Carbone E, Dong X, Hey J.
[Elicitation of preferences under ambiguity](https://doi.org/10.1007/s11166-017-9256-0).
Journal of Risk and Uncertainty 2017
DOI: https://doi.org/10.1007/s11166-017-9256-0

DOI link to article:

https://doi.org/10.1007/s11166-017-9256-0

Date deposited:

15/06/2017

Newcastle University ePrints - eprint.ncl.ac.uk

CrossMark

Elicitation of preferences under ambiguity

Enrica Carbone[1] · Xueqi Dong[2] · John Hey[3]

Abstract This paper is about behaviour under *ambiguity*—that is, a situation in which probabilities either do not exist or are not known. Our objective is to find the most empirically valid of the increasingly large number of theories attempting to explain such behaviour. We use experimentally-generated data to compare and contrast the theories. The incentivised experimental task we employed was that of allocation: in a series of problems we gave the subjects an amount of money and asked them to allocate the money over three accounts, the payoffs to them being contingent on a 'state of the world' with the occurrence of the states being ambiguous. We reproduced ambiguity in the laboratory using a Bingo Blower. We fitted the most popular and apparently empirically valid preference functionals [Subjective Expected Utility (SEU), MaxMin Expected Utility (MEU) and α-MEU], as well as Mean-Variance (MV) and a heuristic rule, Safety First (SF). We found that SEU fits better than MV and SF and only slightly worse than MEU and α-MEU.

Keywords Allocations · Ambiguity · Preferences under ambiguity

JEL Classifications C9 · D81

✉ John Hey
    john.hey@york.ac.uk

    Enrica Carbone
    Enrica.CARBONE@unina2.it

    Xueqi Dong
    xueqi.dong@newcastle.ac.uk

[1]   Seconda Universita degli Studi di Napoli, Caserta, Italy

[2]   Newcastle University Business School at Newcastle University, Newcastle upon Tyne, UK

[3]   Department of Economics and Related Studies, University of York, Helsington, York YO10 5DD, UK

✌ Springer

1 Introduction

The context of this paper is that of decision-making under ambiguity. Ambiguity is normally considered by decision theorists to be a situation in which probabilities either do not exist or are not known. There are now an increasingly large number of theories of behaviour in such situations, and our objective is to look at a subset of these and determine which appears to be most empirically valid. To test between the theories in this subset we use experimentally generated data, asking subjects to allocate money between several accounts, the payoffs to which are ambiguous. This data allows us to fit the various theories and determine which appears to be the 'best.'

The paper is organised as follows. In Section 2 we summarise the main theories of decision-making under ambiguity, concentrating on those that we think most empirically valid and on which we shall focus. As this paper is about the elicitation of preferences, and because we use a particular elicitation method, we discuss the various alternative elicitation methods in Section 3, and compare their possible properties. In Section 4 we state the problem presented to our subjects and possible solutions to it. In Section 5 we describe our experimental implementation. We feel that this implementation is a complement to, and an extension and a refinement of, two apparently closely-related experiments; these we discuss in Section 6, looking at the differences between the various designs. Our results are reported in Section 7 and Section 8 concludes.

2 Theories of behaviour under ambiguity

There are many theories of behaviour under ambiguity. A useful survey is that of Etner et al. (2012). We shall omit a discussion of dynamic models (such as that of Siniscalchi 2009) and hence updating models. We shall also forgo the *Incomplete Preferences* story of Bewley (1986), the *Contraction* model of Gajdos et al. (2008), the *Variational* model of Maccheroni et al. (2005), and the *Confidence Function* of Chateauneuf and Faro (2009), partly because of the lack of empirical support and partly because of the difficulty of parameterising these models (these two reasons may well be related).

Historically modelling started simple. If probabilities are not known with certainty, the obvious thing to assume is that there is a range of possible probabilities, with a lower and an upper bound. A pessimist would assume that the worst could happen, and would therefore rank decisions on the basis of their worst-case outcomes—the optimal decision being the one with the least-worst outcome. This is the basis of Wald Maxmin (Bewley 1986). Later it was considered an excessively pessimistic rule and generalised by Gilboa and Schmeidler (1989) to α-Maxmin, in which decisions are based on a weighted average of the worst and best outcomes. These models worked with raw monetary payoffs.

Then came the revolution of Expected Utility theory in which outcomes are not evaluated on the basis of their monetary value, but on the *utility* of their monetary value. Two models which made the obvious generalisation of Maxmin and α-Maxmin are Maxmin Expected Utility (Gilboa and Schmeidler 1989) and α-MEU (Ghirardato et al. 2004). In both these theories the decision-maker—the DM—uses the *utility* of the outcomes.

In all the above models 'worst' and 'best' relate to possible outcomes, and cover a world in which the possible outcomes do not have probabilities attached to them, but can be ranked. But taking away probabilities is too much for most theorists. Indeed, economists who 'believe' in Subjective Expected Utility (SEU) can of course continue to assume that it holds.

