Scottish Journal of Political Economy, Vol. 51, No. 3, August 2004

EXPLAINING ATTITUDES TOWARDS AMBIGUITY: AN EXPERIMENTAL TEST OF THE COMPARATIVE IGNORANCE HYPOTHESIS

Paul Dolan and Martin Jones***

Abstract

Many theories have been put forward to explain attitudes towards ambiguity. This paper reports on an experiment designed to test for the existence of Comparative Ignorance when it is tested over events with a range of different likelihoods. A total of 93 subjects valued a series of gambles, one of which was played out for real. The results do not lend support to the theory, although the relationship between risk and ambiguity does appear to correspond with other theories and previous empirical work.

I Introduction

Background

Ambiguity has been a topic of research in experimental economics and decision theory ever since Ellsberg's classic paper (Ellsberg, 1961), which drew on Keynes (1921) and Knight (1921) for inspiration. A variety of experiments have been carried out and many theories have been put forward (see Camerer and Weber, 1992, for a review). The Comparative Ignorance hypothesis (CIH) represents a departure from most of these other theories. Instead of relating behaviour under ambiguity to the internal state of the decision-maker, it relates behaviour to *the context within which the decision-maker operates*. It states that agents will exhibit an aversion to ambiguity when they are able to compare their lack of knowledge in one situation to their relatively greater knowledge in another situation – otherwise agents will exhibit ambiguity aversion to a much lesser extent (Fox and Tversky, 1995).

There are several traditions in the testing of choice under ambiguity, all of which are variants of the ideas originally tested by Ellsberg, but many of which use different definitions of ambiguity and are the result of different interpretations of what is actually meant by ambiguity. The aim of this paper

*University of Sheffield
**University of Dundee

is to combine the ideas coming out of two of these traditions. One is founded on the preference-based notion of ‘subjective’ ambiguity as defined by Fox and Tversky, and the other on the notion of ambiguity as being an objective lack of knowledge of the probability distribution (e.g. Hogarth and Einhorn, 1990; also Becker and Brownson, 1964; Camerer and Weber, 1992). In the subjective case, a person is ambiguity averse when she prefers a risky gamble, A, over ambiguous gamble, B, and the complementary risky gamble, A′, over the complementary ambiguous gamble, B′. This means that the uncertainty is directly linked via preference to events in the gambles rather than being defined in terms of probabilities. It follows that discovering whether someone is ambiguity averse is a matter of a subject valuing both the risky and ambiguous gambles and also their complements.

According to the objective tradition, a person is said to be ambiguity averse when she consistently prefers a risky gamble over an equivalent ambiguous gamble, where ‘equivalent’ means that the risky gamble has a probability of success equal to the mid-point of the range of probabilities in the ambiguous gamble (see Curley and Yates, 1989). In this case, ambiguity is related to a range of possible probability distributions. In this study, we will test whether ambiguity aversion (that comes out of either tradition) varies systematically with different likelihoods of events.

The comparative ignorance hypothesis

The CIH states that ‘ambiguity aversion will be present when subjects evaluate clear and ambiguous prospects jointly, but it will greatly diminish or disappear when they evaluate each prospect in isolation’ (Fox and Tversky, 1995). This is an extension of the earlier Competence hypothesis which states that people will be relatively ambiguity loving in situations where they believe themselves to have particular expertise or knowledge (Heath and Tversky, 1991).[1] The CIH extends this idea from being a lack of one’s own knowledge to a lack of knowledge in general. It is the extent to which people *know* that they lack knowledge in one prospect relative to another that is crucial to the CIH. Ambiguity aversion in this case is defined according to the implied preferences emerging from the Ellsberg Paradox (Tversky and Wakker, 1995). Ambiguity aversion can be said to be present when a person prefers one lottery and its complementary lottery to another lottery and its complement. This pattern of preferences violates expected utility. This is a purely subjective definition in that it is based on preferences rather than an objective lack of knowledge of the outside world.

The CIH is derived from the same general framework as that set up by Tversky and Kahneman’s (1992) Cumulative Prospect Theory (CPT). CPT was originally a theory of choice under risk but was extended to cover choice under ambiguity by incorporating the preference structure implied by Ellsberg’s experiments. In developing this further, Tversky and Fox (1995) distinguish

[1] To support this hypothesis, Heath and Tversky (1991) demonstrated that people were more willing to bet on their answers to questions that they felt relatively confident about than on risky gambles with a similar chance of winning.

between source sensitivity and source preference. *Source sensitivity* refers to the (upper and lower) sub-additivity of the decision weights, while *source preference* is the observation that choices between prospects depend not only on the degree of uncertainty in one particular source but also from *which* source the uncertainty comes. Fox and Tversky argue in their paper that ambiguity aversion is characterised by source preference rather than source sensitivity.[2] It should be noted, however, that CIH stands as a hypothesis on its own, and does not rely on CPT for any of its theoretical elements.

Previous evidence on the CIH comes largely from six studies reported in Fox and Tversky (1995), the first three of which are relevant here, and from a replication study by Chow and Sarin (2001). The principal innovation in each of these studies was to examine the difference between valuations of risk and ambiguity both within and across persons. So, subjects in one group would value *both* a risky *and* an ambiguous lottery, and subjects in two other groups would value *either* a risky *or* an ambiguous lottery. Thus, subjects in the first group (the 'comparative' group) could compare the risky and ambiguous gambles, whilst subjects in the other two groups (the 'non-comparative' groups) would be unable to do so.

The first study of Fox and Tversky (1995) replicated Ellsberg's two-urn experiment: one urn contained 50 black balls and 50 red balls and the other contained 100 balls in an unknown mix of red and black balls. Subjects were asked to choose a colour and then to place a certainty equivalent value on one or both urns on the understanding that they would get a (hypothetical) prize if a ball drawn matched their chosen colour. Some subjects valued both urns (the comparative condition) while others valued just one urn (the non-comparative condition). The results were consistent with the CIH i.e. the difference between the values for the risky and ambiguous gambles was significant in the comparative condition but disappeared in the non-comparative condition. The second study, which was similar to the first experiment except that monetary incentives were used, produced similar results. Chow and Sarin, in their similar first experiment, found that the difference between the values for the risky and ambiguous gambles did not disappear in the non-comparative condition, but was still less than in the comparative condition.

