower rate than σ_ε itself. Although in the present framework the situation is further complicated by the fact that the truncation points for ε_l may differ across choice situations (see Subsection 3.2), the latter relationship is able to explain the above property of the deviation measure.

Finally, we focus on the two parameters of the probability weighting function. It is interesting to note that the deviation measure is smaller for the parameter γ than for the parameter δ throughout. This result suggests that inferences about the curvature of the probability weighting function can be made with higher precision as compared to its elevation. One possible objection to this conclusion is that the marginal distribution functions from which the true parameter values are determined differ. To rule out this explanation, additional simulations (not reported in detail in the present paper) were conducted. They show that the above finding remains valid when the marginal distribution of γ is equated with that of δ or vice versa.

4.3. Correlation: Standard Sample Size

We now turn to our main question, which is whether the correlation coefficient between estimates of preference parameters permits to assess its population counterpart in a distortion-free way. This question is first analyzed assuming that the size of the population of hypothetical subjects amounts to N = 41. To limit the dependence of the results to be obtained on the particular set of error terms generated, forty independent runs (of 41 hypothetical subjects each) were conducted. Unless otherwise noted, the following figures represent mean values of the respective quantities, where the mean is taken over the forty runs.

Under Scenario I, all three correlation coefficients computed on the basis of the true parameter values are close to zero due to the assumption of stochastic independence: $\rho(\alpha,\delta)$ = -0.012, $\rho(\alpha,\gamma)$ = 0.001, and $\rho(\delta,\gamma)$ = -0.022.[11] These figures can also be extracted from Table 3, second column (σ_ε = 0), as the absence of estimation error – i.e. $\hat{\alpha} = \alpha$ and likewise for $\hat{\delta}$ and $\hat{\gamma}$ – permits to assess the correlation between the parameters in the underlying population without error – i.e. $\rho(\hat{\alpha},\hat{\delta})$ = $\rho(\alpha,\delta)$ and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$. The last-mentioned relationship should be kept in mind when interpreting subsequent tables.

[Insert Table 3 about here]

For $\sigma_\varepsilon > 0$, the (mean) correlation coefficient between estimates of preference parameters steadily departs from the correlation coefficient between the parameters in the underlying population. This effect is particularly pronounced for $\rho(\hat{\alpha},\hat{\delta})$ with $\rho(\hat{\alpha},\hat{\delta})$ = -0.420 for σ_ε = 10. Whereas the parameters α and δ are virtually uncorrelated in the underlying population, $\rho(\hat{\alpha},\hat{\delta})$ misleadingly suggests a strong negative relationship between the parameters, which would mean that a higher degree of risk aversion incorporated in the utility function tends to be associated with a lower degree of risk aversion incorporated in the probability weighting function and vice versa.

[Insert Figure 1 about here]

The downward distortion is also found if the data are analyzed at the disaggregated level. Figure 1 comprises two scatterplots in which $\rho(\hat{\alpha},\hat{\delta})$ is plotted against $\rho(\alpha,\delta)$, exemplarily for $\sigma_\varepsilon = 3$ (Figure 1 left) and $\sigma_\varepsilon = 10$ (Figure 1 right), where each point corresponds to one of the 40 runs. As expected, $\rho(\alpha,\delta)$ is distributed quite evenly around the zero line, which stands in sharp contrast to $\rho(\hat{\alpha},\hat{\delta})$. The inequality $\rho(\hat{\alpha},\hat{\delta}) < \rho(\alpha,\delta)$ holds for 39 ($\sigma_\varepsilon = 3$) and 40 ($\sigma_\varepsilon = 10$) out of the 40 cases respectively. Summing up, the simulation results strongly favor the conclusion that the method under study is not capable of assessing the correlation coefficient between preference parameters in the underlying population in a distortion-free way.

To complete the picture, we consider two more special cases in Scenarios II and III: perfect positive and perfect negative correlation between α and δ. Given the boundedness of the correlation coefficient and the fact that the two scenarios just represent the two boundary points, it immediately follows that the correlation coefficient between estimates of preference parameters cannot be an unbiased estimator in the presence of decision error. The purpose of the simulations rather lies in a quantification of the degree of bias, particularly in order to enable a comparison across scenarios.

