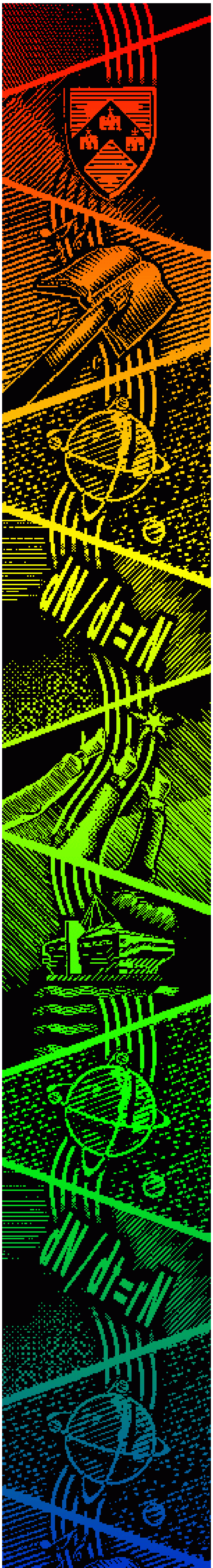


THE UNIVERSITY *of York*

Discussion Papers in Economics

No. 2008/04

The Descriptive and Predictive Adequacy of Theories of Decision Making Under Uncertainty/Ambiguity

By

John D Hey, Gianna Lotito and Anna Maffioletti

Department of Economics and Related Studies
University of York
Heslington
York, YO10 5DD

The Descriptive and Predictive Adequacy of Theories of Decision Making Under Uncertainty/Ambiguity

John D Hey[1], Gianna Lotito[2] and Anna Maffioletti[3]

Abstract

In this paper we examine the performance of theories of decision making under uncertainty/ambiguity from the perspective of their descriptive and predictive power, taking into account the relative parsimony of the various theories. To this end, we employ an innovative experimental design which enables us to reproduce ambiguity in the laboratory in a transparent and non-probabilistic way. We find that judging theories on the basis of their theoretical appeal, or on their ability to do well in testing contexts, is not the same as judging them on the basis of their explanatory and predictive power. We also find that the more elegant theoretical models do not perform as well as simple rules of thumb.

JEL Classifications: D81, C91

Keywords: Ambiguity, Bingo Blower, Choquet Expected Utility, Decision Field Theory, Decision Making, Expected Utility, Hurwicz Criterion, (Gilboa and Schmeidler) MaxMin EU, (Gilboa and Schmeidler) MaxMax EU, (Ghirardato) Alpha-Model, MaxMin, MaxMax, Minimum Regret, Prospect Theory, Uncertainty.

Corresponding Author: John D Hey, Department of Economics and Related Studies, University of York, York, YO10 5DD, UK. e-mail: jdh1@york.ac.uk

[1] LUISS, Rome, Italy and University of York, UK
[2] Università del Piemonte Orientale, Alessandria, Italy
[3] Università di Torino, Italy

1. Introduction

This paper is concerned with decision-making in situations in which probabilities are not knowable by the decision-maker. Such situations are described by economists as uncertain or ambiguous. In 1954, Savage argued that (subjective) Expected Utility theory is an appropriate theory of behaviour in such situations: given the validity of several normatively appealing axioms, he showed that people act *as if* they attached subjective probabilities to the various possible events. However, this conclusion has been frequently challenged as being as descriptively inaccurate. In particular, Ellsberg's 1961 paper has been prominent. In this, he offered his famous (hypothetical) experiment as evidence against the Savage theory. He wrote:

> Let us suppose that you confront two urns containing red and black balls, from one of which a ball will be drawn at random. To "bet on Red I" will mean that you choose to draw from Urn I; and that you will receive a prize *a* (say $100) if you draw a red ball ("if Red I occurs") and a smaller amount *b* (say, $0) if you draw a black ("if not-Red I occurs"). You have the following information. Urn I contains 100 red and black balls, but in a ratio entirely unknown to you; there may be from 0 to 100 red balls. In Urn II, you confirm that there are exactly 50 red and 50 black balls. An observer – who, let us say, is ignorant of the state of your information about the urns – sets out to measure your subjective probabilities by interrogating you as to your preferences in the following pairs of gambles:
> 1. "Which do you prefer to bet on, Red I or Black I: or are you indifferent?" That is, drawing a ball from Urn I, on which "event" do you prefer the $100 stake, red or black: or do you care?"
> 2. "Which would you prefer to bet on, Red II or Black II?"
> **3.** "Which do you prefer to bet on, Red I or Red II?"
> 4. "Which do you prefer to bet on, Black I or Black II?"

Ellsberg noted that "…if you prefer to bet on Red II rather than Red I, and Black II rather than Black I … or if you prefer to bet on Red I rather than Red II, and Black I rather than Black II … you are now in trouble with the Savage axioms." After introspecting about the likely results from conducting an experiment of this type, Ellsberg concluded that many people would be in "trouble with the Savage Axioms". Real experiments with monetary incentives have confirmed that this is indeed the case.

The Ellsberg experiment indicates that there are situations in which decision-makers do not act as if they have subjective probabilities, obeying the usual probability rules, over the various possibilities. It follows, therefore, that in such situations, (subjective) Expected Utility (EU) theory can not be a descriptively accurate account of behaviour. New theories need to be sought.

Theorists have not been slow in coming forward with alternatives, and it is to these alternatives, which we discuss shortly, and their relative superiority over EU, that this paper is addressed. More specifically, we examine how much better these other theories explain and predict behaviour than EU. To this end, we generate data from experiments and use this data to investigate empirically the various theories.

However, in contrast with previous experiments, we do not carry out statistical tests comparing the various theories. This is for two methodological reasons. The first is concerned with the nature of statistical testing. Suppose that some theory (call it GEU) is a generalisation of EU, in the sense that EU is nested within GEU. Then it is inevitably the case that GEU does not explain less of the data than EU[4]. Moreover, let us suppose that EU is not *exactly* true (which we know is the case, like with any theory); then, as long as there is noise in behaviour (which again we know to be the case), then, given enough data, we can always reject EU in favour of GEU at any level of significance. This leads us onto our second methodological point: rejection of EU in favour of some generalisation tells us that there is noise in the data and that we have enough observations for statistical significance – but it does not tell us whether the generalisation is *economically significantly* better as an explanation of behaviour. Statistical significance tells us nothing about economic significance. Nor does it tell us whether the increase in statistical predictability is worth the reduction in theoretical parsimony. Hence, rather than carrying out statistical tests comparing the various theories, we estimate the associated preference functionals, and compare their descriptive and predictive abilities, taking into account their relative parsimony.

The paper includes a further innovation – the nature of our experiment, and in particular, the way we implemented an uncertain/ambiguous situation in the laboratory. In the Ellsberg hypothetical experiment, the ambiguous/uncertain urn was represented to the hypothetical subject

[4] Indeed, it is easy to obtain a theory which explains all the data ("the decision-maker chooses something") but such a theory is valueless.

simply as one which "contains 100 red and black balls, but in a ratio entirely unknown to you; there may be from 0 to 100 red balls". Such a description has been followed in numerous real experiments. A good example is the experiment reported recently in Halevy (2007), in which the instructions included the following:

> Consider the following scenario. There are four boxes, each containing 10 chips, which can be either red or black. The composition of chips in the boxes is as follows:
> *Box 1*: Contains 5 red chips and 5 black chips.
> *Box 2*: The number of red and black chips is unknown. It could be any number between 0 red chips (and 10 black chips) and 10 red chips (and 0 black chips).
> *Box 3*: The number of red and black chips is determined as follows: one ticket is chosen from a bag containing 11 tickets with the numbers 0 to 10 written on them. The number written on the drawn ticket will determine the number of red chips in the third box. For example, if the ticket drawn is 3, then there will be 3 red chips and 7 black chips.
> *Box 4*: The composition of chips in this box is determined in a manner similar to box 3, but instead of 11 tickets in the bag, there are 2, with the numbers 0 and 10 written on them. Therefore, the box may contain either 0 red chips (and 10 black chips) or 10 red chips (and 0 black chips).

Here Box 2 is the "uncertain box", corresponding to the "uncertain Urn I" of Ellsberg.

The implementation of the uncertain box/urn in the laboratory is not straightforward, particularly with modern practices where openness and transparency are paramount. Is the above description sufficient? Should we tell the subjects more? What do we tell the subjects if they ask about Urn I/Box 2? How, in fact, do we compose Urn I/Box 2? – if, indeed, as we have to, we are going to make a drawing from it. Will we tell the subjects how we composed it? Will it be on display in the laboratory?

Some of the possible ways of composing Box 2 are those which are described as Box 3 and Box 4 above – though clearly the descriptions of Boxes 2, 3 and 4 differ one from the other, and could be perceived as different by the subjects. What is crucial about Box 2 is that probabilities are not knowable. This is the essential point. In Boxes 3 and 4, probabilities are knowable – though subjects might not know how to calculate them. But that is a different point. We would argue that Boxes 3 and 4 do not capture the essential point that the probabilities are not knowable, and also induce the subjects to think in terms of second-order probabilities.

Going back to the original Ellsberg experiment, there is an additional problem, which leads to an another reason why people might "prefer to bet on Red II rather than Red I, and Black II rather than Black I": simply that they do not trust the experimenter. If they assume that the experimenter wants to spend as little money as possible on the experiment, they will naturally try and imagine ways that the experimenter could have rigged Urn I to save money. Urn I becomes the 'suspicious urn' – and not the ambiguous/uncertain urn – whether or not there are grounds for suspicion.