The third study was derived from Ellsberg's second experiment in which subjects were faced with a bag containing 30 balls, 10 of which are known to be white and 20 of which could be any mixture of red and blue. The subjects were asked to value four lotteries based on which ball was picked out of the bag: lottery 1 had a prize of \$50 for a white ball, lottery 2 had \$50 for a red ball, lottery 3 had \$50 for a blue ball or a white ball, and lottery 4 had \$50 for a red or a blue ball. This effectively gives bets on an unambiguous probability of 1/3, a probability range between 0 and 2/3, a probability range between 1/3 and 1 and an unambiguous probability of 2/3. Some subjects valued all four lotteries whilst others valued either both of the unambiguous or both of the ambiguous

[2] Tversky and Fox (1995) give source preference and source sensitivity equal weighting and they describe their effects as being independent of each other.

lotteries. This ensured that the latter two groups valued lotteries with complementary events while the first group valued two sets of lotteries with complementary events, one set of which was ambiguous and the other set of which was risky.[3] Again, the results were consistent with the CIH, particularly for the higher probability lotteries. Chow and Sarin again produced qualitatively similar results in their third study, but there was still a difference between the values for the risky and ambiguous gambles in the non-comparative condition.

The relationship between risk and ambiguity attitudes

Much of the research into the relationship between attitudes towards risk and ambiguity revolves around the notion of a weighting function that takes the place of probabilities in decision functions. According to CPT, for example, these risky or ambiguous weights are calculated using cumulative events of increasing likelihood rather than individual events. However, rather than focus on the links between preference and decision weights, this paper is more concerned with the *relationships between risky and ambiguous weights*.

There is, though, a wide range of empirical evidence that provides insights into what such a relationship might look like. Specifically, the literature suggests that the relationship might be one in which risk is preferred to ambiguity with events of moderate and high likelihood (Curley and Yates, 1985, 1989; Cohen, Jaffray and Said, 1985; Hogarth and Einhorn, 1990; Camerer and Weber, 1992) and vice versa (though to less of an extent) at events of low likelihood (Curley and Yates, 1985; Kahn and Sarin, 1988; Camerer and Weber, 1992). This literature tends to use the 'objective' notion of ambiguity where ambiguity aversion occurs simply when a risky gamble is preferred to the ambiguous gamble with the equivalent likelihood of winning. A similar set of results, from the subjective point of view, can be derived from work by Tversky and Wakker (1995) and Tversky and Fox (1995), who look at the relationship between decision weights under risk and ambiguity. The general conclusion is that ambiguous weights are more (upper and lower) subadditive than risky weights. This suggests the same general relationship of ambiguity aversion for events of high likelihood and ambiguity loving for events of low likelihood. However, there seems to be a variety of possibilities for what could happen at events with middling likelihoods.

A theoretical model in the 'objective ambiguity' tradition that produces this kind of relationship between attitudes to risk and ambiguity is *Venture Theory* (Hogarth and Einhorn, 1990). Venture Theory is based on an anchoring and adjustment model where the anchor is the true (objective) probability (P_A) and the adjustment weight is obtained as a result of a simulation carried out by the individual. The result of this is shown in Figure 1 which has a venture function for ambiguous decisions (with a weight of $W'(P_A)$) relative to risky ones (with a

[3] This differs in one important respect from the other experiment in that the subject is not allowed to choose which ball to play – however since, by stochastic dominance, one should always choose the gamble with the highest likelihood, this cannot be helped if one wishes to move off 50:50 likelihoods.

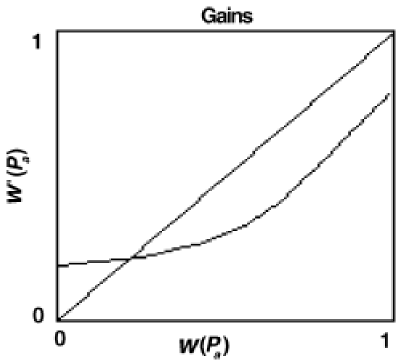

Figure 1. The venture theory relationship function for risky and ambiguous gambles in the domain of gains.

weight of $W(P_A)$).[4] Notice that ambiguous gambles have higher values than risky ones when P_A is low but that the opposite is true when the likelihood of winning is middling or high. While we may not necessarily subscribe to the underlying theory on which Venture theory is based, it is an exemplar of the type of relationship that has been found in the theoretical and empirical literature. So, we refer to a Venture Theory-style relationship only insofar as it may describe the relationship between ambiguity and risk attitudes.

Experimental design

One of the main aims of this study is to test the robustness of the CIH through a range of events of different likelihood. Although Fox and Tversky's (1995) third study, and Chow and Sarin's studies three and four, tested the CIH at probabilities of one-third and two-thirds (and showed that comparative ignorance was stronger when the chances of winning were higher), they did not test what happens when the probabilities were closer to 0 or 1. The experiment in this paper allows subjects to compare ambiguous and risky gambles in a far larger range of likelihoods.

When looking at the subjective form of ambiguity, we cannot really talk about 'low' likelihood events as such, because we need to consider complementary events (so that every 'low' likelihood event has a corresponding complementary event that is of 'high' likelihood). Fox and Tversky sum together the valuations of the complementary events both when the events are ambiguous and when they are risky, and so it is not possible to look for the existence of a Venture Theory style relationship in this case.

We can examine Comparative Ignorance in terms of the 'objective' definition of ambiguity aversion by comparing valuations of equivalent risky and

[4] It should be noted that this diagram is similar to a series of diagrams relating probabilities to decision weights and decision weights to each other starting with Kahneman and Tversky's (1979) prospect theory. However, unlike Kahneman and Tversky's diagram this is a diagram showing the relationship between types of decision weights which originated in Einhorn and Hogarth (1986) and improved in Hogarth and Einhorn (1990).

ambiguous events with different likelihoods. Comparative Ignorance in this case would simply predict that, in general, subjects would exhibit more ambiguity aversion when valuing risky and ambiguous gambles together than when they value the gambles separately. Contrary to the 'subjective' case, it is possible to look at these differences between gambles to see whether, qualitatively, one can map out a Venture Theory relationship between attitudes to risk and ambiguity at different likelihoods.

All of this allows us to investigate the relationship between the comparative ignorance hypothesis, on one hand, and source sensitivity (in the form of a Venture Theory-style function), on the other. Two hypotheses emerge from this:

(i) At high and moderate likelihoods, there would be evidence of Comparative Ignorance.
(ii) At low likelihoods, there may be no evidence of Comparative Ignorance.