[Insert Tables 4 and 5 about here]

The results are displayed in Table 4 ($\rho(\alpha,\delta)$ = 1) and Table 5 ($\rho(\alpha,\delta)$ = -1). As expected, $\rho(\hat{\alpha},\hat{\delta})$ steadily departs from the respective boundary point of the range of possible correlation coefficient values under both scenarios. In Scenario II [III], $\rho(\hat{\alpha},\hat{\delta})$ decreases [increases] monotonically from 1 [-1] until it equals 0.192 [-0.796] for $\sigma_\varepsilon = 10$. It is interesting to note, however, that the margin by which $\rho(\hat{\alpha},\hat{\delta})$ deviates from the respective boundary point is markedly larger for $\rho(\alpha,\delta)$ = 1 than for $\rho(\alpha,\delta)$ = -1. This result is similar in spirit to the downward bias documented for the case of stochastic independence.

Finally, we treat the question of whether the results presented above prove to be stable with respect to changes of the distribution from which the true vector of preference parameters is drawn. In Scenarios IV-VI, the uniform distribution is replaced with the symmetric triangular distribution. It should be noted that the support of the marginal distributions has been extended to keep variance constant.

The results are displayed in Tables 6-8, which have been relegated to the Appendix. From an overall perspective, they are remarkably similar to their counterparts in Tables 3-5. In the stochastic independence case (Scenario IV), $\rho(\hat{\alpha},\hat{\delta})$ is strongly biased downward with $\rho(\hat{\alpha},\hat{\delta})$ = -0.441 for σ_ε = 10. At the disaggregated level, the inequality $\rho(\hat{\alpha},\hat{\delta}) < \rho(\alpha,\delta)$ holds for 39 (σ_ε = 3) and 40 (σ_ε = 10) out of the 40 cases respectively. Under Scenario V [VI], $\rho(\hat{\alpha},\hat{\delta})$ decreases [increases] monotonically from 1 [-1] until it amounts to 0.168 [-0.820] for σ_ε = 10, which means that the deviation from the respective boundary point exhibits the same asymmetry already encountered in the comparison between Scenarios II and III. The main findings are therefore not affected by the change in the distributional assumption.

4.4. Correlation: Large Sample Size

The results reported in the previous subsection provide strong evidence against the applicability of the procedure under study to measure correlation between preference parameters in a distortion-free way. As our interest in the topic was sparked by the work of Abdellaoui et al. (2005), we set N = 41 in an attempt to produce results tailored to our empirical point of departure. Not surprisingly, given this sample size, analyses at the disaggregated level reveal a substantial degree of variation of results across simulation runs, as evident from Figure 1. The present subsection is therefore devoted to investigating the large-sample behavior of the procedure under study. We set N = 2000 for the (single) simulation run.

[Insert Tables 9-11 about here]

The results for Scenarios I-III are presented in Tables 9-11. For the sake of clarity, it should be stressed that the data underlying the present results were generated independently of the standard sample size data analyzed in Subsection 4.3. Nonetheless, the results are highly similar. The "limit" behavior of the procedure under study is thus virtually indistinguishable from the "mean" behavior under traditional sample sizes, provided that the mean is taken over a sufficiently large number of independent simulation runs. Given the proximity of the results, the conclusions drawn in the previous subsection are further confirmed. We refrain from restating them here to avoid repetitions. The same reasoning applies to Scenarios IV-VI. For reasons of completeness, we include the respective results as Tables 12-14 in the Appendix.

5. Discussion and Conclusion

The primary research objective of the present paper was to investigate the issue of interrelation / covariation between different components of the RDEU / CPT preference functional. Prior to obtaining insights in this direction, the suitability of the envisaged measure of covariation had to be verified, which induced us to conduct an extensive simulation study. The simulation results are unequivocal. In the presence of decision error, the proposed procedure does not allow to measure the quantities of interest in a distortion-free way. With the obvious exception of the lower end of the range of possible correlation coefficient values, we find a strong downward distortion.

In the course of our investigation, it has become apparent that decision error plays a central role. Stochastic preference models take explicit account of decision error by adding a stochastic component to a deterministic core theory. Whereas the choice of RDEU / CPT as the core theory seems easy to justify on the basis of its popularity and descriptive success, the debate about how to suitably incorporate randomness is not settled yet.