To get round this problem, suspicion must be removed from Urn I; moreover drawing a ball from Urn I must be done in an open and transparent way. It follows that Urn I must be on display during the experiment. But how can this be so – in such a way that the probabilities can not be knowable? We used a *Bingo Blower*.

Bingo Blowers used to be common in the UK, though they are no longer so, having been replaced by electronic machines. Bingo Blowers are still used in the United States, however. In our case, the Bingo Blower is a rectangular-shaped, glass-sided, object some 3 feet high and 2 feet by 2 feet in horizontal section[5]. Inside the glass walls are a set of (table-tennis) balls – in *continuous motion* – being moved about by a jet of air from a fan in the base. When the time comes to eject one of the balls from the Blower, a transparent tube is tilted and a ball expelled at random up the tube through the pressure of the air created by the fan. It is all physical, there is nothing electronic and it is not manipulatable. There is no way that the identity of the ball being ejected can be selected – by either the experimenter or the subject. Moreover – and this is crucial to our design – all the balls inside the Blower can at all times be seen by people outside, but – unless the number of balls in the Blower is low – the number of balls of differing colours *can not be counted:* they are continually moving around. Hence the balls in the bingo can be seen but not counted, and the

[5] The Bingo Blower can be seen in action at http://www.luiss.it/fur2006/hey/ambiguity/1st.avi (Treatment 1), http://www.luiss.it/fur2006/hey/ambiguity/2nd.avi (Treatment 2) and http://www.luiss.it/fur2006/hey/ambiguity/3rd.avi (Treatment 3). You might need *xvid.exe* installed on your computer. You can get this latter from http://www.xvid.org/.

information available is not sufficient to calculate probabilities[6]. The probabilities exist but are unknowable. We have created a situation of genuine uncertainty/ambiguity.

2. The Theories under Investigation

This section discusses the theories of decision-making under ambiguity/uncertainty that we investigate with our experimental data. We leave technical details to Appendix 2, and confine ourselves here to a descriptive outline.

We note at the outset that there are two broad classes of theory: those that lead directly to a decision rule; and those that proceed indirectly through the use of a preference functional. In the first class, there is no explicit preference functional; in the second, the preference functional is primal – it determines the decision: the decision-maker is perceived as choosing the lottery which maximises the value of the preference functional. This distinction will be used later – when we talk about the stochastic assumptions underlying our empirical work.

We start with the preference functional that is most familiar – that of *(subjective) Expected Utility* theory. According to this theory, agents attach subjective probabilities (which satisfy the usual probability laws) to the various possible events and choose the lottery which yields the highest expected utility (where the subjective probabilities are used to compute the expectations). As we have already noted, the Ellsberg experiment casts doubt on this theory, particularly on the implication that the decision-maker acts as if he or she is attaching probabilities to the various events. As a consequence, there are now many alternative theories. In differing respects, these alternatives amend the axioms underlying the theory so that the implication of well-defined probabilities no longer follows. Different modifications lead to different theories. If, instead of the Savage Axioms, which implicitly assume a benevolent non-strategic Nature, one assumes that the decision-maker always assumes that the worst thing will happen (and hence that Nature is a malevolent opponent), we get the theory known as *MaxMin* – the decision-maker chooses the

[6] Though, of course, there *are* objective probabilities which the subjects could know if they could count the numbers of balls of different colours – but they cannot (unless the total number of balls in the Blower is small).

lottery for which the worst outcome is best. If, on the contrary, the decision-maker assumes that Nature is benevolent, we get the theory known as *MaxMax* – the decision-maker chooses the lottery for which the best outcome is best. The first of these rules is ultra-pessimistic, the second ultra-optimistic. Both of them are in our first class of theories – where the theory provides a decision rule, rather than a preference functional. A further alternative within this first class is the *Minimax Regret* criterion where the decision-maker is assumed to worry about the *regret* arising from his or her decision (rather than the utility of the payoff *per se*) and chooses that lottery for which the maximum regret is minimized. Again this is an ultra-pessimistic decision criterion. A criterion which is a mixture of pessimism and optimism is that known as the *Hurwicz Criterion* which suggests that the decision-maker chooses the lottery for which a weighted average of the worst and best outcome is maximised. This criterion belongs to the second class of theories that we have defined above, as it defines a preference functional, on the basis of which decisions are taken.

While the modern derivation of EU is axiomatic, the others described above (MaxMin, MaxMax, Minimax Regret and the Hurwicz criterion) are not; these latter criterion come from an age when axioms were considered less important than they are now. However, the spirit of the Hurwicz criterion has been resurrected in a new axiom-based theory proposed by Ghirardato *et al* (2004) which is a generalisation of the theory proposed in Gilboa and Schmeidler (1989). Following Ghirardato *et al*, we refer to this model as the Alpha model. Ghirardato *et al*'s axioms imply that, although the decision-maker does not know the true probabilities, he or she acts as if he or she believes that the true probabilities lie within some compact convex set, and that decisions are made on the basis of a weighted average of the minimum expected utility over this set and the maximum expected utility over this set. We note that, if this convex set is the whole probability space, then this Alpha theory reduces to that of Hurwicz. Note also that if the weight attached to the minimum expected utility is one (and that to the maximum expected utility is zero) then the Alpha theory reduces to that termed by Gilboa and Schmeidler (1989) the Maximin Expected Utility model (with a non-unique prior). Conversely, if the weight attached to the minimum expected utility is zero, then we get what we might term the "Gilboa and Schmeidler" Maximax Expected Utility model.

This approach leads us to think of the probabilities as not being unique, but belonging to some set. An alternative is that the probabilities themselves come from some distribution – so that the decision-maker does not know the actual probabilities but has a subjective distribution of their possible values. This is precisely what Halevy is implementing in his Boxes 3 and 4. Note that in Box 3, the probability of drawing a black chip is not known, but it could equally well be one of 0/10, 1/10, 2/10, 3/10, 4/10, 5/10, 6/10, 7/10, 8/10, 9/10 and 10/10. Similarly in Halevy's Box 4 the probability is not known – but it could equally well be either 0 or 1. So with these two boxes we have a two-stage probability tree – at the first stage the probability of a black chip is determined and at the second whether the chip is black or not.

Of course, if the decision-maker used the reduction of compound lotteries axiom, then he or she would regard Boxes 3 and 4 identical and exactly the same as Box 1 – the probability of a black chip is ½ in all three boxes. However, if the decision-maker can not, or does not, use reduction, then he or she may well regard Boxes 3 and 4 as different and both of them as different from Box 1. The issue then is how the decision-maker processes two-stage probability trees if he or she does not use reduction. Halevy (2007) has an extended discussion of this case and refers the reader to Segal (1987) and to Klibanoff *et al* (2005). Segal assumes that the decision-maker processes the compound lottery by backwardly-inducting using certainty equivalents. Of course, if the decision-maker's preferences were those of EU (and the decision-maker had a unique utility function used at all stages in the decision process) this procedure would lead to exactly the same decisions as the procedure using reduction – so Segal assumes non-EU preferences. In contrast, Klibanoff *et al* assume that the decision-maker has EU preferences on both first and second stage lotteries – but that these EU preferences differ – thus again not implying the same behaviour as a decision-maker who follows the reduction of compound lotteries axiom.

We do not investigate the Segal and Klibanoff *et al* preference functionals for the simple reason that in our experiment there is just a single stage. While Halevy, in Boxes 3 and 4, in an attempt to create ambiguity in a way congruent to the theorists, introduces two-stage lotteries, we do not – as these seem contrary to the original Ellsberg ambiguous urn. In our Bingo Blower, we have created the ambiguous (unknown probabilities) urn without the need for two-stage lotteries. Hence we do not consider those theories which rely on a two-stage lottery to create and model ambiguity[7].

A further theory which we investigate is Decision Field Theory (DFT) – proposed by Busemeyer and Townsend (1993). This could be described as a stochastic EU model with heteroscedasticity. We shall explain this in detail after we have discussed stochastic specifications in section 4. Finally we should note that we do not consider psychological theories that involve competence – such as that by Fox and Tversky (1995) – since the competence of subjects (in assessing the actual probabilities) is irrelevant in our experimental setting[8].

We do investigate what we term Prospect Theory, though our interpretation of it may not be what Kahneman and Tversky (1989) had in mind. In that original version of Prospect Theory, decision-makers were envisaged as attaching unique probabilities to the various singular events (or using 'objective' probabilities if they exist), and then weighting these by some probability-weighting function. The preference functional that emerges is the same as Expected Utility theory, in that it is a weighted average of the utilities from the lottery, but the weights do not sum to unity. It can be thought of as the preference functional of an EU person with non-additive probabilities attached to the various singular events. There is a problem with this original version of Prospect Theory, however: decision-makers using it may violate dominance. Such considerations led to the development of the model termed Cumulative Prospect Theory (CPT), where, on the same interpretation, decision-makers look as if they are attaching non-additive 'probabilities' to all the

[7] There is a second reason, on which we elaborate in the concluding section of this paper.
[8] This kind of theory is appropriate in a context where ambiguity is created in the laboratory through the use of real events – such as the temperature in New York or the value of the Japanese stock exchange index at a particular time on a particular day. Here subjects may differ in their knowledge of such events.

various events (including all possible combined events). This looks very similar to Choquet Expected Utility (CEU) theory, but there is one extension: Cumulative Prospect Theory makes a clear distinction between gains and losses with respective to some reference point. Indeed, Choquet Expected Utility theory is nested within Cumulative Prospect Theory. Moreover, in Choquet Expected Utility, decision-makers are perceived as attaching capacities to the various possible events, while in Cumulative Prospect Theory, decision-makers are perceived as attaching probabilities to the various events, which they then weight with some probability weighting function – the capacities in CEU can be interpreted as being the same as the weighted probabilities in CPT.