The first hypothesis derives from Fox and Tversky's idea of comparative ignorance – there would be ambiguity aversion in the comparative condition. This would simply be reinforced by the aversion to ambiguity shown at higher likelihoods. However, at lower likelihoods subjects may be ambiguity-loving and this would work against the fact that subjects are able to compare risky and ambiguous gambles, negating the effects of Comparative Ignorance.

The experiment presented below is, in some respects, an extension of Fox and Tversky's third study in that it uses objective probabilities and objective probability ranges to define its gambles.[5] The experiment here gives the subjects a set of gambles that are ordered in terms of the likelihood of winning, either in ascending or descending order.[6] As well as allowing for an investigation into any possible interactions between gambles that are close to one another in terms of likelihood, this design also allows for any possible order effects to be identified.

II Methods

The experiment

Subjects were asked to value a number of 'risky' and/or 'ambiguous' gambles.[7] They were asked to imagine a 'bag' containing 100 balls, some of which were blue and some of which were green. For risky gambles, they were told the precise

[5] According to Fox and Tversky, their study 3 was susceptible to CI in the non-comparative treatment because the existence of non-ambiguous balls allowed CI to operate. The same problem could apply here.

[6] The breadth of the ambiguity range was a serious experimental problem. We did not wish to be open to the accusation that the range was too narrow, giving little difference between risky and ambiguous gambles, but, at the same time, we wished to have gambles close to zero and 100% that required narrow ranges. In the end, we compromised with narrow ranges at the extremes, but with the four middle probabilities having wider ranges.

[7] Ambiguity is defined here as a range of possible probabilities and *not* a second-order probability (see Camerer and Weber, 1992) for the distinction between the two). This definition has been used by Curley and Yates (1985), amongst others, and is used in study 3 of the Fox and Tversky experiment.

number of blue and green balls in the bag. For ambiguous gambles, they were told that the bag contained a known number of blue balls, a known number of green balls and a number of balls that were green and blue in unknown proportions. The subjects were told that each type of gamble would be resolved by picking a ball out of the imaginary bag such that £10 would be won if the ball was blue and nothing would be won if the ball was green. The risky and ambiguous gambles had computerised layouts similar to those shown in Figure 2. The black area represents the number of blue balls in the bag, the grey area represents the number of green balls and the white area represents the unknown mixture of blue and green balls.

The mechanism used to value the gambles was a variant of the Becker-DeGroot-Marschak preference elicitation device (Becker *et al.*, 1964), was used by Fox and Tversky (1995) in their second study, and by Chow and Sarin (2001). In each of 100 envelopes was a piece of paper with a sum of money (ranging from £0.10 to £10.00 in 10p increments) written on it. At the beginning of the experiment, one envelope was chosen at random (without replacement) for each of the gambles. These envelopes were placed in public view. For each gamble, subjects were asked to enter into the computer the amount of money they would just be willing to receive in exchange for the gamble. They did this on the understanding that, at the end of the experiment, the envelope corresponding to the gamble that they would play out for real would be opened. If the amount of money written on the paper in the envelope was greater than or equal to their certainty equivalent value then they would receive the amount on the paper in the envelope; otherwise they would play out the gamble.

The questions were answered on an individual basis and the subjects were isolated from each other. To enable the subjects to familiarise themselves with what was being asked of them, the experiment began with a set of oral instructions that were read out in conjunction with three practice questions. Subjects in Group 1 were presented with three computer screens, each of which had one risky and one ambiguous gamble on it. Subjects in Group 2a had one risky gamble per screen and those in Group 2b had one ambiguous gamble per

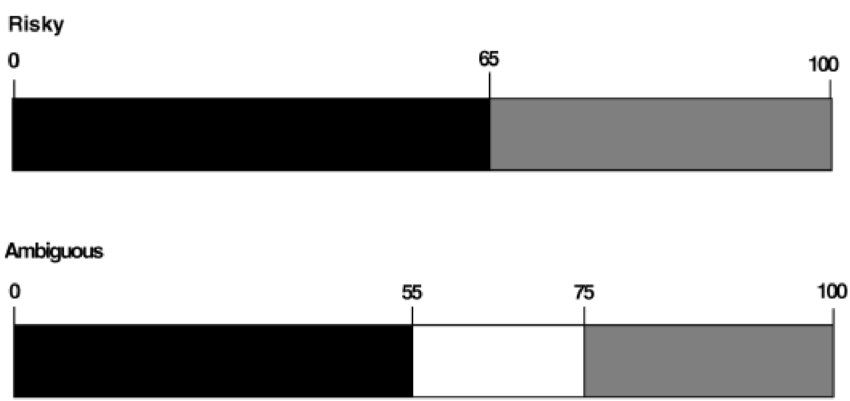

Figure 2. The lay-out of the gambles used in the experiment.

screen. For all subjects, the three practice questions were based around probabilities of 0.5, 0.75 and 0.25. As subjects went through each question on each of the screens, the experiment supervisor talked them through the mechanics of the experiment and described what the diagrams on the screen represented. The mechanism by which the gambles were to be valued was also explained. The main part of the experiment then involved each subject being asked to give certainty equivalent values to 16 gambles (in the case of group 1) or 8 gambles (in the case of groups 2a and 2b). Subjects were told that one gamble would be chosen at random, and played out for real.

At the end of the experiment, each subject rolled a die to determine which gamble would be played out. The certainty equivalent value that the subject had entered for the gamble was compared to the amount in the envelope for that gamble. If the latter was equal to or greater than the former, the subject was paid the amount in the envelope. But if the certainty equivalent value was greater than the amount in the envelope, then the gamble was played out. For risky gambles, a grid with 100 squares containing either the letter 'B' for blue or 'G' for green was used. The letters filled the grid in the proportion of blue and green balls for that gamble. Without seeing the grid, the subject then marked a square on a blank grid. If the corresponding square on the filled in grid had a 'B' in it, then the subject was paid £10; otherwise she was paid nothing.