At the same time, assuming that the DM believes that additive probabilities exist (and uses them) is a strong assumption, particularly in an ambiguous world. A partial softening of that strong assumption (but not interpretable as a total abandonment) is that used in Choquet Expected Utility theory (Schmeidler 1989). In this the DM is thought of as attaching *capacities* to the various outcomes, where crucially these capacities are *non-additive*—so that the capacity attached to the union of two disjoint events, $C(S_1 \cup S_2)$, is not necessarily equal to $C(S_1) + C(S_2)$. To avoid violations of dominance these capacities are associated with *ranked* payoffs. This is very similar to the procedure used in *Rank Dependent Expected Utility* theory (Kahneman and Tversky 1992) though here, weighted probabilities rather than capacities are used.

Some theorists do not like the idea of encoding the ambiguity of an event with a single number (probability or capacity or weighted probability). One route is to say that the probability of some event is not a single number but may be one number from a set of possible probabilities. Clearly this is what the α-model (and its various antecedents) is assuming, but these just work with the worst and the best from this set. A model which goes further is the Smooth Model of Klibanoff et al. (2005), which says that, if the DM cannot attach a single number to a probability, at least he or she can state the set of possible probabilities, and moreover attach probabilities to each member of the set. This is a sort of two-level probability structure, and, if the DM's preference function is linear in the probabilities, it reduces to (subjective) Expected Utility theory. For this reason Klibanoff et al. do not assume that the preference function is linear in the probabilities. We note that while this may be theoretically interesting, it is almost impossible to fit empirically—as one needs to estimate all the possible probabilities and the probabilities attached to them.

The Mean-Variance model (MV), beloved by finance theorists, does not fit neatly into the above categorisation. However, if SEU is used, for example combined with a CARA (Constant Absolute Risk Aversion) utility function, and with normally distributed outcomes, we get a decision rule consistent with MV. Unfortunately, in general, MV violates first-order stochastic dominance (Blavatskyy 2010), and as a consequence is not often used by decision theorists. Nevertheless it is a widely used decision rule in finance, and is essentially simple—relying only on a calculation of a mean and a variance of some prospect. Of course to calculate these, the DM needs to know the probabilities, or at least act as if he or she knows the probabilities.

For the various reasons discussed above, we decided to estimate SEU because of its simplicity, elegance and popularity, MaxMin Expected Utility (MEU) and its generalisation α-MEU because of their relative simplicity, and Mean-Variance (MV) because of its popularity in finance. In addition, believing that many of these theories complicate an already complex decision problem, we estimated a simple heuristic rule, Safety First (SF); we describe this later.

3 Elicitation methods

There are several methods used by economists to elicit the preference functionals of subjects in situations of uncertainty. These include *Holt-Laury Price Lists* (Holt and Laury 2002), *Pairwise Choice questions* (Hey and Orme 1994), the *Becker-DeGroot-Marschak (BDM) Mechanism* (Becker et al. 1964), the *Bomb Risk Elicitation Task* (Crosetto and Filippin 2013), and the *Allocation Method*, pioneered originally by Loomes (1991), revived by Andreoni and Miller (2002) in a social choice context, and later by Choi et al. (2007) in a risky choice context. Some of these are contrasted and compared in Loomes and Pogrebna (2014) and in Zhou and Hey (unpublished). We describe them briefly here.

In the *Holt-Laury Price List* method, while the detail may vary from application to application, the basic idea is simple: subjects are presented with an ordered list of pairwise choices and have to choose one of each pair. The list is ordered in that one of the two choices is steadily getting better or steadily getting worse as one goes through the list. Because of the ordered nature of the list, subjects *should* choose the option on one side up to a certain point, thereafter choosing the option on the other side. Some experimenters force subjects to have a unique switch point; others leave it up to subjects. The point at which the subject switches reveals their attitude to risk. Some commentators suggest that the switch point is dependent on the construction of the list.

A second method is to give a set of *Pairwise Choices,* but separately (not in a list) and not ordered. Indeed, typically the pairwise choices are presented in a random order. Some argue that this method, whilst being similar to that of Price Lists, avoids some potential biases associated with ordered lists.

A method which is elegant from a theoretical point of view is the *Becker-DeGroot-Marschak Mechanism*. The method centres on eliciting the value to a subject of a lottery—if we know the value that a subject places on a lottery with monetary outcomes, we can deduce the individual's attitude to risk over money. Let us discuss one of the two variants of this mechanism that are used in the literature—where the subject is told that they do not own the lottery, but have the right to buy it. The subject's valuation of the lottery as a potential buyer is the *maximum* price for which they would be willing to buy it. The method works as follows: the subject is asked to state a number; then a random device is activated, which produces a random number between the lowest amount in the lottery and the highest amount. If the random number is less than the stated number, then the subject buys the lottery at a price equal to the random number (and then plays out the lottery); if the random number is greater, then nothing happens and the subject stays as he or she was. If the subject's preference functional is the expected utility functional, then it can be shown that this mechanism is incentive compatible and reveals the subject's true evaluation of the lottery. The problem is that subjects do seem to have difficulty in understanding this mechanism, and a frequent criticism is that subjects understate their evaluation when acting as potential buyers and overstate it when acting as potential sellers.