3. The Implementation

We now return to the detailed experimental design. We have already discussed the basic setting. Inside the Bingo Blower there are balls of different colours. Lotteries, or uncertain choices, are bets on particular colours. Lotteries are played out, or implemented, by ejecting one ball from the Bingo Blower and noting its colour. So, for example, with three different colours in the Blower, a lottery (£100, blue; -£10, pink; £10, yellow) would indicate an uncertain choice with a payoff of £100 if the ball ejected was blue, a loss of £10 if the ball ejected was pink and a payoff of £10 if the ball ejected was yellow.

Now we discuss the precise form of the experiment, and the actual decision tasks posed to the subjects in the experiment. One possibility we could have employed was to get subjects to *value* lotteries defined on the Bingo Blower, using either some kind of auction mechanism, or the Becker-Degroot-Marschak Mechanism, to give an appropriate incentive to subjects. Despite the attractions of obtaining valuations of lotteries (particularly that of the increased information value of each observations), we decided that problems with explaining and implementing the incentive mechanism outweighed these attractions. Instead, we decided to give them a set of *n* pairwise choice questions between (pairs of) lotteries, and incentivate them by choosing at random, after they had expressed their preference on each of the n questions, one of the n questions, and playing out their preferred lottery on that question for real. The reasons for this are that: pairwise choice questions are easier to explain to subjects; easier for them to understand; and less prone to problems of understanding associated with the various mechanisms for eliciting valuations[9].

It was decided that there would be balls of m different colours in the Blower, and lotteries would be defined in terms of amounts of money and an associated colour. We then had to decide on n and m. The value of m clearly determines the number of possible events: with $m = 2$ then there are just four events: $\varnothing, a, b,$ and $a \cup b$ (where a and b are the two colours); if $m = 3$ then there are 8 events: $\varnothing$, a, b, c, $a \cup b$, $a \cup c$, $b \cup c$ and $a \cup b \cup c$ (where a, b and c are the three colours). In general, with m colours, there are 2^m possible events. Clearly $m = 2$ is (relatively) uninteresting, while with m greater than 3 there are more than 15 different events. Since, with some of the models that we are going to fit, the number of parameters that need to be estimated increases linearly with the number of possible events, we needed to keep m low in order to conserve on degrees of freedom. We chose $m = 3$ and hence had 3 colours: pink, blue and yellow. We also decided that there would be three amounts of money – three possible prizes – x_1, x_2 and x_3. This implies that we need, in addition, to estimate one utility value – that of x_2 – normalising the other two to 0 and 1 respectively[10].

We now had to decide on the number n of pairwise choice questions. With 3 colours and with 3 amounts of money there are $3^3 = 27$ possible lotteries that can be composed and hence 27x26/2 = 351 possible different pairwise choice questions. Eliminating those questions in which there is stochastic (first-order) dominance, leaves us with 162 questions. This appears a lot. However, extensive simulation work before the implementation of the experiment convinced us that we needed this many observations in order to get accurate estimates. Moreover, even if we forced

[9] The use of this *random lottery incentive mechanism* has its critics. However, the paper by Hey and Lee (2005), and references therein, show that while the criticisms may be valid in theory, they are not so in practice.
[10] Except in the case of Cumulative Prospect Theory, where there is an additional parameter to estimate.

subjects to spend 30 seconds answering each question (which we did), the substantive part of the experiment lasted just 81 minutes.

Next we had to decide on the three outcomes: x_1, x_2 and x_3. We chose the numbers in the example above; -£10, £10 and £100. In addition we gave the subjects a participation fee of £10. Thus their take-away earnings at the end of the experiment were either £0, £20 or £110. Such high amounts of money were necessary to get an incentive to take the experiment seriously. For a risk-neutral subject, who knew the correct probabilities, his or her expected earnings from the experiment were £46.63 (plus the participation fee), while if this subject just answered the questions at random his or her expected earnings would be £33.33 (plus the participation fee).

Finally we had to decide the actual numbers of balls in the Bingo Blower. We decided that we did not want the same number of each colour. Moreover, we wanted different treatments in which the amount of ambiguity varied. We chose:

Treatment 1: 2 pink, 5 blue, 3 yellow;

Treatment 2: 4 pink, 10 blue, 6 yellow;

Treatment 3: 8 pink, 20 blue, 12 yellow.

In Treatment 1, it is actually possible to count the balls of each colour – so this is a situation of risk. In Treatment 2, it is just about possible to count the number of pink balls and begin to guess the number of yellow balls but it is impossible to count the number of blue balls. In Treatment 3 it is impossible to count the balls of any colour. We would say (though we cannot prove this) that the amount of ambiguity increases as we go through the treatments: it is effectively zero in Treatment 1, positive in Treatments 2 and 3 and higher in Treatment 3 than in Treatment 2.

This completes the description of the design. We recruited 48 subjects – 15 on Treatment 1, 17 on Treatment 2 and 16 on Treatment 3. As we have already noted, this was an expensive experiment: we paid out a total of £2130 – equal to £44.37 per subject. Subjects were recruited using the ORSEE (Greiner 2004) software and the experiment was conducted in the EXEC laboratory at the University of York using purpose-written software written in Visual Basic 6.

In the laboratory the Bingo Blower was on display, in action, in the middle of the room, throughout the whole of the experimental session. In addition, images of the Blower in action were projected *via* a video camera onto two big screens in the lab. Subjects were free at any stage to go closer to the Blower to examine it as much as they wanted. At the beginning of the experiment, subjects were taken into the laboratory and given written Instructions (available in Appendix 1). They were then allowed to turn to the computer, which repeated the Instructions. The experimenter then responded to any questions, and the subjects were allowed to begin answering the 162 pairwise choice questions[11]. The software was designed so that they had to spend a *minimum* of 30 seconds before they could move on to the next question (though they could take more time if they wanted). When they had answered all 162 questions, they called over an experimenter and drew a numbered ticket from a box containing tickets numbered from 1 to 162. The computer then recalled their answer to that question. At that point the subject and the experimenter went over to the Bingo Blower and expelled one ball. The colour of the ball, the question picked at random and their answer to that question determined their payment. They filled in a brief questionnaire, were paid, signed a receipt and were free to go.

4. The Stochastic Assumptions

We use maximum likelihood techniques to estimate the parameters of the theories that we are investigating. This requires us to specify the stochastic nature of our data. As we have noted earlier, we have two broad classes of theory – those that lead directly to a decision rule and those that lead indirectly to a rule through the medium of a preference functional. We discuss these in turn.

For those theories which lead directly to a decision rule we have a simple stochastic story: subjects implement the decision implied by the rule with some probability *(1-t)* and 'tremble' with probability *t*. Using the interpretation common in the literature (see Moffatt and Peters 2001), we

[11] The order of the questions and the left-right juxtaposition of the two lotteries on the screen were randomised.

assume that when the decision-maker trembles, he or she chooses amongst the available options at random. Hence with just two choices, subjects take the decision implied by the rule with probability *(1-t/2)* and take the other decision with probability *t/2*.

For those theories which imply a preference functional over the two lotteries, we assume that the decision-maker measures his or her preferences with some error. If we denote the preference function by *V(.)* and the two lotteries in any pairwise choice by L and R (Left and Right), then in the absence of any measurement error the subject would choose L(R) if $V(L) >(<) V(R)$, that is, if $V(L) – V(R) >(<) 0$.[12] With measurement error the subject chooses L(R) if $V(L) – V(R) + \varepsilon >(<) 0$ where ε represents the error in the measurement in the difference between the two lotteries. Finally we have to make an assumption about the distribution of ε – for all of the theories except one we will assume that this is normally distributed with zero mean and constant variance s^2.

