

# Changes in Individual and/or Aggregate?

John D. Hey<sup>a</sup> and Noemi Pace<sup>b</sup>

**Forthcoming in Applied Economics Letters**

<sup>a</sup>Department of Economics, University of York, Heslington, York YO10 5DD, UK, e-mail: john.hey@york.ac.uk;

<sup>b</sup>Department of Economics, University Ca' Foscari of Venice, Cannaregio 873, 30121 Venezia, Italy, e-mail: n.pace@unive.it.

## **Abstract**

The majority of papers reporting the results of experimental tests on individual behaviour concentrate on the average or aggregate behaviour of the subjects, even though different subjects may exhibit different kinds of behaviour. At the same time it is well-known that subjects are noisy in their responses, and thus may change their behaviour during an experiment. The question then arises as to whether this noise causes *average* behaviour to change during an experiment. We take the data from an experiment on behaviour under ambiguity, which try to identify which preference functional subjects are using, and analyse it in two halves. We find that, while some individual preferences seem to change from the first to the second half, on *average*, preferences are unchanged (though there is a reduction in the noise exhibited by subjects in the second half). This is a reassuring message for those who want to draw inferences from experiments.

**Keywords:** ambiguity, errors, noise, parameter estimates, precision, preferences.

**JEL codes:** D81

## 1. Introduction

In experiments the interest is usually in the aggregate or average behaviour of subjects; that is why most experiments have a 'large' number of subjects. At the same time it is well-known that subjects' behaviour is noisy; that is why most experiments have a 'large' number of questions. Normally it is assumed that the noise in behaviour is random, and not systematic, but it may be the case that individual preferences (or at least perceived or estimated preferences) change during an experiment, as subjects learn about the experiment or about themselves. The central issue examined in this paper is whether these changes are solely at an individual level or whether they influence the average conclusions taken from an experiment. We address this issue using data from an experiment aimed at inferring the preferences in conditions of ambiguity, where subjects are categorised in terms of the ambiguity preference functional they appear to be using.

## 2. The context

Our analysis is set in the context of an experiment designed to find the preferences of subjects under ambiguity. Ambiguity is a situation in which either objective probabilities do not exist or the decision-maker does not know them. The experiment is reported in detail in [Hey and Pace \(2014\)](#). We give here only sufficient detail to appreciate the objective of that experiment, the methodology used and the inferences that we wished to draw, and how we are using this experiment as a vehicle for consideration in this present paper.

The experiment was designed to see which of five models of behaviour under ambiguity<sup>1</sup> appear to be the most empirically appealing, both from a descriptive and a predictive point of view. From the results of that experiment two of the models can be eliminated from consideration because of their low descriptive and predictive power, so we concentrate here on the three best-fitting models: Subjective Expected Utility theory (SEU), Alpha Expected Utility (AEU) and Vector Expected Utility (VEU).<sup>2</sup> SEU is well-known; AEU models the decision-maker as perceiving a set of possible probabilities, and choosing on the basis of a weighted average of the lowest and highest expected utility over this set; VEU models the decision-maker as maximising a penalised expected utility where the penalty depends on the perceived amount of ambiguity. These three models were parameterised, the number of parameters differing from model to model: SEU (2) the perceived additive probabilities); AEU (4) the perceived lower bounds on the probabilities defining the set of possible probabilities and a parameter,  $\alpha$ , representing the weight placed by the subject on the minimum expected utility; and VEU (3) the perceived additive probabilities and a parameter,  $\delta$ ,

---

<sup>1</sup> For a survey of such models see Etner *et al* (2012).

<sup>2</sup> For details of AEU and VEU see Ghirardato, Maccheroni and Marinacci (2004) and Siniscalchi (2009) respectively.

indicating the perception of the ambiguity. All these models embody a utility function and we took a functional form (Constant Relative Risk Aversion - CRRA) with one parameter,  $r$ , indicating the risk-aversion of the subject. In addition, recognising that subjects always express any decision with some error or noise, we specified the noise in the subjects' responses by using a beta distribution, with one parameter,  $s$ , indicating the *precision* in the subject's responses.<sup>3</sup>