In the *Bomb Risk Elicitation Task* (BRET) subjects decide how many boxes to collect out of 100, one of which contains a bomb. Earnings increase linearly with the number of boxes accumulated but are zero if the bomb is also collected. The authors claim that "The BRET requires minimal numeracy skills, avoids truncation of the data, allows the precise estimation of both risk aversion and risk seeking, and is not affected by the degree of loss aversion or by violations of the Reduction Axiom."

The *Allocation* method involves giving the subject some experimental money to allocate between various states of the world, with specified probabilities for the various states and, in some implementations, with given exchange rates between experimental money and real money for each of the states.

As we have noted above, the different methods have their advantages and disadvantages. In evaluating and comparing them there is a fundamental problem: the experimenter does not know the 'true' attitude to risk of the subjects, nor their 'true' preference functional. All we can conclude from Loomes and Pogrebna (2014) and Zhou and Hey (unpublished) is that context matters. Further work needs to be done to discover how and why. In the meantime, this paper will use the Allocation method, which is relatively under-used, and, in our opinion, relatively easy for subjects to understand. We describe below the particular allocation problem presented to our subjects.

4 The allocation problem and possible solutions

The problems presented to our subjects took the following form: the subject is given an endowment (which we normalise here to 100, as was the case in our experiment) in cash to allocate to three accounts: one with a certain return (which we normalise to 1); and the other two with uncertain returns, which depend upon which *state of nature* occurs. The number of such states is set at 3, which makes the problem a meaningful[1] one while reducing its complexity. Denote by c_1 and c_2 the allocations to the two uncertain accounts 1 and 2 respectively. This implies that the allocation to the certain account c_0 is given by $c_0 = 100 - c_1 - c_2$. Crucial to the allocation problem are the returns in the uncertain states. Denoting by r_{ij} the *absolute* return on account i if state j occurs, we have the following *returns table*:

	State 1	State 2	State 3
Account 1	r_{11}	r_{12}	r_{13}
Account 2	r_{21}	r_{22}	r_{23}

It follows that the payoff to the subject in state j, denoted by d_j, is given by $d_j = c_0 + r_{1j}c_1 + r_{2j}c_2$ $(j = 1,2,3)$.

The DM's optimal allocations depend upon his or her preferences. If we start with Expected Utility (EU) theory under risk, or Subjective Expected Utility (SEU) under ambiguity, where p_j $(j = 1,2,3)$ is the (subjective) probability of state j occurring, then the DM's objective function is the maximisation of $p_1u(d_1) + p_2u(d_2) + p_3u(d_3)$ where $u(.)$ is the individual's utility function. If instead the DM follows Mean-Variance (MV) theory using probabilities p_j $(j = 1,2,3)$, then the objective is the maximisation of $\mu - r\sigma^2$, where r indicates the attitude to risk and the mean, μ, and variance, σ^2, of the portfolio are given by $\mu = p_1d_1 + p_2d_2 + p_3d_3$ and $\sigma^2 = p_1(d_1-\mu)^2 + p_2(d_2-\mu)^2 + p_3(d_3-\mu)^2$.

---

[1] If there were just 2 states there would not be enough information in the data to allow us to infer preferences.

The above assumes that the subject works with either objective or subjective probabilities. If, however, the DM feels they are in a situation of ambiguity and hence unable to attach unique probabilities to the various states of the world, then to model his or her behaviour we need to turn to one of the new theories of behaviour under ambiguity. In this paper we work with the simplest—MaxMin Expected Utility (MEU) and the α-MEU model. Both of these theories start by assuming that, while the DM cannot attach unique probabilities to the various states, he or she works with a *set of possible probabilities*. The theories do not say how this set is specified. We assume what appears to be the simplest: this set is all possible probabilities defined by (non-negative) *lower bounds* $\underline{p}_1$, $\underline{p}_2$ and $\underline{p}_3$ (where $\underline{p}_1 + \underline{p}_2 + \underline{p}_3 \leq 1$) on the probabilities. If you like, it is a little triangle properly within the Marschak-Machina triangle.

MEU postulates that the objective function of the DM is to choose the allocation which maximises the minimum expected utility over this set of possible probabilities. The α-MEU model generalises this to maximising the weighted average of the minimum and maximum expected utility over this set. More precisely, the α-MEU model's objective function is the maximisation of $\alpha \min \left(\underline{p}_1 \leq p_1, \underline{p}_2 \leq p_2, \underline{p}_3 \leq p_3 \right)$ $[p_1 u(d_1) + p_2 u(d_2) + p_3 u(d_3)] + (1-\alpha) \max \left(\underline{p}_1 \leq p_1, \underline{p}_2 \leq p_2, \underline{p}_3 \leq p_3 \right) [p_1 u(d_1) + p_2 u(d_2) + p_3 u(d_3)]$- MEU is the special case when $\alpha = 1$.

Finally, we investigate a simple rule motivated in part by informally enquiring of the subjects how they had reached their decisions and in part by the data. We call this the Safety First (SF) rule: allocations were made first such that their payoff in *all* states would be above some threshold $\underline{w}$ and then maximising the payoff in the most likely state.[2] When fitting this model, we estimate the parameter $\underline{w}$.