This one exception is a preference functional that we have not yet described – that implied by Decision Field Theory (DFT) – see Busemeyer and Townsend (1993). This is described by them as *Random SEU Theory*. In the absence of measurement error it is exactly the same as deterministic EU theory. With error, it is as we have discussed in the paragraph above – except for the fact that the theory prescribes a *heteroscedastic* error – so that the *s* in the paragraph above is not independent of the decision task. We note that neither one of EU and DFT is nested within the other. We give a brief description of DFT here though we refer the reader to Busemeyer and Townsend (1993) for the details. Consider a choice problem between two lotteries: suppose the lottery on the left of the screen leads to outcomes x_a, x_b and x_c if respectively the colour drawn is *a, b* or *c*; similarly suppose that the lottery on the right leads to outcomes y_a, y_b and y_c if respectively the colour drawn is *a, b* or *c*. (Each of x_a, x_b , x_c y_a, y_b and y_c are one of -£10, £10 and £100.) To save notational complexity, let us use v_a, v_b, v_c, w_a, w_b and w_c (just in this paragraph) to refer to the associated utilities. According to DFT, the *attention weight* (similar to a probability) for any colour is a random variable, and the associated Expected Utility for each lottery is therefore a random

[12] Being indifferent if *V(L) = V(R).*

variable. To distinguish DFT from EU, Busemeyer and Townsend use the expression *valence* instead of Expected Utility. The valence for the left lottery is given by $V = P_a v_a + P_b v_b + P_c v_c$ and that of the right lottery $W = P_a w_a + P_b w_b + P_c w_c$ where the attention weights, P_a, P_b and P_c, are random variables centred on the individual's subjective probabilities p_a, p_b and p_c. So the key variable on which the decision will be taken, $V - W$, is a random variable with mean

$$d = (p_a v_a + p_b v_b + p_c v_c) - (p_a w_a + p_b w_b + p_c w_c)$$

and variance σ^2. This variance is crucially not constant, but, in the theory expounded by Busemeyer and Townsend, is given by the expression

$$s^2\{p_a(v_a - w_a)^2 + p_b(v_b - w_b)^2 + p_c(v_c - w_c)^2 - (V-W)^2\}.$$

It can be interpreted as the weighted variance of the difference between the utilities of the outcomes conditional on the colours. Busemeyer and Townsend call it the *variance of the valence difference.* An interesting special case is when both lotteries are certainties and one dominates the other. In this case, we have that $v_a = v_b = v_c = v$, that $w_a = w_b = w_c = w$, and that $v = w + d$. In this case it follows that the variance σ^2 is zero, so that the subject never makes an error and always chooses the dominating lottery. Note that this property is not implied by the EU specification with a homosecedastic error term – here violations of dominance are possible (though not observed in the experiment, as there are no choice problems in which one lottery dominates the other). More generally, the theory implies that the more dispersed are the outcomes for particular colours, then the more likely it is that the subject makes a mistake. Consider Figure 1, in which two pairwise choices are displayed. In the left-hand pair, the differences in the outcomes on pink and yellow are much greater than the differences in the outcomes on blue and yellow in the right-hand pair. The theory implies that, other things being equal (which is not necessarily true in this example), subjects are more likely to make an error in the left-hand choice pair than in the right-hand choice pair. The explanation for this property is based on psychological principles, which are discussed in detail in the Busemeyer and Townsend paper.

5. The Estimated Preference Functionals

We consider 12 different specifications: (subjective) Expected Utility theory (EU), Choquet Expected Utility theory (CEU), Prospect Theory (PT), Cumulative Prospect Theory (CPT), Decision Field Theory (DFT), Gilboa and Schmeidler's MaxMin Expected Utility theory (GS Min), "Gilboa and Schmeidler's" MaxMax Expected Utility theory (GS Max), Ghirardato *et al*'s Alpha theory (Alpha), MaxMin, MaxMax, Minimum Regret (MinReg) and Hurwicz's criterion (Hurwicz). It should be noted that EU is nested within PT which is nested within CEU which in turn is nested within CPT. EU and DFT are 'overlapping' specifications, to use Vuong's (1989) terminology. The GS Min and GS Max models are nested with the Alpha model and Hurwicz within the Alpha model. The other specifications are not nested inside any other specifications.

Because subjects are quite clearly different, and our intention is not to find the 'best-fitting specification' across all subjects, we fit these 12 specifications to the data subject by subject. Moreover, because we are interested in the out-of-sample predictive ability of the various fitted specifications, we estimate using three different data sets: first, using all 162 observations on each subject; second, using the first 152 of these (which vary from subject to subject as they were presented them in a randomised order), keeping 10 for out-of-sample predictions; third, using the first 142 of these, keeping 20 for out-of-sample predictions. We fitted the models to the data using GAUSS. The resulting maximised log-likelihoods are reported in Table 1.1 (all 162 observations used for estimation), Table 1.2 (152 observations used for estimation) and Table 1.3 (142 observations used for estimation). We note that full detail on the estimated parameters, and on their standard errors, are available on request.

The issue[13] now is how to evaluate and compare these maximised log-likelihoods, given that the number of parameters estimated varies from specification to specification. To be precise we have

[13] An alternative way of proceeding would be to carry out various pairwise tests between the various specifications. To be precise we could carry out likelihood ratio tests between the nested models, Voung (non-nested) tests between the non-nested models and Vuong (overlapping) tests between the overlapping models (see Vuong 1989). The problem here is that with 12 competing specifications, there is a total of 66 possible pairwise tests that we could carry out for each subject, and the results may well be conflicting. That is, we could, for example, find that specification 1 is not significantly better (or not significantly closer to the true model) than specification 2, that specification 2 is not

the following parameters, where *u* denotes the utility value of the middle outcome (see above), *s* is the standard deviation of the error term (for those decision rules which operate through a preference functional), and *t* is the tremble probability (for those decision rules which do not have an associated preference functional)

1. EU (4 estimated parameters): *u,* two probabilities (one for each of the three colours, with the third being the residual of the sum of the other two from 1), and *s.*
2. CEU (8 estimated parameters): *u*, six capacities (one on each colour separately and one on each pairwise combination of colours), which are not necessarily additive, and *s*.
3. PT (5 estimated parameters): *u,* three weighted probabilities (one for each of the three colours) which do not necessarily sum to 1, and *s.*
4. CPT (9 estimated parameters): *u, v,*[14] six weighted probabilities (one on each colour separately and one on each pairwise combination of colours) which are not necessarily additive, and *s*.
5. DFT (4 estimated parameters): *u,* two probabilities (one for each of the three colours, with the third being the residual of the other two from 1), and *s.*
6. GS Min (5 estimated parameters): *u,* three minimum probabilities (one for each of the three colours), and *s*
7. GS Max (5 estimated parameters): *u,* three maximum probabilities (one for each of the three colours), and *s.*
8. Alpha (6 estimated parameters): *u,* three bounding probabilities[15] (one for each of the three colours), *α* and *s.*

significantly better (or not significantly closer to the true model) than specification 3, but that specification 1 is significantly better (or significantly closer to the true model) than specification 3. Or we could, for example, find that specification 1 is significantly better (or significantly closer to the true model) than specification 2, that specification 2 is significantly better (or significantly closer to the true model) than specification 3, but that specification 1 is not significantly better (or not significantly closer to the true model) than specification 3. What do we then conclude?

[14] We adopt £0 as the reference point and normalize by putting *u(£0) = 0* and *u(£100) = 1*. *u* is equal to *u(£10)* and *v* is equal to *u(-£10)*.

[15] Ghirardato *et al* do not specify how the convex set should be characterised. To make the characterisation as parsimonious as possible, as well as treating the three colours symmetrically, we assumed that the convex space, in a triangle with the probability of one colour on the horizontal axis and the probability of a second colour on the vertical

9. MaxMin (1 estimated parameter): *t*.

10. MaxMax (1 estimated parameter): *t*.

11. MinReg (1 estimated parameter): *t*.

12. Hurwicz (3 estimated parameters): *u, α* and *s*.

We need to somehow correct the log-likelihoods for their degrees of freedom. One possibility is to use the Akaike Information Criterion. However, given the fact that we have many observations, and given Hansen's comments (see Hansen 1999) on the unreliability of the AIC and the better performance of the Bayesian Information Criterion (BIC) in selecting the best model, we report in Tables 2.1 through 2.3 the associated values of the BIC. An asterisk indicates the model with the highest value of the BIC[16]. Tables 3.1 and 3.2 summarise the results. We will concentrate attention on these latter tables.

First, we note that of the 48 subjects, one specification is best over all three data sets for 40 of the subjects. The remaining 8 are not clear, having one 'best' specification on two of the data sets and another on the third. To help classify these 8, we refer to Table 5, which reports the values of the BIC for the two competing specifications, as well as their predictive ability. For subject 5, both DFT and EU have the same predictive ability, and there seems no way that we can declare one as being better than the other on the basis of the BICs. For subject 12, however, GS Min seems better on the basis of predictive ability. The same is true for PT for subject 17, GS Min for subject 21, Alpha (marginally) for subject 24, and GS Min for subject 48. Subjects 32 and 46 are similar to subject 5, and it is difficult to declare an 'overall winner' for these subjects. Summarising, we get Table 3.2, from which it seems that we can dismiss CEU, CPT, MinReg and Hurwicz as plausible empirical models. The Alpha model does not perform too well, penalised by its high number of parameters, as is GS Max. What is interesting is that PT and DFT emerge as the empirically best models, though the 'rules of thumb' models, MaxMin, MaxMax (and also GS Min, which is not a

---

axis, was bounded by a vertical line, a horizontal line and a line parallel to the hypoteneuse. In estimating this model, we found the best-fitting values for these three bounds.

[16] The calculations are carried out to more than the two decimal places reported in Tables 2, and the asterisk is awarded on the basis of the unrounded values.

rule of thumb) perform almost as well as EU. There is also very slight evidence that these rules of thumb are increasingly used as ambiguity increases – going from Treatment 1 (essentially a situation of risk) through Treatment 2 to Treatment 3[17].

An alternative way to compare the performance of the various specifications is to look at their predictive abilities. Tables 4.1 through 4.3 report the percentage of correct predictions both in and out of sample. Clearly the numbers in Table 4.1 are highly correlated with the maximised log-likelihoods reported in Table 1.1, and similarly the in-sample predictive abilities reported in Tables 4.2 and 4.3 are highly correlated with the maximised log-likelihoods reported in Tables 1.2 and 1.3 respectively. Note that these predictiction numbers are not corrected for the numbers of parameters involved in the estimation.