We opted for the white noise error model in view of its parsimony and its favorable empirical results. It must be mentioned, however, that even this error model is not free from criticism, for instance for its problems to predict the very low frequency of violations of FOSD commonly observed (Loomes and Sugden 1998). Furthermore, it is by no means clear that the white noise model is necessarily descriptively superior to its contenders, especially the random preference model (Loomes and Sugden 1995, Carbone 1997). More recently, Loomes et al. (2002) proposed hybrid error models obtained by adding a tremble à la Harless and Camerer (1994) to either the white noise model or random preference model.

Even if attention is restricted to the white noise error model, generalizations of our approach are conceivable. In our simulation, the error term ε_l is assumed to be homoscedastic, i.e. the variance of ε_l is assumed to be constant (across subjects and choice tasks).[12] Hey (1995) and Buschena and Zilberman (2000) show that allowing for heteroscedasticity can lead to a significant improvement in the quality of fit to experimental data. Our choice of a homoscedastic error term can be defended on at least two grounds. First, the heteroscedastic formulation that emerges as the best in Hey (1995) involves time taken to answer a particular choice question, which obviously has no analogue in a simulation framework. Second, Buschena and Zilberman (2000) find that the benefits from introducing a heteroscedastic specification are significantly reduced under a non-EU core theory like RDEU (as compared to EU).

Taken together, the issue of how to suitably model decision behavior for simulation purposes may be controversial. With regard to the research question addressed in the present paper via the simulation methodology, however, this controversy probably is of secondary importance. Although we have considered a multiplicity of different constellations of simulation parameters, the results are remarkably uniform. A further expansion of the scope of the simulation, such as incorporating the aspects of decision behavior discussed above, can therefore be expected to lead to a low incremental gain in insight only.

From a methodological point of view, we hope to have advanced simulation as a useful technique for tackling selected problems in decision research. The simulation methodology has been applied for instance by Bleichrodt and Pinto (2000) to study the impact of decision error on their experimental results, by Carbone and Hey (1994) to study the distinguishability between different preference functionals under risk, and by Jia et al. (1998) to study the influence of different attribute weighting schemes on the quality of decisions in a multiattribute decision setting. One possibly fruitful application of the simulation program developed for the present paper might consist of assessing the performance of alternative utility elicitation methods in the presence of decision error.

From a conceptual point of view, the results of the simulations are somewhat dissatisfactory. While we are confident to have provided a thorough examination of the suitability of the proposed method, the findings refuting its suitability necessarily leave open the more fundamental underlying research question: How do particular components of the RDEU / CPT preference functional covary "in the real world"? We continue to consider this a worthwhile topic and thus hope that future research will shed light on it.

Appendix

[Insert Tables 6-8 about here]

[Insert Tables 12-14 about here]

Notes

1 Our interest in this research question stems from additional computations on the basis of the data in Abdellaoui et al. (2005). There we find a highly pronounced correlation (across subjects) between the estimate of the utility function parameter and the estimate of the parameter governing elevation of the probability weighting function.

2 It is widely acknowledged that there are significant interpersonal differences in risk preferences (e.g., Tversky and Kahneman 1992, Gonzalez and Wu 1999).

3 Schmidt and Zank (2002) do not directly utilize the result obtained by Chew et al. (1987) as their assumptions imposed on the preference functional differ.

4 Ambiguity attitude is defined as “the difference between … behavior with respect to risk and … behavior with respect to uncertainty” (Cohen et al. 1985, Remark 2 on p. 217).

5 In the simulation, we set $r = 0$, $R = 50$, $x_0 = 100$, and $p = 0.5$.

6 In the simulation, we set n = 6 like in Abdellaoui (2000), Bleichrodt and Pinto (2000) and Abdellaoui et al. (2005).

7 A more detailed exposition, including a discussion of alternative explanations, can be found in Hey (1999).

8 Carbone and Hey (1994), p. 241, argue that “indifference question data … are usually rejected by economists on motivational grounds”. Notwithstanding objections to the “matching”-methodology (Bostic et al. 1990), it is worth observing that indifference data can also be obtained through a series of choice questions (like in Abdellaoui et al. 2005), which should invalidate the above argument.