5 Our experimental implementation

Subjects were presented with a total of 65[3] allocation problems, in each of which they were asked to allocate 100 in experimental cash to two accounts or to keep some of the 100 as cash. In each of these they were shown a returns table. An example is the following:

	Pink	Green	Blue
Account 1	1.7	0.9	0.6
Account 2	0	0.1	3.1

The colours represent the possible states of the world and relate to the way that ambiguity was implemented in the experiment. In the laboratory there was a Bingo Blower with pink, green and blue balls blowing around in continuous motion. Subjects could see the balls, and get a rough idea of the numbers and relative proportions of each colour, but when at the end of the experiment they were asked to eject one ball, they

---

[2] It was clear from the Bingo Blower that there were more balls of one colour than either of the other two, though the precise numbers could not be known.

[3] These problems (and the number of them) were chosen after intensive pre-experimental simulations based on results from a pilot experiment, and were chosen to maximise the power of our estimates.

could not be sure of the probability of getting a ball of a particular colour. (There were actually 10 pink, 20 green and 10 blue balls in the Blower, so the objective probabilities were 0.25, 0.5 and 0.25.) Subjects were paid on a randomly chosen problem, with their payment being determined by the payoff (given their chosen allocations) for the state implied by the colour of a ball randomly ejected from the Blower.

A screen shot from the experiment can be seen in Fig. 1[4]; the 'returns table' was called the 'Payoff Table.' The triangle shows the set of all allowable allocations; as the subject moved his or her cursor around the triangle the 'Portfolio' entries on the screen dynamically changed, and the implied payoffs for each colour were shown in the entries under 'Portfolio Payoff.' Subjects were forced to spend a minimum time of 30 seconds before registering their choice on any problem; there was a maximum time of 180 seconds per problem, and if they had not registered their choice by that time, it was taken to be an allocation of zero to the two uncertain accounts. The instructions given to the subjects, and other material related to this experiment, can be found at https://www.york.ac.uk/economics/research/centres/experimental-economics/research/unpublished/.

In the experiment we did not allow the subjects to make negative allocations (which they might have wanted to do to maximise their objective function). We enforced this rule to avoid the possibility of subjects incurring losses in the experiment. This meant that what we observe in the data are not optimal allocations, but optimal *constrained* allocations. In order to fit the various models to the data we need to compute (for any given set of parameters) the optimal constrained allocations. While explicit analytical solutions are obtainable for the optimal unconstrained allocations for some of the preference functionals, they are not easily obtained for the optimal constrained ones. As a consequence we calculate them numerically.

There was also an additional 'constraint' on the allocations that subjects could make. In the experiment, the endowment in each problem was 100, and subjects were forced to implement allocations to the nearest integer. Given the non-negativity constraint this implied a set of 5151 possible allocations. Searching over these 5151 possible allocations proved to be a more efficient method of finding the optimal constrained allocations than using some built-in function, because of the complexity of the problem.

6 Similar experiments

Before we proceed to our data analysis, we must comment on the similarities and differences between this paper and two other closely related papers, Ahn et al. (2014) and Hey and Pace (2014). Table 3 summarises the main differences and we amplify here.

We first address the comparison of our work with Ahn et al.'s. The nature of the accounts is different. Their accounts are Arrow Securities—each security only pays off in one particular state while the other two do not. That is to say for each state, there is only one security payoff. Our accounts are general, each paying off in each state of the world. Besides, one of their states of world is risky; all three of our states are

---

[4] We should note that we worded the instructions so that the decision problem represented an investment problem rather than an allocation problem, as we thought that our subjects would be more familiar with the former. This is also reflected in Figure 1.
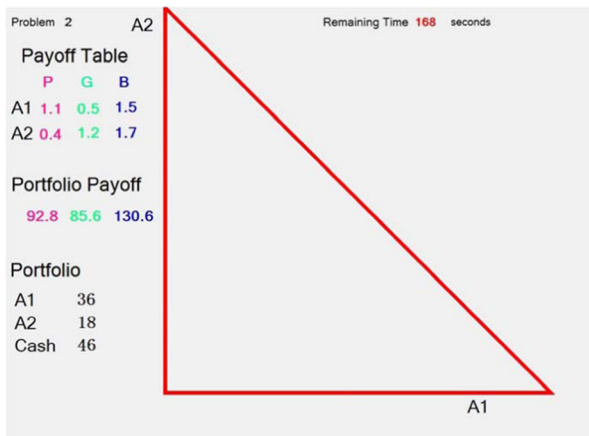
Fig. 1 A screen shot from the experiment

ambiguous. We think our setup is closer to reality. The econometric techniques differ: Ahn et al. use non-linear least squares (with an implicit normality assumption); we use maximum likelihood with what appears to be an appropriate stochastic specification. Allocation problems are the outcomes of optimisation, which is subject to high cognitive capacity and could potentially be highly noisy. We think our error specification eliminates this possible drawback associated with the allocation method; moreover we are able to estimate and interpret the magnitude of the noise. Ahn et al. implement ambiguity in the laboratory using traditional Ellsberg urns, whereas we use a Bingo Blower. The experimental interface differs: they use a three dimensional representation; we use a simpler two dimensional representation. They investigate different model specifications (kinked and smooth); we estimate particular preference functionals. The preference functionals also include a specific utility function: Ahn et al. use CARA only—we fit both CARA and Constant Relative Risk Aversion (CRRA).