As can be seen from these tables, the model which has the lowest value of the BIC is not necessarily the model which predicts best out of sample. When the estimation is done using the first 152 observations, the model which fits the best also is the best predictor of the remaining 10 observations for 30 of the 48 subjects. The remaining 18 subjects are listed in Table 6.1, which shows, in the second column the best fitting model (according to the BIC) and its predictive ability (the percentage of the 10 observations correctly predicted), and in the other columns the models with a better predictive ability, and their respective predictive abilities. For example, for subject 3, the model which fits the best according to the BIC is PT (see Table 3) and it correctly predicts 8 out of the 10 observations. However, MaxMin and MaxMax correctly predict all 10 while MinReg correctly predicts 9 out of the 10. Table 6.2 gives the same information when the first 142 observations are used for estimation and the remaining 20 for prediction; here it will be noted that the best-fitting model is also the best-predicting model for exactly half of the 48 subjects.

We use the data from the previous tables to produce Tables 7. In these tables (Table 7.1 for when 152 observations are used for estimation and Table 7.2 for when 142 observations are used for estimation) the model which performs the best on the estimation criterion is listed down the

[17] We would like to argue that Treatment 3 is more ambiguous than Treatment 2, but despite referring to the arguments of Einhorn and Hogarth (1985) we find ourselves unable to unambiguously prove this.

rows and the model(s) which perform the best on the prediction criterion is(are) listed along the columns. Where there are *n* models which predict equally well and best, we assign a value *1/n* to the cell entry. So the first row of Table 7.1 indicates that, when using 152 observations for estimation and 10 for prediction, EU is both the best explanatory model and the best predictor for 3 of the subjects. In 2 other cases when EU is the best explanatory model, CEU is the better predictor for one subject while for the other subject CEU, CPT and Alpha are all equally good and the best predictors and better than EU. Clearly, given the way that the table has been constructed, the sum of the row totals and the sum of the column totals are both 48 – the number of subjects. We also report in these tables an index, which is constructed by dividing the diagonal element by the row total. This gives an indication of the relative predictive ability of the model. Based on this criterion, we get a ranking of the models: DFT, MaxMax, (GS Min and MaxMin), EU, GS Max, PT and (CEU, CPT, Alpha, MinReg and Hurwicz) – the latter block having an index value of 0 – from Table 7.1, and a ranking: GS Max, MaxMax, MaxMin, DFT, EU, PT and (CEU, CPT, GS min, Alpha, MinReg and Hurwicz) – again this latter block having an index value of 0 – from Table 7.2. It is interesting to note that PT, which does well on the absolute ranking, does badly by this criterion, unlike DFT which does well on both.

It is important to note, however, that the ranking discussed in the paragraph above is a *relative* ranking. One should also take into account the *absolute* values of the diagonal elements. These give us a ranking of the models: DFT, PT, MaxMin, (EU and MaxMax), GS Min, GS Max, (CEU, CPT, Alpha, MinReg and Hurwicz) from Table 7.1; and DFT, PT, MaxMax, (EU andMaxMin), GS Max and (CEU, CPT, GS Min, Alpha, MinReg and Hurwicz) from Table 7.2. We note that DFT and PT emerge as the best models on this criterion. It is interesting to note that the more recent and more sophisticated models perform relatively poorly as judged in this light. This could be because of the relatively simple context of the experiment that we performed and the large number of parameters of the more sophisticated models.

6. Discussion and Conclusions

Table 3.2 essentially gives the bottom line. It seems that we can dismiss CEU, CPT, Minimax Regret and the Hurwicz Criterion. GS Max is best only once and Alpha only one-and-a-half times. Of the remaining models, clearly the best are PT and DFT, while EU, GS Min, MaxMin and MaxMax score almost equally. As we go from Treatment 1 through Treatment 2 to Treatment 3, we note some interesting effects[18]: there is a very slight evidence of an increase in the 'rules of thumb', MaxMin and MaxMax; and interestingly and counter-intuitively, EU seems to improve, with a corresponding decline in PT.

We should comment on these findings, and, in particular, the high success rate of PT and DFT, and the poor showing of the more sophisticated models. If we take together EU, PT, GS Min and DFT, these four models explain best the behaviour of over 35 of our 48 subjects. In a sense these four models are close cousins: they all involve the simple weighting of the utilities (by additive probabilities in the case of EU and DFT and by weighted probabilites in the case of PT; they all involve decisions made on the basis of 'expected utility' with EU and DFT differing only in that EU (as modelled here) has a homoscedastic error term while DFT has a particular heteroscedastic error term. GS Min is similar to EU but has a bias towards pessimism. It is clear that these are simple rules to apply – just comparing an average of the utilities of the two lotteries in each pairwise choice question. DFT performs better than EU, as a consequence of the assumption made about the variance of the error term: here it seems that DFT captures well the idea that error rates increase as the lotteries become more different. So DFT, while adopting the same simple rule as EU, takes into account in a formal sense the variation in behaviour. The fact that the performance of EU improves and that of PT declines as we go from Treatment 1 through Treatment 2 to Treatment 3 is interesting – and a trifle mysterious – and seems to go against our idea that ambiguity increases as we go throught the treatments. However, there are some people who would argue that ambiguity decreases, certainly from Treatment 2 to Treatment 3. In the latter, the balls

[18] Though we note that we have too few observations to draw any definite conclusions about treatment effects.

almost become a blur of colour and perhaps it is easier to see the proportions of the various colours: instead of counting the balls of different colours subjects may be simply estimating the proportions. We are not sure that we are convinced by this argument, but it is a possible explanation. In order to investigate it further, we would need to carry out an auxiliary experiment in which we try to discover directly how the subjects are perceiving the different treatments.

Perhaps a discussion of these auxiliary experiments might be of interest, as it points to a possible way forward. There are two cases to consider: one in which subjects are trying to count the *numbers* of balls of each colour, with the intention then of calculating the proportions of each colour; and one in which subjects think directly in terms of *proportions*. Let us consider the first case first. In each treatment, subjects may have subjective beliefs about the numbers of balls of each colour. This could be represented by a joint discrete probability distribution $f(n_a,n_b,n_c)$ representing the subjective probability that there are n_a balls of colour a, n_b balls of colour b and n_c balls of colour c. Alternatively it could be represented by three discrete marginal probability distributions $f_d(n_d)$ ($d = a, b, c$). The latter can be obtained from the former, though not the former from the latter. In the case of Treatment 1, where the balls can be counted, and there are 2 pink, 5 blue and 3 yellow, then if a is pink, b is blue and c is yellow, we have that $f_a(2) = 1$ and $f_a(n_a) = 0$ for $n_a \neq 2$; $f_b(5) = 1$ and $f_b(n_b) = 0$ for $n_b \neq 5$; and $f_c(3) = 1$ and $f_c(n_c) = 0$ for $n_c \neq 3$. In Treatment 2, where perhaps the pink balls (there are 4 of them) can be counted, and the yellow balls can almost be counted, we may have something like: that $f_a(4) = 1$ and $f_a(n_a) = 0$ for $n_a \neq 4$; $f_b(n_b) = 1/5$ for $n_b = 8, 9, 10, 11, 12$ and $f_b(n_b) = 0$ for $n_b \neq 8, 9, 10, 11, 12$; and $f_c(n_c) = 1/3$ for $n_c = 5, 6, 7$ and $f_c(n_c) = 0$ for $n_c \neq 5, 6, 7$. This would embody the idea that the subject is sure that there are 4 pink balls, thinks that the number of blue balls is either 8, 9 , 10, 11 or 12 (all of these being equally likely) and thinks that the number of yellow balls is either 5, 6 or 7 (all of these being equally likely). Obviously these are subjective perceptions and may vary from subject to subject. If, however, our conjecture about ambiguity increasing from Treatment 1 throught Treatment 2 to Treatment 3 is correct, we would expect that the subjective marginal distributions[19] $f_d(.)$ for $d =$ a, b and c all get more dispersed as we go through the treatments.

How do we test this directly? We would need to somehow elicit these distributions, in a way where the elicitation was not dependent on (or not influenced by) the preferences of the subjects. This seems difficult – the use of *scoring rules* assumes risk neutrality of the subjects. An alternative would be to try to elicit solely the *lower and upper bounds* on these distributions. Let us denote by $\underline{n}_d$ and $\overline{n}_d$ these lower and upper bounds (d = a, b, c). In the example, above, we have that $\underline{n}_a = 4$ $\overline{n}_a = 4$ $\underline{n}_b = 8$ $\overline{n}_b = 12$ $\underline{n}_c = 5$ $\overline{n}_c = 7$. If our conjecture about increasing ambiguity is correct, we would expect that the difference between these upper and lower bounds would relatively increase as we went through the treatments. For example,

$\underline{n}_a = 2$ $\overline{n}_a = 2$ $\underline{n}_b = 5$ $\overline{n}_b = 5$ $\underline{n}_c = 3$ $\overline{n}_c = 3$ in Treatment 1
$\underline{n}_a = 4$ $\overline{n}_a = 4$ $\underline{n}_b = 8$ $\overline{n}_b = 12$ $\underline{n}_c = 5$ $\overline{n}_c = 7$ in Treatment 2
$\underline{n}_a = 7$ $\overline{n}_a = 9$ $\underline{n}_b = 14$ $\overline{n}_b = 26$ $\underline{n}_c = 8$ $\overline{n}_c = 16$ in Treatment 3

If this was true for all subjects, we could conclude that perceived ambiguity increased through the treatments[20], though this is not saying that *objective* ambiguity increases through the treatments.