9 It should be noted that stochastic independence between α and δ constitutes a property of the multivariate distribution from which the vectors of preference parameters for individual subjects are drawn. When $\rho_{\alpha,\delta}$ is computed for particular population of hypothetical subjects, it usually does not equal zero exactly.

10 One caveat about notation is in order. As explained before, the error term ε_l is assumed to follow a doubly truncated normal distribution with symmetric truncation points. In the following, σ_ε^2 will denote the variance of the normal distribution from which the truncated normal distribution is derived. The variance of the truncated normal distribution is smaller than σ_ε^2 for all finite truncation points; for details see Johnson et al. (1994).

11 The correlation coefficients do not equal zero exactly for reasons explained in Footnote 8.

12 Strictly speaking, the error term in our simulation does not satisfy homoscedasticity exactly. Due to differences in the truncation point (see Subsection 3.2) across choice tasks, the variances of the respective error terms are not exactly equal.

References

Abdellaoui, Mohammed. (2000). “Parameter-free elicitation of utility and probability weighting functions,” *Management Science* 46, 1497-1512.

Abdellaoui, Mohammed, Frank Vossmann, and Martin Weber. (2005). “Choice-based elicitation and decomposition of decision weights for gains and losses under uncertainty,” *Management Science* (forthcoming).

Allais, Maurice. (1953). “Le comportement de l’homme rationnel devant le risque: Critique des postulats et axiomes de l’école américaine,” *Econometrica* 21, 503-546.

Anderhub, Vital et al. (2001). “On the interaction of risk and time preferences: An experimental study,” *German Economic Review* 2, 239-253.

Birnbaum, Michael H. and Juan B. Navarrete. (1998). “Testing descriptive utility theories: Violations of stochastic dominance and cumulative independence,” *Journal of Risk and Uncertainty* 17, 49-78.

Bleichrodt, Han and Jose L. Pinto. (2000). “A parameter-free elicitation of the probability weighting function in medical decision analysis,” *Management Science* 46, 1485-1496.

Bostic, Raphael, Richard J. Herrnstein, and R. Duncan Luce. (1990). “The effect on the preference-reversal phenomenon of using choice indifferences,” *Journal of Economic Behavior and Organization* 13, 193-212.

Buschena, David and David Zilberman. (2000). "Generalized expected utility, heteroscedastic error, and path dependence in risky choice," *Journal of Risk and Uncertainty* 20, 67-88.

Carbone, Enrica. (1997). "Investigation of stochastic preference theory using experimental data," *Economics Letters* 57, 305-311.

Carbone, Enrica and John D. Hey. (1994). "Discriminating between preference functionals: A preliminary Monte Carlo study," *Journal of Risk and Uncertainty* 8, 223-242.

Carbone, Enrica and John D. Hey. (1995). "A comparison of the estimates of expected utility and non-expected-utility preference functionals," *Geneva Papers on Risk and Insurance Theory* 20, 111-133.

Chateauneuf, Alain and Michèle Cohen. (1994). "Risk seeking with diminishing marginal utility in a non-expected utility model," *Journal of Risk and Uncertainty* 9, 77-91.

Chateauneuf, Alain and Peter P. Wakker. (1999). "An axiomatization of cumulative prospect theory for decision under risk," *Journal of Risk and Uncertainty* 18, 137-145.

Chew, Soo Hong, Edi Karni, and Zvi Safra. (1987). "Risk aversion in the theory of expected utility with rank dependent probabilities," *Journal of Economic Theory* 42, 370-381.

Cohen, Michèle, Jean-Yves Jaffray, and Tanios Said. (1985). "Individual behavior under risk and under uncertainty: An experimental study," *Theory and Decision* 18, 203-228.

Ellsberg, Daniel. (1961). "Risk, ambiguity, and the Savage axioms," *Quarterly Journal of Economics* 75, 643-669.

Etchart-Vincent, Nathalie. (2004). “Is Probability Weighting Sensitive to the Magnitude of Consequences? An Experimental Investigation on Losses,” *Journal of Risk and Uncertainty* 28, 217-235.

Fox, Craig R. and Amos Tversky. (1998). “A belief-based account of decision under uncertainty,” *Management Science* 44, 879-895.