As a consequence of these differences, what we can conclude naturally differs—though there is one important point of intersection. Ahn et al. write "we cannot reject SEU preferences for over 60% of subjects." As we will see, our Tables 1 and 2 point to a similar conclusion: the more general models are significantly better than SEU for a rather small proportion of subjects.

Table 1 Percentage of subjects significant using the likelihood ratio test

	Significant at 5%	Significant at 1%
(A) CARA-better group		
MEU v SEU	18%	14%
α-MEU v MEU	13%	9%
α-MEU v SEU	25%	13%
(B) CRRA-better group		
MEU v SEU	11%	11%
α-MEU v MEU	21%	11%
α-MEU v SEU	27%	11%

Table 2 Clarke tests

(A) CARA-better group

(a) Comparisons between SF, SEU, MEU and α-MEU (5% significance level)

SEU v SF			**MEU v SF**			**α-MEU v SF**		
SEU better than SF	SF better than SEU	Neither better than the other	MEU better than SF	SF better than MEU	Neither better than the other	α-MEU better than SF	SF better than α-MEU	Neither better than the other
70%	5%	25%	70%	2%	28%	70%	2%	28%

(b) Comparisons between MV and SEU, MEU and α-MEU (5% significance level)

SEU v MV			**MEU v MV**			**α-MEU v MV**		
SEU better than MV	MV better than SEU	Neither better than the other	MEU better than MV	MV better than MEU	Neither better than the other	α-MEU better than MV	MV better than α-MEU	Neither better than the other
50%	7%	43%	48%	4%	48%	52%	5%	43%

(B) CRRA-better group

(a) Comparisons between SF, SEU, MEU and α-MEU (5% significance level)

SEU v SF			**MEU v SF**			**α-MEU v SF**		
SEU better than SF	SF better than SEU	Neither better than the other	MEU better than SF	SF better than MEU	Neither better than the other	α-MEU better than SF	SF better than α-MEU	Neither better than the other
63%	0%	37%	63%	0%	37%	63%	0%	37%

(b) Comparisons between MV and SEU, MEU and α-MEU (5% significance level)

SEU v MV			**MEU v MV**			**α-MEU v MV**		
SEU better than MV	MV better than SEU	Neither better than the other	MEU better than MV	MV better than MEU	Neither better than the other	α-MEU better than MV	MV better than α-MEU	Neither better than the other
74%	0%	26%	79%	0%	21%	84%	0%	16%

We now proceed to a discussion with respect to Hey and Pace (2014). First of all, the subjects are facing different decisions in the experiments. In Hey and Pace's experiment, subjects have two types of problems. In type 1 problems, subjects can only invest in two of the three accounts. In type 2 problems subjects could invest between one account and the other two accounts. In our experiment, subjects are free to invest in all three accounts. Hey and Pace chose to implement their experiment in that particular way because it makes the analytic solution for the optima much easier. The optimisation for us is more demanding and there is no analytic solution. That is why we solve our optima numerically by grid search over the whole integer space. The experimental interface differs: their subjects use one sidebar to indicate one particular allocation while our subjects move mouse cursors—the locations of which indicate the allocations for all three accounts. The preference functionals that have been estimated are also different; besides the common ones, we particularly estimated Mean Variance preference and the Safety First heuristic rule. The preference functionals also include a specific utility function: Hey and Pace only use CRRA—we fit both CARA and CRRA.

We feel that this paper represents a complement to, and an extension and a refinement of, these closely-related papers: we focus in on the apparently empirically-relevant preference functionals, and broaden the set of utility functionals used in them; we use a potentially informatively-richer experiment, and we use appropriate econometric techniques. It can be seen as a fusion of the best parts of these two papers with significant added elements. Table 3 should make this clear.

7 Stochastic specification

(This section can be safely skipped by those mainly interested in the results.)

The object of the paper is to fit preference functionals to the experimental data and see which best explains the behaviour of the subjects. We do this subject by subject, as we believe that subjects are different. Our data are the actual allocations in each problem, denoted by x_1, x_2 and x_3 (where $x_1 + x_2 + x_3 = 100$). Each preference functional specifies, given the underlying behavioural parameters, an optimal constrained allocation on any problem. Let us denote these by x_1^*, x_2^* and x_3^*; again these add to 100. These depend upon the underlying behavioural parameters. It would be pleasing if $x_i = x_i^*$ for all i, for a particular preference functional and particular parameters, as this would enable us to identify the best preference functional. But this is unlikely to happen—the reason being, as is well-known, that subjects make errors when implementing their decisions. (An alternative explanation is that none of the preference functionals explain behaviour.)

So we need to admit the possibility of errors. We need also to model how these are generated. As both x and x^* are bounded (between 0 and 100) we proceed as follows. First we introduce new variables y and y^* which are the corresponding x's divided by 100. So $y_i = x_i/100$ and $y_i^* = x_i^*/100$ for $i = 1,2,3$. These are bounded between 0 and 1. The obvious candidate distribution is the *beta* distribution which takes values over 0 and 1. Furthermore, it seems natural to first assume that the actual allocations, whilst noisy, are not biased, so that each y_i has a mean of y_i^* (and hence that each x_i has mean x_i^*). Now a variable with a beta distribution has two parameters α and β, and the mean and variance of the variable are respectively $\alpha/(\alpha + \beta)$ and $\alpha\beta/[(\alpha + \beta)^2(\alpha + \beta + 1)]$.