A similar mechanism was used for ambiguous gambles, but first the ambiguity had to be resolved. Each subject was first given two potential proportions of blue and green balls together with a 'critical' number between 1 and 6. A die was then rolled and, if the number on the die was less than the critical number, the first proportion was used; otherwise the second proportion was used. Whichever proportion was selected, the appropriate grid was used in the same way as for risky gambles. This method was used as a way of operationalising the idea of ambiguity without introducing any idea of a second order probability into the subjects' perceptions. The subjects were not told beforehand what the two proportions were and were only told in the instructions that these two proportions were chosen 'arbitrarily' by the experimenter. The critical number was also unknown to the subjects and, likewise, was said to have been chosen 'arbitrarily'. No more information was given apart from this. There was, therefore, no explicit implication of one particular probability distribution in the experimental design.[8]

There were three experimental groups. Group 1 was the comparative group, and so subjects were presented with a risky *and* an ambiguous gamble on the screen at the same time. The probability of the risky gamble was equal to the centre of the range of possible probabilities in the ambiguous gamble. Group 2 was the non-comparative group – subjects in 2a valued only risky gambles and those in 2b valued only ambiguous gambles. Each group was split into two

[8] Arguably the subjects may have been 'suspicious' of the experimenter's motives. However, this is an empirical hypothesis about the subjects' attitude to the experiment and so needs to be tested in another experiment. It could be the case, for example, that the subjects were overly optimistic about the chances of winning rather than suspicious. There were no signs of suspicion in the experiment or in subjects' reactions to the experiment.

subgroups. The first subgroup did the tasks in descending order of the likelihood of a blue ball being drawn (referred to as the 'down' subgroup) while the second subgroup did the experiment in ascending order of the likelihood of a blue ball being drawn (referred to as the 'up' subgroup). Table 1 shows the questions presented to the groups. The numbers refer to the number (or range) of blue balls in the bag, the remainder (i.e. up to 100) being taken up by green balls.

Hypotheses tested

It is assumed that individuals in the experiment have stochastic preferences in the sense of Loomes and Sugden (1995). Under this theory, individuals' preferences vary not as a result of errors as such, but as a result of preferences themselves being random. Preference orderings are drawn from a probability distribution of such preference orderings. In this case, the randomness comes from variations within the individuals themselves, some of which may derive from errors in subjects' calculations, individual idiosyncracies etc. For the purposes of testing there are two possible 'core theories' – expected utility (EU) and procedural invariance.

Before testing the CIH, a number of other tests for procedure invariance were performed (Tversky, Sattath and Slovic, 1988). In the context of this paper, procedure invariance means that subjects' behaviour does not vary with equivalent elicitation procedures. Comparisons can be made between the valuations given to the risky gambles by Groups 1 and 2a and between the valuations given to the ambiguous gambles by Groups 1 and 2b to see if there was a significant difference in how they were valued. Since subjects in each group were split into 'up' and 'down' subgroups, it is also possible to test whether responses are affected by the order in which questions are presented to respondents. A further order test was carried out on Group 1. Here, the *differences* between the values for risky and ambiguous gambles at each risk level were correlated with one another to see whether the differences between one pair of gambles influenced the differences between subsequent pairs of gambles.

Table 1
Question parameters

Risky gambles Group 1 and 2a	Ambiguous gambles Group 1 and 2b
95	90–100
90	80–100
80	60–100
60	40–80
40	20–60
20	0–40
10	0–20
5	0–10

Note:
The numbers refer to the probability of picking a blue ball, and hence of winning £10.

The CIH was tested in the following manner across subjects. Subjects in Group 1 were asked to evaluate risky and ambiguous gambles together whilst those in Groups 2a and 2b were asked to evaluate them in isolation from each other. Two sets of tests were carried out, one that tested the subjective view of ambiguity aversion and the other that tested the objective view. In the former, following the procedure in Fox and Tversky's third study, the valuations of the complementary risky gambles were added together (e.g. the values of the 5% and 95% gambles are added together), and likewise for the ambiguous gambles. The second set of tests, for the objective view of ambiguity, involved comparing directly equivalent risky and ambiguous gambles (e.g. the values of the 5% risky gamble are compared with the values of the 0–10% ambiguous gamble), both within Group 1 and across Groups 2a and 2b. For both sets of tests, according to the CIH, Groups 2a and 2b should value the risky and ambiguous gambles closer together than those in Group 1. According to the 'objective' view of ambiguity, if there is a Venture Theory-type relationship, ambiguous gambles should be valued higher than risky ones for low likelihood events but risky gambles should be valued higher than ambiguous ones thereafter.

The null hypothesis for those tests involving procedure invariance is that the elicitation procedure makes no difference, whatever the underlying model. A similar assumption was made for direct tests of Comparative Ignorance since CIH is effectively a test across two different types of treatment, but all using the same elicitation method. By contrast, EU is used as the null hypothesis in all tests of ambiguity aversion. In those gambles involving ambiguity, EU in its basic form is not a sufficient null hypothesis because of the vagueness of the probabilities and hence of the probability distributions underlying them within the ambiguity ranges. In these cases, EU is supplemented with the idea that, when faced with ambiguity, subjects place equal likelihood on there being either a blue or a green ball in the mixture. This effectively splits the ambiguity range in half, thus making it equal to the corresponding risky gamble (see also Curley and Yates, 1985). While this cannot be argued to be fully rational it can be seen as being plausible for the purposes of creating a null hypothesis.

To test for normality, the Kolmogorov-Smirnov goodness-of-fit test was applied to the various subgroups of data. Generally speaking, the distributions of values in Group 1 passed the test, whilst those in Groups 2a and 2b failed the test. There were many cases of multiple peaks in the frequency distributions and there was considerable clustering around whole numbers (£1, £2, £3 etc.) despite the fact that subjects could value the gambles in 10p increments. For these reasons, non-parametric statistics were used. Therefore, if i and j are the groups (or subgroups) being compared, the test is:

$$\mathrm{H}_0\colon M_i = M_j$$

$$\mathrm{H}_1\colon M_i \neq M_j$$

where $M_a(a = i,j)$ is the median value for a particular group or subgroup. All tests were performed using the Wilcoxon Test (in within-subject comparisons) and the Mann-Whitney U Test (in between-subject comparisons). Significant differences are reported at the $p < 0.05$ level. When assessing correlations between two variables, Spearman's Rank Correlation coefficient (Rho) was used. A two-way ANOVA with interaction effects was used in order to look at the interaction between the type of gamble valued (risky or ambiguous) and the experimental condition (comparative or non-comparative).[9]

III Results

The sample and data quality

The subjects were undergraduate students at the University of Newcastle. The experiment took place in March 1998 over a period of two weeks. The subjects were recruited via an e-mail database and consisted of students from all subjects across all three years. The sample did include some students who had done economics or psychology but most of these would not have taken a course in decision theory. A total of 99 subjects took part in the experiment. Six subjects were excluded on the grounds that they had a pronounced tendency to value the gambles inversely to the likelihood of picking a blue ball, thus leaving 93 in the final dataset. Because of different numbers of subjects turning up to each experimental session, there were between 28 and 34 subjects in each of the three groups and between 13 and 17 subjects in each of the 'up' and 'down' subgroups.