Goldstein, William M. and Hillel J. Einhorn. (1987). “Expression theory and the preference reversal phenomena,” *Psychological Review* 94, 236-254.

Gonzalez, Richard and George Wu. (1999). “On the shape of the probability weighting function,” *Cognitive Psychology* 38, 129-166.

Harless, David W. and Colin F. Camerer. (1994). “The predictive utility of generalized expected utility theories,” *Econometrica* 62, 1251-1289.

Hey, John D. (1995). “Experimental investigations of errors in decision making under risk,” *European Economic Review* 39, 633-640.

Hey, John D. (1999). “Estimating (risk) preference functionals using experimental methods.” In Luigi Luini (ed.), *Uncertain decisions: Bridging theory and experiments*. Boston: Kluwer.

Hey, John D. and Chris Orme. (1994). “Investigating generalizations of expected utility theory using experimental data,” *Econometrica* 62, 1291-1326.

Jia, Jianmin, Gregory W. Fischer, and James S. Dyer. (1998). “Attribute weighting methods and decision quality in the presence of response error: A simulation study,” *Journal of Behavioral Decision Making* 11, 85-105.

Johnson, Norman L., Samuel Kotz, and Narayanaswamy Balakrishnan. (1994). *Continuous univariate distributions*, 2nd ed. New York: Wiley.

Kilka, Michael and Martin Weber. (2001). "What determines the shape of the probability weighting function under uncertainty?," *Management Science* 47, 1712-1726.

Lattimore, Pamela K., Joanna R. Baker, and Ann D. Witte. (1992). "The influence of probability on risky choice: A parametric examination," *Journal of Economic Behavior and Organization* 17, 377-400.

Levy, Moshe and Haim Levy. (2001). "Testing for risk aversion: A stochastic dominance approach," *Economics Letters* 71, 233-240.

Loomes, Graham, Peter G. Moffatt, and Robert Sugden. (2002). "A microeconometric test of alternative stochastic theories of risky choice," *Journal of Risk and Uncertainty* 24, 103-130.

Loomes, Graham and Robert Sugden. (1995). "Incorporating a stochastic element into decision theories," *European Economic Review* 39, 641-648.

Loomes, Graham and Robert Sugden. (1998). "Testing different stochastic specifications of risky choice," *Economica* 65, 581-598.

Mossin, Jan. (1968). "Aspects of rational insurance purchasing," *Journal of Political Economy* 76, 553-568.

Quiggin, John. (1982). "A theory of anticipated utility," *Journal of Economic Behavior and Organization* 3, 323-343.

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

Rothschild, Michael and Joseph E. Stiglitz. (1970). "Increasing risk: I. A definition," *Journal of Economic Theory* 2, 225-243.

Schmeidler, David. (1989). "Subjective probability and expected utility without additivity," *Econometrica* 57, 571-587.

Schmidt, Ulrich and Horst Zank. (2002). "Risk aversion in cumulative prospect theory," Discussion Paper #206, University of Manchester.

Segal, Uzi. (1987). "The Ellsberg paradox and risk aversion: An anticipated utility approach," *International Economic Review* 28, 175-202.

Segal, Uzi and Avia Spivak. (1990). "First order versus second order risk aversion," *Journal of Economic Theory* 51, 111-125.

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

Tversky, Amos and Craig R. Fox. (1995). "Weighing risk and uncertainty," *Psychological Review* 102, 269-283.

Tversky, Amos and Daniel Kahneman. (1992). "Advances in prospect theory: Cumulative representation of uncertainty," *Journal of Risk and Uncertainty* 5, 297-323.

von Neumann, John and Oskar Morgenstern. (1944, 1947). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Wakker, Peter P. (1994). "Separating marginal utility and probabilistic risk aversion," *Theory and Decision* 36, 1-44.

Wakker, Peter P. (2004). "Preference axiomatizations for decision under uncertainty." In Itzhak Gilboa (ed.), *Uncertainty in economic theory: Essays in honor of David Schmeidler's 65th birthday*. London: Routledge.

Wakker, Peter P. and Daniel Deneffe. (1996). "Eliciting von Neumann-Morgenstern utilities when probabilities are distorted or unknown," *Management Science* 42, 1131-1150.