Table 3 Differences between this paper and those of Ahn et al. (2014) and Hey and Pace (2014)

Topic	This paper	Ahn et al. (2014)	Hey and Pace (2014)
Econometrics	Maximum Likelihood assuming Beta with bias for two random variables	Non-linear least squares (NLLS) – normality implicit	Maximum Likelihood assuming Beta with bias for one random variable
Models	SEU, Mean-variance, MaxMin, α-MEU, Safety First	Kinked and Smooth – others mentioned in an Appendix	SEU, CEU, AEU*, VEU, COM** (* same as α-MEU **Contraction Model)
Utility functions	CARA and CRRA	CARA	CRRA
Setting	3 states of the world (colours) – all 3 ambiguous	3 states of the world – 1 risky, 2 ambiguous	3 states of the world (colours) – all 3 ambiguous
Subjects' decisions	Allocate between three accounts	Allocate between three accounts	Allocate either (1) between one account and another account or (2) between one account and the other two accounts
Accounts and returns	1 certain account – with a return of 1.00 in all 3 states of the world, and 2 ambiguous accounts – both pay off something in each state of the world. Asset prices are 1	The accounts are 3 Arrow securities – each pays 1.00 in just 1 state of the world. Asset prices are not 1.	1 certain account – with a return of 1.00 in all 3 states of the world, and 2 ambiguous accounts – both pay off something in each state of the world. Asset prices are 1
Ambiguity implementation	Bingo Blower	Subjects told that the ambiguous states 'were selected with unknown probabilities that sum to $^2/_3$'	Bingo Blower
Finding optimal allocations	Numerical search over (integer) grid.	Calculated analytically (possible because of the assumptions)	Calculated analytically (possible because of the assumptions)
Experimental interface	Visual Studio program in which allocations for given cursor position in a triangle are shown and the implications shown alongside	3-dimensional representation with planes inserted for prices	Visual Studio program in which subjects indicate preferred allocations with a slider
Number of problems	65	50	76

Italics indicate where our paper differs from *one* of the other two.

Bold italics indicate where our paper differs from **both** of them.

Taking y_1 first, if we assume that its distribution is beta with parameters $\alpha_1 = y_1*(s-1)$ and $\beta = (1-y_1*)(s-1)$, this guarantees that the mean of y_1 is y_1* and that its variance is $y_1*(1-y_1*)/s$. The parameter s here is an indicator of the *precision* of the distribution: the higher is s the more precise is the DM and the less noisy are the allocations.

Notice, however, that the variance of the distribution depends upon y_1*—the closer it is to the bounds, the smaller it is, and at the bounds it becomes zero. This implies that this distribution cannot rationalise any non-zero allocation if the optimal is zero, nor can it rationalise any observation not equal to 1 if the optimal is 1. To get round this problem, we modify our definitions of the parameters α_1 and β_1. Instead of $\alpha_1 = y_1*(s-1)$ and $\beta = (1-y_1*)(s-1)$ we postulate that $\alpha_1 = y_1'(s-1)$ and $\beta = (1-y_1')(s-1)$ where $y_1' = b/2 + (1-b)y_1*$. There is a new parameter, b, which indicates the *bias* of the actual allocation, so that now the mean of y_1 is not y_1* but instead $b/2 + (1-b)y_1*$. If b is zero then it is not biased, and as b increases the bias increases.

Now we turn to y_2. We must take into account that this must be between 0 and $1-y_1$. Hence $y_2/(1-y_1)$ is between 0 and 1. Here again a beta distribution is the natural candidate and we assume that the distribution of $y_2/(1-y_1)$ is beta with parameters α_2 and β_2 given by $\alpha_2 = y_2'(s-1)/(1-y_1)$ and $\beta_2 = (1-y_2')(s-1)/(1-y_1)$ where $y_2' = b/2 + (1-b)y_2*$. Clearly if $y_1 = 1$, this method is not applicable, and so in this case we assume that the error is made solely on y_1. In all cases the third allocation, y_3, is the residual.

Finally, in order to proceed to the likelihood function we should remember that allocations could only be made in integers. We assume that subjects rounded their intended allocations. So, for example, the likelihood of an observation of x_1 is equal to the cumulative probability from x_1–0.5 to $x_1 + 0.5$. The general form of the sum of log-likelihood function for all 65 problems can therefore be written as

$$\mathscr{L} = \sum_j^{65} log(L_1 L_2)$$

Here

$$\begin{cases} L_1 = F\left(\frac{x_1+0.5}{100}, \alpha_1, \beta_1\right) - F\left(\frac{x_1-0.5}{100}, \alpha_1, \beta_1\right) \\ L_2 = F\left(\frac{x_2+0.5}{100-x_1}, \alpha_2, \beta_2\right) - F\left(\frac{x_2-0.5}{100-x_1}, \alpha_2, \beta_2\right) \text{ when } x_1 \neq 100 \\ \qquad\qquad 1 \text{ when } x_1 = 100. \end{cases}$$

where $F(x,\alpha,\beta)$ is the *cdf* of a beta distribution with parameters α and β. These parameters are specified above.