How might we elicit lower and upper bounds in an incentive compatible way that does not make any assumptions about the preference functional of the individual? Let us discuss this in the context of *proportions* rather than *numbers*. One possibility is the following: to elicit the upper bound[21] $\overline{p}$, we could ask the subject to state a number (which should turn out to be his or her upper bound), and then, if this number is less than the true proportion *P*, we give them nothing, whereas if it is greater than *P*, we we will play out a binary lottery which yields them some big prize with probability $1-\overline{p}$ and nothing with probability $\overline{p}$. This gives them an incentive to keep their stated $\overline{p}$ as low as possible but greater than the true *P*. Unfortunately, this would only work if there was

[19] Or the joint distribution *f(.,.,.)*.
[20] Though we note that increased relative differences between the lower and upper bounds is not necessarily the same as increase relative dispersion of the marginal distributions.
[21] To elicit their lower bound we could follow a similar procedure, *mutatis mutandis*.

sufficient subjective probability mass at the upper bound – for otherwise the subject would have an incentive to state a number lower than the upper bound with the objective of increasing the probability of winning the big prize. Moreover, this elication procedure would require extra experimental sessions with the same subjects. Additionally, it elicits only the lower and upper bounds of the marginal distributions, and not the whole of the marginal distributions and clearly not the joint distribution of all three colours. Also, in the context of eliciting bounds on the *numbers,* it would seem to be the case that we might have to reveal to subjects the total number of balls in the Bingo Blower.

However we should note that, rather than elicit subjects' perceptions of the possible numbers, or proportions, of balls of different colours, we would like to have some *objective* statement about the relative amounts of ambiguity in the different treatments.

Interestingly, the discussion above has led us to a world in which subjects might be thinking about second-order probability distributions. It will be recalled that in section 2, we commented that we would not be investigating theories of decision making under ambiguity (such as those of Segal (1987) and Klibanoff *el al* (2005)) which use second-order probabilities, because they did not seem appropriate to the context of our experiment. The discussion in the paragraphs above suggest that they may in fact be appropriate. But there is a more serious objection (and probably our real reason for us not considering these types of models): if one follows the route of these two-stage probability distributions, we would then have to estimate the first-stage probability distributions (in the absence of an independent procedure to elicit them). Without any restriction on the form of these first-stage distributions, we have a serious problem with degrees of freedom. We would have to estimate either the (subjective) joint distribution, denoted $f(.,.,.)$ above, or the three marginal distributions $f_a(.)$,$f_b(.)$ and $f_c(.)$. The number of extra parameters to be estimated is high: in Treatment 3 (even if we make the bold assumption that the subject can count the total number of balls), the number of extra parameters is equal to the number of balls of each colour that are possible – probably a number in excess of 8+20+12 = 40. Given that the more sophisticated models that we have already considered fare badly, probably because of their extra parameters, we would expect that these models based on second-order probability distributions would do considerably worse.

We suspect that this, and the fact that CEU and CPT do badly, is probably a consequence of the fact that subjects are trying to simplify a complex problem. In this context, subjects can not (and seem to not) use sophisticated rules with many parameters. Instead they use simple rules (like EU, PT and DFT which involve only the calculation of one utility value and two or three probabilities). Or a simple pessimistic rule – GS Min. Or they use even simpler rules – such as MaxMin and MaxMax. At the end of the day, it would seem that over 45 of our 48 subjects are using such simple rules. Sophistication is not possible in this context.

References

Busemeyer J and Townsend J (1993), "Decision Field Theory: A Dynamic-Cognitive Approach to Decision Making in an Uncertain Environment", *Psychological Review*, 100, 432-59.

Choquet G (1955), "Theory of Capacities", *Ann. Inst. Fourier* (Grenoble), 5, 131-295.

Einhorn H J and Hogarth R M (1985), "Ambiguity and Uncertainty in Probabilistic Inference", *Psychological Review*, 92, 433-461.

Ellsberg D (1961), "Risk, Ambiguity and the Savage Axioms", *Quarterly Journal of Economics,* 75, 643-669.

Fox C R and Tversky A (1995). "Ambiguity Aversion and Comparative Ignorance", *Quarterly Journal of Economics, 110,* 585-603.

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

Gilboa I and Schmeidler D (1989), "Maxmin Expected Utility with a Non-Unique Prior", *Journal of Mathematical Economics*, 18, 141-153.

Greiner B (2004), "The Online Recruitment System ORSEE 2.0 – A Guide for the Organization of Experiments in Economics", University of Cologne Discussion Paper (www.orsee.org).

Halevy Y (2007), "Ellesberg Revisited: An Experimental Study", *Econometrica*, 75, 503-536.

Hansen B E (1999), "Discussion of 'Data mining reconsidered'", *Econometrics Journal*, 2, 192-201.

Hey J D and Lee J, "Do Subjects Separate (or Are They Sophisticated)?" *Experimental Economics*, 8, 233-265, 2005.

Kahneman D and Tversky A (1979) "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, 47, 263-291.

Klibanoff P, Marinacci M and Mukerji S (2005), "A Smooth Model of Decision Making under Ambiguity", *Econometrica*, 73, 1849-1892.

Luce R D and Raiffa H (1957), *Games and Decisions*, Wiley: New York.

Moffatt P G and Peters S A (2001), "Testing for the presence of a tremble in economic experiments", *Experimental Economics*, 4, 221-228.

Savage L J (1954), *Foundations of Statistics*, Wiley: New York.

Schmeidler D (1989), "Subjective Probability and Expected Utility Without Additivity", *Econometrica*, 57, 571-587.

Segal U (1987), "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach", *International Economic Review*, 28, 175-202.

Tversky A and Fox C R (1995)."Weighing Risk and Uncertainty", *Psychological Review*, 102, 269-283.

Vuong Q (1989), "Likelihood ratio tests for model selection and non-nested hypotheses", *Econometrica*, 57, 307–333.

Table 1.1: Maximised Log-Likelihoods - Number of observations used for estimation 162