Wu, George, Jiao Zhang, and Richard Gonzalez. (2004). "Decision under risk." In Derek J. Koehler and Nigel Harvey (eds.), *The Blackwell Handbook of Judgment and Decision Making*. Oxford: Oxford University Press.

Tables

Table 1: Simulation scenarios

Scenario number	Marginal distribution of α	Marginal distribution of δ	Marginal distribution of γ	Correlation structure
I	uniform (0.33, 1.45)	uniform (0.19, 1.11)	uniform (0.17, 1.03)	α, δ, γ mutually independent
II	uniform (0.33, 1.45)	uniform (0.19, 1.11)	uniform (0.17, 1.03)	$\rho_{\alpha,\delta} = +1$; γ independent
III	uniform (0.33, 1.45)	uniform (0.19, 1.11)	uniform (0.17, 1.03)	$\rho_{\alpha,\delta} = -1$; γ independent
IV	triangular (0.10, 0.89, 1.68)	triangular (0.00, 0.65, 1.30)	triangular (0.00, 0.60, 1.20)	α, δ, γ mutually independent
V	triangular (0.10, 0.89, 1.68)	triangular (0.00, 0.65, 1.30)	triangular (0.00, 0.60, 1.20)	$\rho_{\alpha,\delta} = +1$; γ independent
VI	triangular (0.10, 0.89, 1.68)	triangular (0.00, 0.65, 1.30)	triangular (0.00, 0.60, 1.20)	$\rho_{\alpha,\delta} = -1$; γ independent

Table 2: Precision of the estimation

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
mean of $\lvert\hat{\alpha} - \alpha\rvert$	0.000	0.039	0.074	0.102	0.125	0.161	0.188	0.211
mean of $\lvert\hat{\delta} - \delta\rvert$	0.000	0.043	0.076	0.099	0.117	0.142	0.158	0.172
mean of $\lvert\hat{\gamma} - \gamma\rvert$	0.000	0.023	0.042	0.056	0.067	0.084	0.098	0.111

Table 3: Correlation coefficients between parameter estimates (Scenario I, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	-0.012	-0.080	-0.202	-0.275	-0.318	-0.371	-0.401	-0.420
$\rho(\hat{\alpha},\hat{\gamma})$	0.001	0.008	0.033	0.053	0.066	0.093	0.116	0.127
$\rho(\hat{\delta},\hat{\gamma})$	-0.022	-0.041	-0.060	-0.075	-0.093	-0.107	-0.132	-0.142

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 4: Correlation coefficients between parameter estimates (Scenario II, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	1	0.966	0.859	0.731	0.616	0.435	0.299	0.192
$\rho(\hat{\alpha},\hat{\gamma})$	0.001	0.002	0.015	0.029	0.041	0.052	0.058	0.064
$\rho(\hat{\delta},\hat{\gamma})$	0.001	-0.018	-0.043	-0.063	-0.082	-0.100	-0.105	-0.099

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 5: Correlation coefficients between parameter estimates (Scenario III, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	-1	-0.962	-0.923	-0.886	-0.858	-0.824	-0.809	-0.796
$\rho(\hat{\alpha},\hat{\gamma})$	0.001	0.013	0.046	0.070	0.088	0.124	0.151	0.173
$\rho(\hat{\delta},\hat{\gamma})$	-0.001	-0.015	-0.047	-0.064	-0.079	-0.107	-0.145	-0.163

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 6: Correlation coefficients between parameter estimates (Scenario IV, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	0.014	-0.062	-0.165	-0.237	-0.291	-0.370	-0.414	-0.441
$\rho(\hat{\alpha},\hat{\gamma})$	0.012	0.036	0.057	0.071	0.085	0.104	0.125	0.134
$\rho(\hat{\delta},\hat{\gamma})$	-0.024	-0.053	-0.068	-0.086	-0.106	-0.125	-0.138	-0.145

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 7: Correlation coefficients between parameter estimates (Scenario V, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	1	0.967	0.874	0.754	0.628	0.430	0.278	0.168
$\rho(\hat{\alpha},\hat{\gamma})$	0.012	0.024	0.036	0.039	0.039	0.047	0.054	0.056
$\rho(\hat{\delta},\hat{\gamma})$	0.012	0.006	-0.006	-0.027	-0.047	-0.063	-0.077	-0.080