The data were obtained from an experiment with 76 allocation questions, each asking the subjects to allocate a given number of tokens between three ambiguous events – the colour of a ball drawn from a Bingo Blower<sup>4</sup> – with known and given exchange rates between tokens and money for each colour and for each problem.<sup>5</sup> Subjects were paid on a randomly selected question: the subject's answer to that question was recalled from the computer; a ball was drawn from the Bingo Blower; and the subject was paid the amount the subject allocated to the colour of the ball drawn multiplied by the appropriate exchange rate. The average payment was £19.56. There were two treatments which varied in the amount of ambiguity implied by the Bingo Blower: Treatment 1, in which the Blower contained just 10 balls; Treatment 2: in which it contained 40 balls. Treatment 1 is the least ambiguous treatment, as the balls of each colour could be counted; Treatment 2 is quite clearly an ambiguous situation as the balls cannot be counted. We report here the results for both treatments combined since they seem to suggest that there is no treatment effect: subjects treat the two situations as being equally ambiguous.

The data can be used to fit each preference functional by maximum likelihood, hence obtaining estimates of the parameters of each model, of the risk-aversion parameter,  $r$ , and of the precision parameter,  $s$ , for each subject. In Hey and Pace (2014) we estimated these parameters using the decisions on a subset of 60 of the questions, and used the resulting estimates to predict behaviour on the remaining 16 questions. Here we consider all the 76 questions, as our objective in this present paper is different.

### 3. The Data Analysis

We split, for each subject, the data into two halves: the first half consisting of the first 38 allocation problems that the subject tackled; the second half consisting of the last 38 questions. We then, for

---

<sup>3</sup> We actually investigated three specifications – CARA (Constant Absolute Risk Aversion) with Fechner error, CRRA with Fechner error and CRRA with Beta distributed errors - but the details need not concern us here. Suffice it to say that the one reported here, CRRA with Beta distributed errors, was the best-fitting one.

<sup>4</sup> This is a container with glass walls in which a number of balls are blown around by a jet of air. Subjects can see the balls but cannot count them (unless the total number of balls is small), but can get a rough idea of the relative proportions of the different colours. Videos of the Bingo Blower can be seen using the *Ambiguity* tab on the [ongoing projects](#) page of the [EXEC website](#). At the end of the experiment, subjects individually ejected a ball by moving the tube in the Blower.

<sup>5</sup> These exchange rates varied between the colours and between problems, thus giving us data which could be used to infer the preferences of the subjects.

each subject, fit the 3 preference functionals separately, and hence obtain estimates of the various parameters – and the maximised log-likelihood – for each half of the questions. This enables us to check as to whether the best-fitting preference functional and the estimated parameters of any given preference functional differ in the two halves, therefore enabling us to test the assertions in the Introduction, both at an individual level and at the aggregate level.

Let us start at the individual level. We first ask whether the best-fitting preference functional for individual subjects changes from the first half to the second. We rank the preference functionals by the value of the fitted Bayesian Information Criterion (BIC).<sup>6</sup> Table 1 shows the results in the form of a transition matrix. The entries are the proportion of subjects who are ranked first in the first half on a preference function and are ranked first in the second half on a preference functional. If the rankings remain constant across the two halves, there should be ‘1’s down the main diagonal and ‘0’s elsewhere. That is not the case.

**Table 1: Transition matrix**

Rankings by BIC	Model at 1st ranking in second half		
Model at 1st ranking in first half	SEU	AEU	VEU
SEU (60)	0.52	0.12	0.37
AEU (13)	0.31	0.31	0.38
VEU (56)	0.34	0.04	0.63

Note: SEU: Subjective Expected Utility; AEU: Alpha Expected Utility; VEU: Vector Expected Utility. The figures in brackets show the number of individuals for which a preference functional ranks first.

The largest numbers occur down the main diagonal, though they are not all one’s. The table is not quite symmetric – there being more moves from AEU to SEU than vice-versa, and slightly more moves from AEU to VEU than vice-versa. So there are some changes at the individual level. However if we look at the distribution of subjects over the three functionals we get Table 2.

**Table 2: Distribution of subjects over the three functionals**

First (second) half	Percentages			Average Ranking
	1st	2nd	3th	
SEU	47(42)	34(32)	19(26)	1.73(1.84)
AEU	10(10)	19(27)	71(63)	2.61(2.53)
VEU	43(48)	47(41)	9(11)	1.66(1.63)

<sup>6</sup> This is given by  $k \log(n) - 2LL$  where  $k$  is the number of estimated parameters,  $n$  the number of observations and  $LL$  the maximised log-likelihood.