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	-16.39	-3.86	-13.54	-3.86	-13.40	-25.67	-14.14	-11.81	-35.17	-78.20	-80.24	-70.88
2	-31.93	-25.06	-26.31	-25.06	-35.06	-30.41	-26.31	-25.98	-69.96	-48.36	-54.85	-93.90
3	-26.20	-16.16	-22.38	-15.78	-25.64	-25.99	-22.38	-22.37	-72.27	-72.27	-79.28	-102.37
4	-24.08	-14.21	-24.06	-14.21	-25.83	-24.06	-16.37	-16.37	-71.14	-42.16	-53.01	-76.19
5	-6.27	-3.95	-6.27	-3.95	-6.24	-6.27	-6.27	-5.99	-52.13	-85.15	-87.86	-74.86
6	-17.18	-8.85	-12.83	-8.82	-20.64	-12.83	-16.31	-12.83	-10.32	-67.49	-67.55	-69.99
7	-38.00	-30.18	-31.09	-29.45	-37.78	-33.07	-31.09	-30.15	-53.92	-37.60	-49.01	-94.27
8	-27.44	-21.07	-22.04	-20.56	-28.77	-26.56	-22.04	-21.15	-71.14	-75.41	-82.04	-101.89
9	-26.52	-12.15	-16.17	-12.15	-28.94	-26.52	-16.17	-16.17	-64.83	-79.87	-83.70	-97.59
10	-33.32	-30.54	-32.73	-30.53	-31.64	-33.21	-32.73	-31.88	-55.64	-69.96	-72.74	-92.87
11	-33.21	-23.60	-25.75	-20.98	-34.75	-29.22	-25.75	-23.63	-53.92	-68.75	-75.08	-98.30
12	-41.77	-32.55	-36.88	-30.06	-42.22	-36.82	-36.88	-35.03	-58.92	-76.38	-82.89	-97.04
13	-16.80	-15.98	-16.81	-16.81	-8.87	-16.57	-16.81	-16.21	-51.98	-84.68	-86.82	-74.86
14	-30.48	-20.23	-25.99	-19.85	-30.54	-24.39	-26.00	-22.16	-77.31	-63.43	-70.24	-100.34
15	-30.26	-25.83	-29.81	-24.05	-32.30	-35.57	-29.21	-29.21	-37.60	-79.05	-81.16	-76.72
16	-69.01	-65.65	-68.57	-65.65	-70.01	-69.01	-69.01	-69.01	-88.29	-84.03	-87.25	-106.89
17	-17.91	-10.87	-12.74	-10.87	-17.16	-15.22	-12.74	-11.07	-48.36	-72.27	-76.19	-92.52
18	-29.59	-16.09	-17.78	-15.42	-29.79	-27.56	-17.78	-16.55	-68.75	-66.18	-72.74	-102.07
19	-21.64	-17.80	-21.62	-17.80	-19.46	-19.24	-21.64	-19.24	-46.24	-75.09	-77.98	-86.40
20	-24.63	-23.01	-24.34	-22.66	-23.01	-23.24	-24.34	-23.19	-37.60	-75.41	-78.29	-83.04
21	-47.05	-35.70	-41.57	-35.26	-49.78	-39.97	-41.57	-38.85	-61.98	-72.27	-79.28	-101.79
22	-53.73	-49.25	-50.00	-48.89	-54.81	-52.57	-50.00	-49.53	-65.95	-65.95	-73.66	-101.84
23	-28.33	-27.47	-28.04	-27.47	-30.47	-27.83	-28.04	-27.79	-17.80	-65.95	-65.93	-80.38
24	-54.65	-44.46	-49.79	-44.46	-57.35	-49.57	-49.79	-46.18	-65.95	-67.24	-74.80	-104.30
25	0.00	0.00	0.00	0.00	0.00	0.00	-1.39	0.00	-84.03	-48.36	-60.01	-74.67
26	-22.42	-21.56	-22.34	-19.88	-21.91	-21.84	-22.34	-21.83	-46.37	-71.14	-73.93	-85.96
27	-35.53	-29.22	-32.82	-28.59	-33.69	-32.84	-32.82	-31.29	-60.47	-82.11	-85.93	-96.36
28	-32.16	-28.97	-30.16	-26.72	-29.46	-30.78	-30.16	-29.89	-58.92	-85.15	-87.25	-89.12
29	-35.83	-27.13	-33.02	-27.13	-36.27	-31.97	-32.65	-30.89	-79.50	-69.70	-75.90	-102.06
30	-53.13	-48.15	-52.88	-48.15	-56.29	-53.13	-52.88	-52.44	-66.74	-45.99	-56.14	-87.27
31	-31.26	-26.64	-27.82	-26.64	-34.03	-30.23	-27.82	-26.71	-61.98	-76.38	-80.24	-99.02
32	-37.35	-19.50	-20.47	-19.46	-36.73	-34.97	-20.80	-19.85	-63.43	-53.92	-63.18	-98.99
33	-22.35	-17.07	-22.05	-16.28	-27.21	-30.69	-22.05	-20.19	-58.92	-72.27	-76.19	-90.54
34	-29.79	-23.97	-27.31	-23.97	-28.55	-29.38	-29.79	-27.31	-32.64	-75.41	-78.29	-76.32
35	-28.80	-23.08	-28.78	-23.08	-28.97	-27.54	-28.78	-27.54	-55.64	-35.17	-49.14	-85.98
36	-36.08	-22.43	-32.25	-21.31	-36.69	-36.09	-34.00	-34.00	-32.57	-55.48	-56.47	-83.34
37	-27.40	-18.62	-21.40	-18.62	-23.83	-35.50	-26.67	-21.11	-50.28	-82.11	-83.70	-72.48
38	-25.36	-24.30	-25.31	-24.30	-25.70	-25.31	-25.24	-25.19	-78.20	-84.03	-87.86	-99.75
39	-26.48	-20.57	-21.33	-20.57	-22.51	-21.31	-21.33	-20.94	-69.96	0.00	-18.02	-67.57
40	-13.80	-12.22	-13.80	-11.59	-9.39	-13.08	-13.80	-13.08	-17.83	-74.40	-76.19	-70.88
41	-6.65	-3.98	-6.47	-2.19	-7.72	-110.04	-6.47	-6.47	-69.96	0.00	-21.39	-67.57
42	-26.98	-15.93	-16.82	-15.93	-24.85	-25.70	-16.82	-16.19	-61.98	-79.87	-83.70	-96.67
43	-33.45	-12.33	-13.28	-8.63	-33.20	-26.64	-13.28	-12.16	-52.13	-39.92	-49.14	-97.08
44	-25.07	-24.74	-25.05	-24.67	-25.73	-25.05	-25.05	-25.03	-80.65	-50.28	-60.01	-81.91
45	-4.38	-3.90	-4.38	-3.82	-2.50	-4.38	-4.18	-4.18	-50.28	-84.61	-87.25	-74.86
46	-39.89	-37.39	-37.81	-35.89	-40.37	-39.12	-37.81	-37.40	-74.40	-58.92	-65.71	-96.46
47	-26.76	-18.18	-20.58	-18.18	-25.58	-23.30	-20.58	-19.61	-42.16	-66.18	-72.74	-92.97
48	-22.67	-16.88	-22.20	-17.19	-22.28	-17.98	-22.20	-17.32	-77.31	-79.05	-84.48	-102.37

Table 1.2: Maximised Log-Likelihoods - Number of observations used for estimation 152

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	-16.38	-3.86	-13.51	-3.86	-13.38	-25.41	-14.13	-11.76	-34.51	-72.78	-73.47	-65.57
2	-29.83	-23.74	-25.25	-23.70	-33.10	-27.73	-25.25	-24.56	-66.63	-45.35	-53.43	-89.90
3	-22.97	-14.15	-18.51	-13.98	-22.05	-22.76	-18.51	-18.49	-70.20	-70.20	-75.72	-97.33
4	-22.57	-10.62	-22.50	-10.65	-23.90	-22.57	-15.40	-15.40	-65.46	-36.86	-48.09	-71.21
5	-6.27	-3.95	-6.27	-3.95	-6.24	-6.27	-5.99	-5.99	-48.96	-79.88	-81.68	-71.39
6	-17.13	-8.81	-12.83	-8.81	-20.54	-12.83	-16.31	-12.83	-10.16	-62.47	-63.91	-67.05
7	-36.80	-27.67	-28.66	-27.06	-36.60	-31.51	-28.66	-28.66	-48.96	-36.77	-45.74	-90.36
8	-26.91	-21.01	-21.92	-20.53	-28.28	-26.06	-21.92	-21.77	-66.63	-69.90	-76.28	-94.24
9	-25.16	-12.15	-14.82	-12.15	-27.34	-25.16	-18.18	-15.04	-60.27	-74.50	-77.96	-92.78
10	-29.84	-24.18	-27.48	-24.18	-27.80	-27.68	-27.48	-25.15	-47.26	-64.23	-66.99	-86.44
11	-32.56	-23.12	-25.16	-20.51	-34.09	-28.79	-25.16	-23.16	-50.89	-62.96	-69.16	-92.14
12	-40.29	-30.66	-34.70	-28.67	-40.91	-36.37	-34.70	-33.13	-57.57	-70.20	-76.97	-93.08
13	-15.91	-15.10	-15.91	-15.26	-8.69	-15.71	-15.91	-15.26	-49.11	-79.86	-81.94	-70.70
14	-29.90	-20.17	-25.66	-19.82	-29.95	-24.21	-25.67	-22.01	-67.19	-56.97	-61.14	-93.15
15	-28.33	-24.79	-27.97	-23.20	-29.71	-32.96	-27.89	-27.89	-32.12	-74.01	-76.05	-72.04
16	-67.36	-64.21	-67.01	-64.26	-68.53	-67.36	-67.32	-67.32	-84.34	-80.21	-83.43	-100.62
17	-16.34	-10.80	-12.58	-10.80	-14.91	-14.35	-26.63	-10.98	-45.35	-65.46	-69.43	-86.21
18	-29.09	-15.69	-17.33	-15.14	-29.34	-26.99	-17.33	-16.18	-64.48	-64.48	-70.86	-97.18
19	-19.19	-15.76	-19.12	-15.76	-17.28	-17.68	-19.19	-17.68	-41.28	-69.90	-72.76	-79.36
20	-24.60	-22.99	-24.32	-22.62	-22.99	-23.22	-24.32	-23.17	-36.96	-71.22	-72.76	-78.49
21	-45.45	-29.89	-39.46	-29.89	-47.82	-37.78	-39.46	-36.97	-58.85	-68.85	-75.72	-97.12
22	-50.43	-45.11	-45.78	-45.11	-51.74	-49.65	-45.78	-45.66	-61.17	-61.17	-68.89	-95.78
23	-26.92	-25.60	-26.54	-25.37	-29.00	-26.10	-26.54	-26.05	-17.54	-62.72	-62.55	-77.58
24	-49.44	-39.37	-43.03	-39.37	-51.44	-45.94	-40.70	-39.81	-62.72	-62.72	-70.30	-97.79
25	0.00	0.00	0.00	0.00	0.00	0.00	-1.39	0.00	-79.31	-47.26	-58.35	-70.65
26	-22.09	-21.06	-22.01	-19.64	-21.55	-21.51	-22.01	-21.51	-45.35	-66.63	-70.58	-82.64
27	-34.84	-29.00	-32.58	-28.43	-32.95	-32.17	-32.58	-30.95	-56.97	-75.27	-79.09	-90.58
28	-27.91	-25.78	-26.83	-20.68	-25.83	-26.85	-26.83	-26.52	-52.60	-79.31	-81.29	-80.40
29	-35.00	-25.51	-32.13	-25.51	-34.92	-31.11	-31.62	-30.13	-76.06	-65.46	-72.76	-96.13
30	-47.12	-43.22	-45.79	-43.22	-49.98	-47.12	-45.79	-44.96	-62.94	-40.70	-51.05	-84.45
31	-30.22	-24.17	-26.57	-24.17	-33.00	-28.42	-28.27	-24.70	-60.27	-71.86	-74.45	-94.85
32	-35.66	-19.10	-20.10	-19.05	-35.34	-33.62	-20.49	-19.51	-60.49	-49.26	-56.95	-92.83
33	-21.33	-17.04	-21.02	-16.15	-26.05	-28.76	-21.02	-19.96	-55.84	-66.63	-70.58	-86.01
34	-28.62	-22.91	-26.24	-22.91	-27.67	-28.22	-28.62	-26.24	-32.12	-72.19	-75.09	-71.59
35	-27.22	-21.51	-27.20	-21.51	-27.38	-25.97	-27.20	-25.97	-51.36	-34.76	-46.37	-78.58
36	-32.39	-17.35	-28.58	-15.83	-33.15	-32.39	-28.58	-28.58	-32.20	-54.61	-55.64	-76.94
37	-24.99	-18.28	-20.26	-18.28	-19.78	-20.26	-24.47	-20.12	-45.35	-77.47	-79.12	-67.11
38	-23.93	-23.28	-23.93	-23.28	-24.28	-23.93	-23.93	-23.93	-72.44	-78.29	-82.12	-94.56
39	-21.09	-20.30	-21.02	-20.30	-21.96	-21.00	-21.02	-20.71	-63.98	0.00	-14.18	-63.68
40	-11.65	-10.84	-11.65	-10.09	-7.97	-11.49	-11.65	-11.49	-17.50	-68.26	-71.10	-66.53
41	-5.99	-3.98	-5.94	-2.39	-6.63	-37.95	-5.94	-5.94	-65.19	0.00	-17.70	-64.15
42	-26.67	-15.80	-16.65	-15.97	-24.64	-25.47	-16.65	-16.00	-60.71	-74.36	-79.05	-90.48
43	-32.30	-8.13	-10.41	-6.02	-32.20	-85.69	-10.41	-10.41	-50.89	-39.11	-47.96	-92.02
44	-24.36	-23.94	-24.29	-23.94	-25.12	-24.29	-24.29	-24.24	-76.83	-47.54	-57.13	-76.31
45	-4.38	-3.59	-4.38	-3.28	-2.50	-4.38	-3.92	-3.92	-47.26	-78.12	-80.60	-69.10
46	-31.28	-28.99	-29.59	-27.71	-29.54	-31.03	-29.59	-29.52	-71.22	-52.77	-59.88	-87.87
47	-25.11	-17.89	-19.69	-17.89	-23.79	-22.53	-19.69	-19.26	-36.86	-61.64	-68.23	-85.95
48	-21.21	-15.99	-20.79	-16.12	-20.88	-17.43	-20.79	-16.68	-71.53	-71.53	-77.96	-95.70