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 8: Correlation coefficients between parameter estimates (Scenario VI, N = 41)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	-1	-0.957	-0.928	-0.907	-0.889	-0.862	-0.839	-0.820
$\rho(\hat{\alpha},\hat{\gamma})$	0.012	0.044	0.066	0.094	0.113	0.145	0.162	0.165
$\rho(\hat{\delta},\hat{\gamma})$	-0.012	-0.039	-0.071	-0.098	-0.122	-0.157	-0.173	-0.181

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 9: Correlation coefficients between parameter estimates (Scenario I, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	0.012	-0.068	-0.175	-0.247	-0.293	-0.335	-0.373	-0.378
$\rho(\hat{\alpha},\hat{\gamma})$	-0.005	0.003	0.025	0.048	0.057	0.082	0.111	0.122
$\rho(\hat{\delta},\hat{\gamma})$	-0.030	-0.044	-0.062	-0.086	-0.067	-0.067	-0.128	-0.118

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 10: Correlation coefficients between parameter estimates (Scenario II, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	1	0.968	0.872	0.753	0.639	0.445	0.298	0.186
$\rho(\hat{\alpha},\hat{\gamma})$	-0.005	-0.004	0.006	0.017	0.025	0.032	0.039	0.051
$\rho(\hat{\delta},\hat{\gamma})$	-0.005	-0.021	-0.039	-0.055	-0.067	-0.081	-0.086	-0.095

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 11: Correlation coefficients between parameter estimates (Scenario III, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	-1	-0.962	-0.927	-0.905	-0.871	-0.811	-0.788	-0.750
$\rho(\hat{\alpha},\hat{\gamma})$	-0.005	0.004	0.036	0.082	0.094	0.125	0.154	0.116
$\rho(\hat{\delta},\hat{\gamma})$	0.005	0.002	-0.027	-0.065	-0.065	-0.085	-0.145	-0.135

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 12: Correlation coefficients between parameter estimates (Scenario IV, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	0.026	-0.041	-0.136	-0.219	-0.278	-0.344	-0.374	-0.409
$\rho(\hat{\alpha},\hat{\gamma})$	-0.003	0.005	0.025	0.037	0.050	0.065	0.073	0.094
$\rho(\hat{\delta},\hat{\gamma})$	-0.023	-0.038	-0.067	-0.090	-0.112	-0.073	-0.085	-0.157

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 13: Correlation coefficients between parameter estimates (Scenario V, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	1	0.968	0.878	0.760	0.643	0.437	0.276	0.154
$\rho(\hat{\alpha},\hat{\gamma})$	-0.003	0.001	0.008	0.011	0.011	0.011	0.024	0.033
$\rho(\hat{\delta},\hat{\gamma})$	-0.003	-0.007	-0.016	-0.028	-0.039	-0.059	-0.068	-0.065

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Table 14: Correlation coefficients between parameter estimates (Scenario VI, N = 2000)

	$\sigma_\varepsilon = 0$	$\sigma_\varepsilon = 1$	$\sigma_\varepsilon = 2$	$\sigma_\varepsilon = 3$	$\sigma_\varepsilon = 4$	$\sigma_\varepsilon = 6$	$\sigma_\varepsilon = 8$	$\sigma_\varepsilon = 10$
$\rho(\hat{\alpha},\hat{\delta})$	-1	-0.948	-0.927	-0.860	-0.878	-0.813	-0.752	-0.729
$\rho(\hat{\alpha},\hat{\gamma})$	-0.003	0.016	0.039	0.061	0.082	0.099	0.109	0.121
$\rho(\hat{\delta},\hat{\gamma})$	0.003	-0.014	-0.025	-0.015	-0.046	-0.049	-0.054	-0.067

Note: $\rho(\hat{\alpha},\hat{\delta}) = \rho(\alpha,\delta)$ for $\sigma_\varepsilon = 0$ (and likewise for $\rho(\hat{\alpha},\hat{\gamma})$ and $\rho(\hat{\delta},\hat{\gamma})$).

Figures

Figure 1: Scatterplot $\rho(\hat{\alpha},\hat{\delta})$ vs. $\rho(\alpha,\delta)$ for $\sigma_\varepsilon = 3$ (left) and $\sigma_\varepsilon = 10$ (right)