The relatively low number of exclusions suggests that, on the whole, the experiment was well understood by the subjects. To look at this more closely, the valuations for each question were ranked in order and the Spearman's Rank Correlation Coefficient was used to show how each individual's rankings are correlated with one another across questions. It would be expected that if a subject gave a relatively high (low) value to one question, they would give a relatively high (low) value to other questions. Overall, the results suggest that this is indeed the case, particularly in Group 1 where 84% of all correlation coefficients are significant at the 5% level. The comparable figures for Groups 2a and 2b are 55% and 44%, respectively. As would be expected, the rankings that are not significantly correlated with one another involve comparisons of questions that are separated by a number of other questions. This suggests that answers to the questions were consistent rather than being done at random.

Procedure invariance

The first test of procedure invariance involved comparing evaluations of identical gambles in comparative and non-comparative situations. When comparing the certainty equivalent values of the risky gambles from Group 1 with those from Group 2a, there was only one significant difference (in question 5) between the two. There were no significant differences between the values given to the ambiguity gambles by Groups 1 and 2b. Overall, then, the subjects

[9] A non-parametric version of the interaction effect is not available for unbalanced sets of data (see Sawilowsky, 1990, for examples of non-parametric interaction statistics), so parametric statistics are the best alternative available.

displayed procedure invariance when evaluating identical gambles in comparative and non-comparative conditions.

The second test of procedural invariance involved a comparison of the values given by 'up' and 'down' subgroups. The results are summarised in Table 2, In Group 1, the 'up' subgroup gave significantly higher values than the 'down' subgroup to 10 of the 16 questions. There is one significant difference (out of eight comparisons) in Group 2a and three (out of eight) in Group 2b. The 'up' subgroup produced higher median values than the 'down' subgroup in 27 of the 32 comparisons and the difference was £1 or more on 14 occasions. On only one occasion (in Group 2a) did the 'down' subgroup produce a higher median value than the 'up' subgroup. All of this suggests that there is a powerful order effect in the data. A test was also carried out for the effects of order on the differences in value in Group 1. In general there were no signs of a systematic order effect – only two of the 32 correlations were significant (the coefficient between the eighth and seventh questions and that between the fifth and fourth questions). This showed that while there was indeed an order effect in valuations between

Table 2
Differences between 'up' (starting with a low probability first) and 'down' (starting with a high probability first) subgroups

Group 1

Risk level	Group 1 risk 'up'	Group 1 risk 'down'	Z-statistic	Group 1 ambiguity 'up'	Group 1 ambiguity 'down'	Z-statistic
95	9.00	8.50	– 1.974*	9.30	8.40	– 2.733*
90	9.00	8.35	– 1.970*	8.00	7.50	– 1.319
80	8.00	7.25	– 2.258*	7.00	5.50	– 1.988*
60	6.00	4.50	– 2.833*	6.00	4.00	– 2.965*
40	5.00	3.50	– 3.082*	4.00	2.25	– 2.456*
20	3.00	2.00	– 2.270*	3.00	2.00	– 1.888
10	1.50	1.00	– 1.417	1.50	1.00	– 1.212
5	1.50	1.00	– 1.433	1.00	1.00	– 1.516

Groups 2a and 2b

Risk level	Group 2a risk 'up'	Group 2a risk 'down'	Z-statistic	Group 2b ambiguity 'up'	Group 2b ambiguity 'down'	Z-statistic
95	9.10	8.50	– 1.134	8.00	8.00	– 0.363
90	8.60	7.00	– 1.455	6.50	6.00	– 0.433
80	6.00	7.00	– 0.255	5.00	4.00	– 1.040
60	5.10	4.60	– 1.294	4.00	2.60	– 1.850
40	4.00	3.00	– 2.600*	3.50	1.80	– 2.125*
20	2.00	2.00	– 0.714	2.50	1.00	– 2.389*
10	1.00	1.00	0.743	2.00	0.70	– 2.548*
5	1.00	0.80	0.472	1.00	0.70	– 1.737

Notes:
Numbers are median certainty equivalent values for gambles shown in Table 1. Statistics are asymptotic normal z-statistics (those marked * are significant at $p < 0.05$).

the 'up' and 'down' subgroups, there was no such effect between questions when it came to valuation *differences*.

Comparative ignorance

The median certainty equivalent values and inter-quartile ranges, together with the results from the ('subjective' and 'objective') tests of the differences between the risky and ambiguous gambles are shown in Tables 3 and 4. In Table 3, the median differences for the comparative and non-comparative conditions are broadly comparable. In Table 4, a comparison of the differences in median values for individual questions (in columns 4 and 8) shows that, for gambles with a high and middling likelihood of winning, the differences are *larger* in the non-comparative case than in the comparative case. For gambles involving a low likelihood of winning, the median differences are broadly comparable across conditions. These results do not support the CIH, which suggests that the differences between the values of risky and ambiguous gambles should be greater in the comparative condition, whatever the level of likelihood. The ANOVA tests for interaction effects for each summed valuation and for each level of likelihood showed that there were no significant interaction effects whether one is looking at subjective or objective ambiguity. This demonstrates that the effect of the experiment itself is the same in the comparative and non-comparative conditions.

The relationship between risk and ambiguity

In relation to 'objective' ambiguity, Figure 3(i) shows a boxplot of the differences between the values given to risky and ambiguous gambles by subjects in Group 1. It can be seen that the first five differences have their 95% confidence intervals above zero whereas the last three are around zero. These results suggest ambiguity aversion at high and middling likelihood events and ambiguity-neutrality at low likelihood events. Figure 3(ii) shows all pairs of values for risky and ambiguous gambles (ranked according to the value of the risky gamble). The generally higher values given by group 1 to risky as compared to ambiguous gambles is clear. These results provide mixed support for a Venture Theory-type relationship which predicts that risky gambles will be valued more highly at high and middling likelihood events (which they are) and that ambiguous gambles will be valued more highly at low likelihood events (which they are not).[10] These results also go against an alternative assumption of consistent ambiguity aversion at all likelihoods which is quite often assumed when discussing Keynesian and Knightian uncertainty.