In terms of the first place, SEU has moved marginally down and VEU marginally up; in terms of the average ranking the reverse is true, but the differences are small. On average preferences do not appear to change, though at the individual level there are changes.

Given any preference function, a second question needs to be asked. Do the estimated parameters change from half to half? There are three types of parameter: the risk-aversion parameter; the precision parameter; and parameters which describe the *perception* of the ambiguity by the subject. These latter vary from preference functional to preference functional: they are the estimated probabilities for SEU; the lower bounds on the probabilities and the minimum/maximum weight for AEU; and the perceived probabilities and the delta parameter for VEU.

In general these average perceptions do not change significantly from half to half; tests are given in Table 3. Apart from the average perceived probability of colour 1 in SEU, and the perceived lower bound on that probability for AEU, there are no significant changes in the averages; the true probability of colour 1 was, in fact, 0.2, so there was a slight movement towards the correct value in both cases, but the magnitude of this movement was small.<sup>7</sup>

**Table 3: Test on parameters of the models**

SEU	First half	Second half	t-stat diff
$p_1$	0.232	0.218	2.20
$p_2$	0.328	0.331	-0.49
$p_3$	0.440	0.451	-1.18
AEU	First half	Second half	t-stat diff
$\underline{p}_1$	0.206	0.192	2.15
$\underline{p}_2$	0.302	0.299	0.35
$\underline{p}_3$	0.413	0.417	-0.44
$\alpha$	0.346	0.378	-0.97
VEU	Second half	Second half	t-stat diff
$p_1$	0.232	0.221	1.65
$p_2$	0.328	0.329	-0.03
$p_3$	0.439	0.450	-1.15
$\delta$	-0.027	-0.025	-0.28

Interestingly, and importantly as far as the elicitation of the risk attitude is concerned, the average estimated value risk-aversion parameter does not change significantly from half to half; Table 4 gives the details.

<sup>7</sup> The true probabilities, defined by the proportion of balls of different colour, were  $p_1 = 0.2$ ,  $p_2 = 0.3$  and  $p_3 = 0.5$ .

**Table 4: Estimates of  $r$ , the risk aversion parameter**

	First half	Second half	t-stat diff
<b>SEU</b>	2.24	2.85	-1.91
<b>AEU</b>	2.68	3.52	-1.36
<b>VEU</b>	2.34	2.91	-1.09

One interesting parameter remains: the *precision* parameter  $s$ . Here there *are* differences in the average precision as Table 5 shows: the precision is significantly higher on average in the second half of the experiment.

**Table 5: Changes in the parameter of precision**

	First half	Second half	t-stat diff
SEU	21.89	27.21	-3.59
AEU	26.04	32.70	-3.30
VEU	25.27	32.09	-3.41

Taken together with the previous results, it tells us that, given the preference functional, preferences on average do *not* change during the experiment, and that subjects do not get noisier as the experiment proceeds: on the contrary they express their preferences more precisely.

### **Conclusions**

Our results tell us that, while there may be changes at the individual level during the experiment, the *average* findings are robust. This is one reason why experiments should use a ‘large’ number of subjects. More importantly perhaps, the precision with which preferences are expressed rises during the experiment. Nevertheless, there is noise in subjects’ behaviour, so there remains the need for a ‘large’ number of questions in order to get accurate estimates.

### **References**

Etner, J, Jeleva, M, and Tallon J M, 2012, Decision Theory under Ambiguity, *Journal of Economic Surveys*, 26, 2: 234-270.

Ghirardato, P., Maccheroni F., and Marinacci M, 2004, “Differentiating Ambiguity and Ambiguity Attitude, *Journal of Economic Theory*, 118, 2, 133-173.

Hey, J D, and Pace, N, 2014, “The Explanatory and Predictive Power of Non Two-Stage-Probability Theories of Decision Making Under Ambiguity”, forthcoming *Journal of Risk and Uncertainty*. A version is available at the EXEC website.

Siniscalchi, M, 2009, Vector Expected Utility and Attitudes Toward Variation, *Econometrica*, 77, 3, 801-855.