We use Matlab to find the estimates of our parameters (which are r, s, b the underlying probabilities or the lower bounds on them), and the goodness-of-fit of the various preference functionals.

8 Results

We have explored a number of different specifications and we report here just the best. Our primary concern is about the best fitting preference functional; we start with that.

We measure the goodness-of-fit by the Maximised Log-Likelihood (MLL), but we need to correct for the number of parameters in the preference functional—the number of degrees of freedom in the estimation.

We have already mentioned the preference functionals we have fitted. Each of these involves a utility function; we have taken two utility functionals. The first is the Constant Absolute Risk Aversion (CARA) form so that utility $u(x)$ is proportional to $-e^{-rx}$. The second is the Constant Relative Risk Aversion (CRRA) form so that utility $u(x)$ is proportional to x^{1-r}. In order to compare the goodness-of-fit of the different specifications, we need to distinguish between pairs of preference functionals one of which is nested within the other, and pairs of preference functionals where neither is nested within the other. We use the *Likelihood Ratio Test* for the former and the Clarke test for the latter. We note that SEU is nested within both MEU and α-MEU and that MEU is nested within α-MEU, but that none of the other functionals are nested within any other.

We had a total of 77 subjects. We omit 2 from the analysis that follows as they were extremely risk-averse, investing nothing in either risky account.[5] We then divide the remaining 75 subjects into two groups, which we call the CARA-better group and the CRRA-better group, membership of which was determined by the value of the maximised log-likelihood. For 71 of these 75 subjects, one of CARA or CRRA had a higher log-likelihood.[6] There are 56 in the CARA-better group and 19 in the CRRA-better. We then report the results of the Likelihood Ratio and the Clarke tests for each of these groups separately.

When one model is nested within another, the test statistic is

$$\mathrm{T} = 2(\mathscr{L}_1 - \mathscr{L}_0)$$

where $\mathscr{L}_0$ is the maximised log-likelihood of the nested model and $\mathscr{L}_1$ is the maximised log-likelihood of the nesting model. The test statistic has a Chi-square distribution with degrees of freedom equal to the difference in the number of parameters in the two competing models. As α-MEU has one more parameter than MEU and as MEU has one more parameter than SEU, the corresponding degrees of freedom for SEU v MEU, SEU v α-MEU and MEU v α-MEU are 1, 2 and 1 respectively. The results are summarised in Table 1, which reports the percentage of the subjects for which the test was significant. Table 1 Panel A gives the results for the CARA-better group and Table 1 Panel B gives the results for the CRRA-better group.

As the results are similar for the two groups, we put them together and note that both MEU and α-MEU do moderately better than SEU for a small number of subjects, which may not be surprising as the decision problem was one under ambiguity rather than under risk. Nevertheless SEU performs well.

When models are *not* nested one within the other we use the *Clarke Test* (Clarke 2007). The null hypothesis is that the models are equally good, and hence on a particular problem the probability of the log-likelihood for one model being larger than the probability of the other model is ½. That is:

$$H_0: \quad P(L_1 - L_2 > 0) = 0.5$$

---

[5] All the models, with appropriate parameters, can equally well describe the behaviour of these 2 subjects.
[6] The allocation of the final 4 was done on the basis of a majority rule.

Here L_1 and L_2 are the individual log-likelihoods of the 65 problems, which are calculated using the estimated parameters of the two competing models. The test statistic is

$$T = \sum_i^{65} I_i(L_1 - L_2)$$

where

$$I_i\left(L_1 - L_2\right) = \begin{cases} 1, & L_1 - L_2 > 0 \\ 0, & L_1 - L_2 \leq 0. \end{cases}$$

Under the null hypothesis T has a binomial distribution with parameters $n = 65$ and $p = 0.5$. Thus an observation greater than 40 or less than 25 rejects the null hypothesis at the 5% significance level. The results are summarised in Table 2. These are the *percentages* for which the test was significant. Table 2 Panel A gives the results for the CARA-better group and Table 2 Panel B gives the results for the CRRA better-group.

Here there are more noticeable differences between the two groups. In a comparison between SF, SEU, MEU and α-MEU, SF does not perform too well in the CARA-better group, though it does marginally better in the CRRA-better group. In comparisons between MV, SEU, between MEU and MV and between α-MEU and MV, in the CARA-better group SEU is often significantly better than MEU and α-MEU, and very rarely is one of the more general functionals significantly better than SEU. In the CRRA-better group, SEU does even better.