[10] This, interestingly, is a similar *empirical* result to that gained by Hogarth and Einhorn (1986) at the 0.1 level of probability. However, they do not interpret this as being a problem for their theory. This may mean that they believe that the ambiguity-loving effect will become more pronounced at even lower likelihood events. Our results (with a lowest probability level of 0.05) do not support this belief.

Table 3
Differences between responses to risk and ambiguity questions – 'subjective' test of ambiguity aversion

Risk level	G1 risk	G1 ambiguity	Difference	*Z*-statistic	G2a risk	G2b ambiguity	Difference	*Z*-statistic
95+5	10.00 (9.00–11.50)	10.00 (8.90–11.30)	**0.00**	– 2.889*	9.90 (8.47–10.50)	9.25 (4.95–10.93)	**0.65**	– 0.920
90+10	10.00 (8.50–11.50)	9.00 (7.50–11.00)	**1.00**	– 3.649*	9.50 (7.73–10.37)	8.00 (5.30–9.90)	**1.50**	– 1.905
80+20	10.00 (9.00–11.00)	9.30 (8.00–11.00)	**0.70**	– 0.993	8.45 (6.85–10.00)	6.90 (3.87–9.13)	**1.55**	– 2.181*
60+40	10.00 (7.00–10.90)	8.30 (6.00–10.00)	**1.70**	– 3.138*	8.5 (7.00–10.00)	6.00 (3.00–9.00)	**1.50**	– 2.435*

Note:
Numbers refer to median (and inter-quartile range) values – (z-statistics marked * are significant at $p<0.05$).

Table 4
Differences between responses to risk and ambiguity questions – 'objective' test of ambiguity aversion

Risk level	G1 risk	G1 ambiguity	Difference	Z-statistic	G2a risk	G2b ambiguity	Difference	Z-statistic
95	9.00 (8.00–9.50)	9.00 (8.00–9.50)	**0.00**	– 2.283*	9.00 (7.50–9.50)	8.00 (4.00–9.30)	**1.00**	– 1.350
90	8.50 (8.00–9.00)	8.00 (6.00–8.60)	**0.50**	– 3.651*	8.00 (6.13–9.00)	6.00 (3.00–8.00)	**2.00**	– 2.436*
80	8.00 (6.00–8.00)	6.00 (4.00–7.50)	**2.00**	– 4.194*	6.20 (4.50–7.58)	4.00 (2.00–6.00)	**2.20**	– 2.611*
60	6.00 (4.00–6.60)	5.00 (4.00–6.00)	**1.00**	– 2.526*	5.00 (4.00–6.00)	3.50 (2.00–5.00)	**1.50**	– 3.043*
40	4.00 (2.50–5.00)	3.00 (2.00–4.00)	**1.00**	– 2.754*	3.40 (2.85–4.00)	2.10 (1.00–4.00)	**1.30**	– 2.037*
20	2.40 (1.50–4.00)	2.00 (1.00–3.50)	**0.40**	– 0.879	2.00 (1.00–3.38)	2.00 (1.00–3.00)	**0.00**	– 0.386
10	1.00 (0.70–2.50)	1.20 (0.50–2.50)	– **0.20**	– 0.648	1.00 (0.53–2.42)	1.25 (0.50–2.60)	– **0.25**	– 0.307
5	1.00 (0.50–2.00)	1.00 (0.50–2.00)	**0.00**	– 2.002*	0.90 (0.43–1.95)	0.95 (0.50–2.00)	– **0.05**	– 0.228

Note:
Numbers refer to median (and inter-quartile range) values – (z-statistics marked * are significant at $p<0.05$).

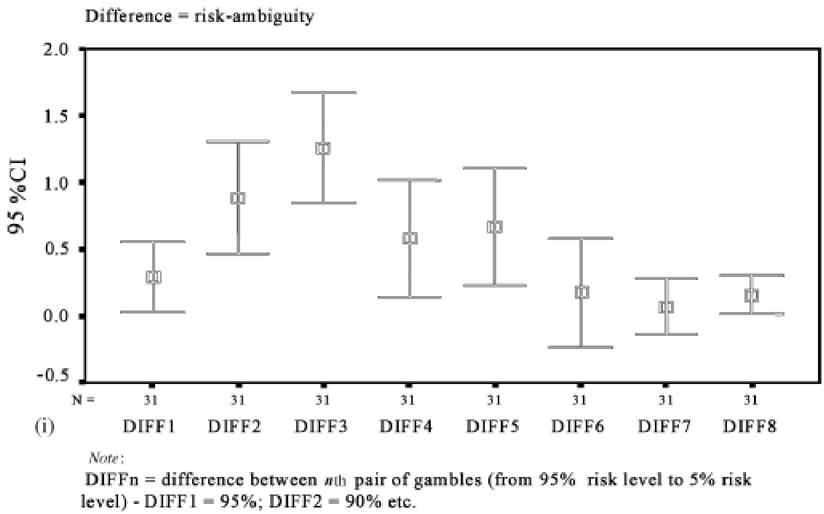

Figure 3(i). Difference between risk and ambiguity: Group 1.

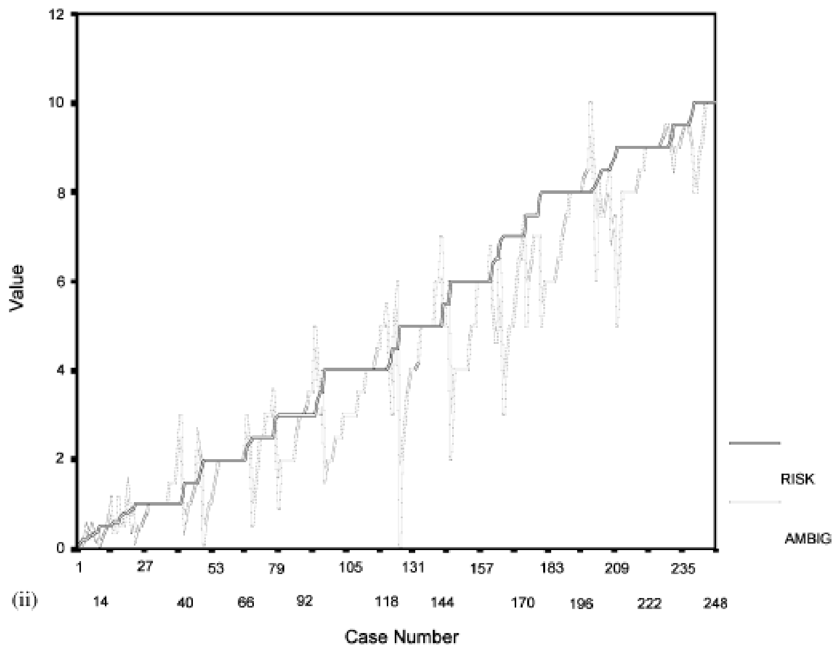

Figure 3(ii). All group 1 valuations (ranked according to value of risky gamble).

The results of the Mann-Whitney test to compare Groups 2a and 2b are qualitatively similar to the Wilcoxon test results for Group 1 (refer back to Table 4). The risky lotteries with the second to fifth highest probabilities of

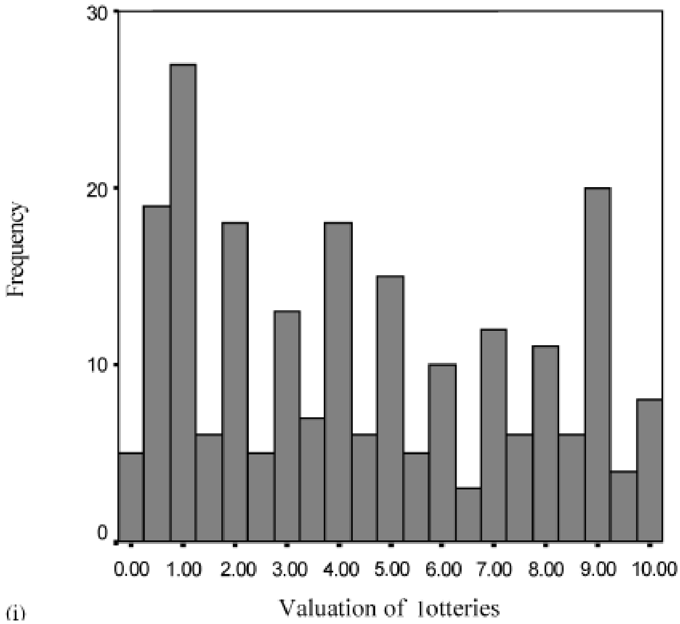

Figure 4(i). All valuations for Group 2a.

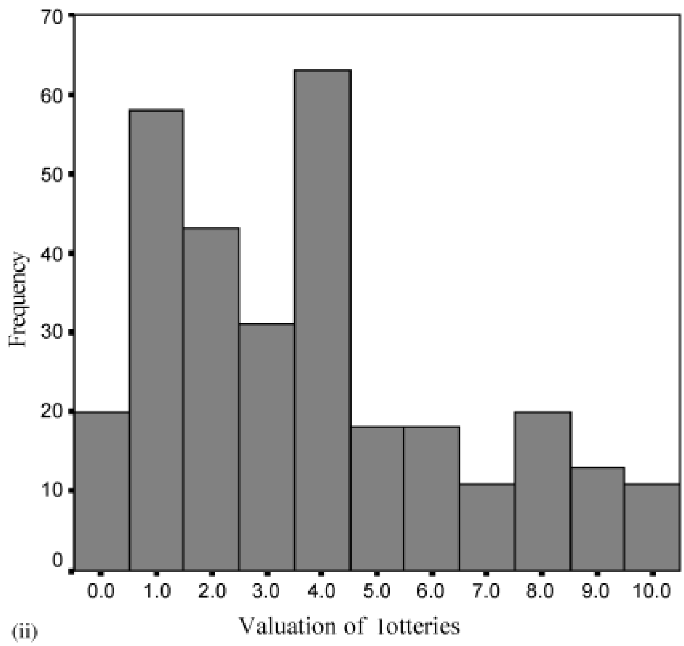

Figure 4(ii). All valuations for Group 2b.

winning have higher values than their ambiguous counterparts. The remaining comparisons do not reach statistical significance. Figures 4(i) and 4(ii) show the distributions of all values given by subjects in Groups 2a and 2b, respectively. It can be seen that the two distributions are different from one another. Group 2a's valuations (for risky gambles) are rather evenly spread with multiple peaks whereas group 2b's valuations (for ambiguous gambles) are highly skewed to the left, indicating relatively lower values. The results of the between-subject comparisons, then, provide limited support for a Venture Theory-type relationship at high and middling likelihood events and, as in the within-subject comparisons, no support for it at low likelihood events.

IV Discussion

The main aim of this paper has been to test the CIH at different levels of likelihood. The range of probabilities chosen allowed for the shape of the function that compares the value of risky and ambiguous gambles to be investigated. It also allowed for the possibility of an order effect to be tested. The data presented here show evidence of an order effect between the 'up' and 'down' subgroups. Specifically, the results suggest that the subjects used an 'anchoring and adjustment' heuristic (see Kahneman, Slovic and Tversky, 1982), by which they are bounded away from certainty or impossibility and then adjust in relation to their previous valuation. Since the gap between the two subgroups tends to remain constant, this does not have a major effect on the rest of the results that involve the comparison of relative differences across questions. However, it does raise questions about how subjects formulate their responses to questions of this kind and suggests that there is a need for other such tests in future studies.

With regards to the CIH, the 'subjective' results do not lend much support to the hypothesis, with some differences being *higher* (not lower) for the non-comparative condition than for the comparative one. In general, the lack of any interaction effect suggests that there is no systematic difference as a result of using different treatments. In 'objective' terms, for high and middling likelihood events, the valuation differences are also in the opposite direction to those predicted by the theory; that is, the non-comparative situation produces *greater* differences between the value of risky and ambiguous gambles than the comparative situation does. For low likelihood events, the two situations produce similar differences. These results are in contrast to those presented by Fox and Tversky (1995) and Chow and Sarin's study 3 (2001) which shows strong support for the CIH but report a *greater* comparative ignorance effect when there was a high (two-thirds) probability of winning as compared to a low (one-third) probability of winning. They also render the two qualitative hypotheses in the 'Experimental Design' section effectively irrelevant, as the variation in the applicability of CIH suggested by these hypotheses are obviously not caused by source sensitivity.