As a side issue, it may be interesting to report on the estimated probabilities for SEU and the estimated lower bounds on the probabilities for MEU and α-MEU; recall that the true probabilities were 0.25 (pink), 0.5 (green) and 0.25 (blue). When the CARA utility functional is the one estimated, the averages (over all subjects) of the estimated probabilities for SEU were 0.262, 0.530 and 0.208, which are very close to the true probabilities (though there was considerable dispersion across subjects). For MEU the average lower bounds were 0.228, 0.507 and 0.190, while for α-MEU they were 0.212,
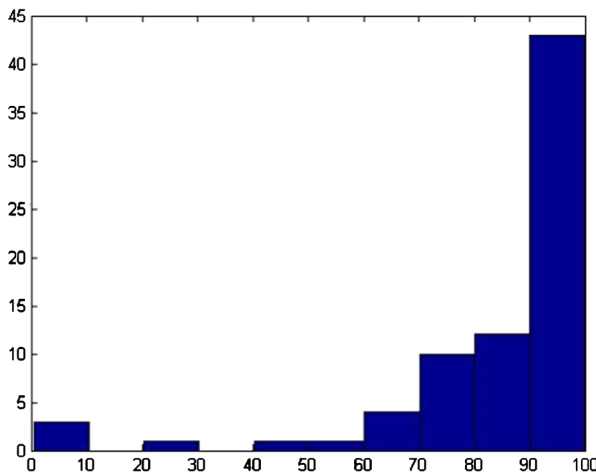

Fig. 2 The distribution of the estimated threshold $\underline{w}$ for SF

0.490 and 0.171. These are (necessarily) lower than the corresponding SEU probabilities, but only marginally so. These figures suggest that while, for some subjects, MEU or α-MEU are *statistically* superior to SEU, the economic importance is marginal. When the CARA utility functional is the one estimated, these numbers are 0.257, 0.514 and 0.229 for SEU; 0.233, 0.503 and 0.233 for MEU; 0.224, 0.462 and 0.198 for α-MEU. These are very similar to those when the CARA functional was that estimated.

While SF does not perform particularly well, it may be of interest to report the estimated values of the threshold w—the distribution is in Fig. 2. It will be seen from this that many subjects had a very high threshold—some approaching 100%. This alternatively could be interpreted as the result of very high risk-aversion, but this will of course by picked up by SEU (or MEU or α-MEU) with a high estimated level of risk-aversion.

9 Conclusions

The main conclusion from the experiment is that MV did rather badly as an explanation of behaviour, possibly as a consequence of it being a special case of SEU. In contrast SEU does rather well, not only compared to MV, but also compared with the generalisations MEU and α-MEU: for relatively few subjects do these latter preference functionals perform better. This indicates that subjects do not use a more complicated preference functional when choosing their allocations in a complicated setting. At the same time our simple rule, SF, does worse than SEU, suggesting some sophistication in subjects' decisions. Finally, it is reassuring for experimentalists that the results of Ahn et al. (2014) and Hey and Pace (2014) are confirmed by our findings, insofar as the experiments are comparable.

References

Ahn, D., Choi, S., Gale, D., & Kariv, S. (2014). Estimating ambiguity aversion in a portfolio choice experiment. *Quantitative Economics, 5*, 195–223.

Andreoni, J., & Miller, J. (2002). Giving according to GARP: an experimental test of the consistency of preferences for altruism. *Econometrica, 70*, 737–753.

Becker, G. M., DeGroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Behavioral Science, 9*, 226–231.

Bewley, T. (1986). *Knightian decision theory: Part I. Discussion paper 807*. Cowles Foundation.

Blavatskyy, P. R. (2010). Modifying the mean-variance approach to avoid violations of stochastic dominance. *Management Science, 56*, 250–257.

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

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

Clarke, K. A. (2007). A simple distribution-free test for nonnested model selection. *Political Analysis, 15*(3), 347–363. https://doi.org/10.1093/pan/mpm004.

Crosetto, P., & Filippin, A. (2013). The "bomb" risk elicitation task. *Journal of Risk and Uncertainty, 47*(1), 31–65.

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

Gajdos, T., Hayashi, T., Tallon, J.-M., & Vergnaud, J.-C. (2008). Attitude toward imprecise information. *Journal of Economic Theory, 140*, 23–56.

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

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

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

Hey, J. D., & Pace, N. (2014). The explanatory and predictive power of non two-stage-probability models of decision making under ambiguity. *Journal of Risk and Uncertainty, 49*, 1–29. doi:10.1007/s11166-014-9198-8.

Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review, 92*, 1644–1655.

Kahneman, D., & Tversky, A. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty, 5*, 297–323.

Klibanoff, P., Marinacci, M., & Mukerji, S. (2005). A smooth model of decision making under uncertainty. *Econometrica, 6*, 1849–1892.

Loomes, G. (1991). Evidence of a new violation of the independence axiom. *Journal of Risk and Uncertainty, 4*, 91–108.

Loomes, G., & Pogrebna, G. (2014). Measuring individual risk attitudes when preferences are imprecise. *The Economic Journal, 124*(576), 569–593.

Maccheroni, F., Marinacci, M., & Rustichini, A. (2005). Ambiguity aversion, robustness, and the variational representation of preferences. *Econometrica, 74*, 1447–1498.

Schmeidler, D. (1989). Subjective probability and expected utility without additivity. *Econometrica, 57*, 571–587.

Siniscalchi, M. (2009). Vector expected utility and attitudes toward variation. *Econometrica, 77*, 801–855.

Zhou, W., & Hey, J.D. (unpublished). Context matters. EXEC discussion paper: https://www.york.ac.uk/economics/research/centres/experimental-economics/research/unpublishedpapers1/#tab-5.