The reasons for these contradictory and surprising findings are unclear. However, there is one obvious difference between the experiment as carried out by Fox and Tversky and our experiment. In study 3 of their paper, all of the participants of the experiment knew that it was possible to have both ambiguous and non-ambiguous gambles constructed from the thirty balls. Indeed Fox and Tversky explicitly state:

> this problem differs from the two-color problem because here the description of the bets. . . involves both clear and vague probabilities. Consequently, we expect some ambiguity aversion even in the non-comparative context in which each subject evaluates only one bet. However, we expect a stronger effect in a comparative context in which each subject evaluates both the clear and vague bets.

In our study, Groups 2a or 2b did not know anything about each other's gambles, and this might go some way towards explaining the difference between our results and those of Fox and Tversky. This general result of Fox and Tversky's is replicated in Chow and Sarin's study 3.

Chow and Sarin's study 4 examined the non-comparative condition only and, here, all the subjects valuing risky gambles were put in a room away from those valuing ambiguous gambles. The results show that a significant amount of ambiguity aversion was present, exceeding that in the non-comparative treatment of their study 3. This suggests that, *even though the experimenters had taken steps to avoid comparisons between gambles*, the amount of ambiguity aversion had increased. This particular experimental condition was replicated in the experiment presented here where the subjects in group 2a had no contact with the subjects in group 2b, so there was no chance of them inadvertently making comparisons between the risky and ambiguous gambles. It is noteworthy, therefore, that the results presented here are similar to those in Chow and Sarin's study 4.

The consequence of this is that, if subjects are aware of alternative gambles in both the comparative and non-comparative conditions but, in the non-comparative condition, are only asked about one type of gamble, then the Comparative Ignorance Hypothesis seems to hold. If they are unaware of these alternative gambles in the non-comparative condition, then Comparative Ignorance does not seem to hold or is not as strong as before. It seems that *awareness* of the gambles is crucial here. One can be aware of a gamble but not be asked to value it or one can be aware of a gamble and be asked to value it. What seems to come from Chow and Sarin's study 4 and from our study, is that comparative ignorance holds when one is aware of the gamble which one is not valuing in the non-comparative treatment, but does not hold when one is not aware of this gamble.

In relation to the relationship between risk and ambiguity with events of different likelihood, there is some support for 'objective' ambiguity aversion at high and middling likelihood events but no support for ambiguity loving at low likelihood events. Furthermore, there is no support for a general presumption of ambiguity aversion at all likelihood levels. From these results, we can conclude that, if Comparative Ignorance does hold, then it might be modified to take account of the likelihood of the events involved but there is no strong evidence of a Venture Theory-type relationship between ambiguity and risk evaluations. All of this leads to at least two avenues for future research. First, more research needs to be done on what causes ambiguity aversion under different knowledge states. And second, more research needs to be conducted on how differing likelihoods of events cause different levels of ambiguity aversion.

Acknowledgments

Martin Jones was supported by the Leverhulme Trust (Grant No. F/125/AF). The authors would like to thank Graham Loomes for advice on the experiment and comments on earlier versions of the paper and they are indebted to Norman Spivey for doing the programming that made the experiment possible. They are also grateful for the statistical advice and support provided by Stephen Walters and the helpful comments of the editor and an anonymous referee are appreciated.

References

Becker, S. W. and Brownson, F. O. (1964). What price ambiguity? Or the role of ambiguity in decision making. *Journal of Political Economy*, **72**, pp. 62–73.

Becker, G., DeGroot, M. and Marschak, J. (1964). Measuring utility by a single response sequential method. *Behavioral Science*, **9**, pp. 226–32.

Camerer, C. and Weber, M. (1992). Recent developments in modeling preferences: uncertainty and ambiguity. *Journal of Risk and Uncertainty*, **5**, pp. 325–70.

Chow, C. C. and Sarin, R. K. (2001). Comparative ignorance and the Ellsberg Paradox. *Journal of Risk and Uncertainty*, **22**, 2, pp. 129–39.

Cohen, M., Jaffray, J.-Y. and Said, T. (1985). Individual behavior under risk and under uncertainty: An experimental study. *Theory and Decision*, **18**, pp. 203–28.

Curley, S. P. and Yates, F. J. (1985). The center and range of the probability interval as factors affecting ambiguity preferences. *Organizational Behavior and Human Decision Processes*, **36**, pp. 273–78.

Curley, S. P. and Yates, F. J. (1989). An empirical evaluation of descriptive sources of ambiguity. *Journal of Mathematical Psychology*, **33**, pp. 397–427.

Einhorn, H. J. and Hogarth, R. M. (1986). Decision making under ambiguity. *Journal of Business*, **59**, pp.2.

Ellsberg, D. (1961). Risk, ambiguity and the savage axioms. *Quarterly Journal of Economics*, **75**, pp. 643–69.

Fox, C. R. and Tversky, A. (1995). Ambiguity aversion and comparative ignorance. *Quarterly Journal of Economics*, **110**, pp. 585–603.

Heath, C. and Tversky, A. (1991). Preference and belief: Ambiguity and competence in choice under uncertainty. *Journal of Risk and Uncertainty*, **4**, pp. 5–28.

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

Kahn, B. E. and Sarin, R. K. (1988). Modelling ambiguity in decisions under Uncertainty. *Journal of Consumer Research*, **15**, pp. 265–72.

Kahneman, D., Slovic, P. and Tversky, A. (1982). *Judgment Under Uncertainty: Heuristics and Biases*. Cambridge: Cambridge University Press.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, **47**, pp. 263–91.

Keynes, J. M. (1921). *A Treatise on Probability*. MacMillan, London.

Knight, F. H. (1921). *Risk, Uncertainty and Profit*. Houghton Mifflin, Boston.

Loomes, G. and Sugden, R. (1995). Incorporating a stochastic element into decision theories. *European Economic Review*, **39**, pp. 641–48.

Sawilowsky, S. (1990). Nonparametric tests of interaction in experimental design. *Review of Educational Research*, **60**, 1, pp. 91–126.

TVERSKY, A. and FOX, C. R. (1995). Weighing risk and uncertainty. *Psychological Review*, **102**, 2, pp. 269–83.
TVERSKY, A. and KAHNEMAN, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty*, **5**, pp. 297–323.
TVERSKY, A. and WAKKER, P. (1995). Risk attitudes and decision weights. *Econometrica*, **63**, 6, pp. 1255–80.
TVERSKY, A., SATTATH, S. and SLOVIC, P. (1988). Contingent weighting in judgment and choice. *Psychological Review*, **95**, pp. 371–84.

Date of receipt of final manuscript: October 2003