Table 1.3: Maximised Log-Likelihoods - Number of observations used for estimation 142

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	-15.59	-3.63	-12.96	-3.63	-12.98	-23.15	-13.29	-11.22	-34.07	-71.03	-70.58	-59.71
2	-28.05	-22.89	-24.50	-22.78	-31.28	-25.87	-24.50	-23.63	-64.06	-42.20	-50.06	-84.20
3	-20.52	-12.20	-15.47	-11.92	-19.01	-20.34	-15.47	-15.42	-65.09	-65.09	-70.51	-91.06
4	-21.36	-10.52	-21.29	-10.54	-22.99	-21.36	-14.99	-14.99	-61.86	-33.70	-44.94	-69.09
5	-5.70	-3.95	-5.70	-4.57	-6.22	-5.70	-5.46	-5.46	-47.85	-75.07	-77.40	-66.42
6	-16.48	-8.54	-12.46	-8.54	-19.67	-12.46	-15.77	-12.46	-10.03	-60.69	-62.23	-62.87
7	-32.78	-24.46	-25.58	-24.02	-32.46	-28.01	-25.58	-24.62	-48.02	-33.88	-40.70	-83.86
8	-24.52	-20.03	-20.88	-19.80	-25.81	-23.50	-20.88	-20.63	-59.72	-65.40	-70.86	-85.78
9	-24.13	-11.57	-13.75	-11.57	-26.49	-26.31	-13.75	-13.75	-59.14	-71.85	-74.92	-86.52
10	-29.02	-23.62	-26.77	-23.62	-26.70	-27.11	-26.77	-24.63	-46.24	-62.41	-64.93	-81.39
11	-29.75	-22.03	-23.19	-19.70	-31.56	-26.67	-23.19	-22.04	-46.24	-59.97	-65.80	-85.63
12	-36.31	-27.19	-30.13	-25.18	-37.12	-31.99	-30.13	-28.87	-51.36	-64.65	-71.53	-87.25
13	-14.34	-13.64	-14.34	-14.34	-8.53	-14.34	-14.34	-13.64	-48.02	-76.03	-78.29	-66.54
14	-29.27	-18.32	-25.12	-18.38	-29.55	-22.69	-25.13	-20.76	-63.45	-55.04	-59.14	-88.16
15	-28.01	-24.31	-27.68	-23.11	-29.16	-31.38	-27.64	-27.64	-31.65	-69.27	-70.90	-67.27
16	-62.64	-58.29	-61.91	-58.08	-63.21	-62.64	-62.56	-62.56	-78.98	-76.51	-78.92	-94.92
17	-16.33	-10.80	-12.57	-10.80	-14.91	-14.34	-26.46	-10.96	-44.10	-61.86	-65.80	-81.38
18	-27.52	-13.83	-16.14	-13.60	-27.70	-25.52	-16.14	-14.49	-59.97	-61.21	-67.48	-91.26
19	-14.00	-13.10	-13.91	-13.10	-13.06	-13.46	-14.00	-13.46	-38.25	-63.27	-66.90	-73.85
20	-21.55	-20.04	-21.36	-20.04	-19.95	-20.65	-21.36	-20.61	-28.82	-62.98	-65.52	-73.07
21	-41.11	-28.45	-35.81	-28.45	-43.32	-35.55	-35.81	-33.85	-52.55	-62.98	-69.93	-89.88
22	-48.43	-43.76	-44.12	-43.76	-49.66	-48.00	-44.12	-44.10	-56.87	-58.19	-65.80	-88.29
23	-26.13	-24.85	-25.90	-24.63	-28.06	-25.24	-25.90	-25.24	-17.26	-58.19	-56.34	-73.09
24	-44.89	-36.69	-39.98	-36.69	-46.30	-42.24	-39.98	-36.94	-60.21	-57.55	-65.46	-89.98
25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-72.39	-42.20	-53.51	-67.11
26	-19.82	-17.58	-19.73	-17.09	-18.96	-18.38	-19.73	-18.38	-44.39	-64.65	-67.19	-76.40
27	-32.20	-25.61	-30.01	-25.14	-30.79	-29.24	-30.01	-28.03	-50.81	-68.39	-71.39	-84.56
28	-27.23	-25.11	-26.10	-19.37	-25.49	-26.23	-26.10	-25.84	-51.55	-74.91	-76.38	-75.42
29	-31.44	-20.06	-28.89	-20.06	-32.16	-27.33	-28.46	-26.59	-71.06	-59.47	-66.60	-90.20
30	-33.48	-26.75	-28.42	-26.75	-31.32	-33.48	-28.61	-28.61	-60.44	-35.44	-46.09	-78.28
31	-26.84	-20.17	-22.39	-19.92	-30.31	-25.98	-23.22	-20.69	-58.91	-68.93	-71.19	-88.57
32	-31.20	-10.21	-13.98	-10.21	-31.07	-24.18	-14.00	-11.41	-51.18	-42.33	-50.40	-86.90
33	-21.13	-16.68	-20.75	-15.74	-25.23	-28.46	-19.72	-19.72	-52.55	-61.86	-65.80	-80.85
34	-27.90	-22.41	-25.51	-22.41	-26.94	-27.40	-27.90	-25.51	-31.49	-68.59	-71.19	-67.54
35	-25.12	-19.59	-25.03	-19.39	-25.09	-24.44	-25.12	-24.44	-48.02	-31.49	-43.10	-70.97
36	-28.54	-15.01	-25.18	-14.68	-29.10	-28.54	-25.92	-25.18	-26.44	-53.32	-52.60	-71.25
37	-23.06	-17.45	-19.09	-17.45	-18.60	-19.09	-22.64	-18.93	-40.70	-74.73	-75.30	-61.42
38	-21.57	-20.91	-21.49	-20.91	-21.86	-21.31	-21.57	-21.31	-65.74	-71.96	-75.78	-87.54
39	-20.15	-19.58	-20.13	-19.58	-20.76	-20.13	-20.14	-20.07	-61.86	0.00	-13.98	-59.99
40	-11.65	-10.79	-11.65	-10.09	-7.97	-11.49	-11.65	-11.49	-17.18	-63.45	-66.31	-63.38
41	-5.68	-3.47	-5.64	-1.04	-6.43	-77.07	-5.64	-5.64	-60.69	0.00	-17.43	-59.20
42	-25.79	-15.15	-16.29	-15.15	-23.72	-24.52	-16.29	-15.53	-57.55	-69.82	-74.50	-84.91
43	-31.86	-8.13	-10.41	-5.97	-31.62	-120.61	-10.41	-9.36	-49.55	-38.25	-44.67	-87.53
44	-21.95	-21.42	-21.81	-21.39	-23.17	-21.81	-21.72	-21.64	-69.82	-42.59	-50.73	-70.32
45	-4.38	-3.62	-4.38	-3.27	-2.50	-4.38	-3.92	-3.92	-46.24	-73.87	-76.38	-64.33
46	-29.81	-28.06	-28.68	-27.54	-28.04	-29.73	-28.68	-28.67	-66.07	-47.69	-54.86	-80.62
47	-21.56	-14.98	-15.94	-14.98	-20.45	-18.83	-15.94	-15.19	-31.49	-55.92	-62.72	-79.61
48	-16.54	-12.26	-15.44	-11.80	-15.20	-14.90	-15.44	-14.90	-66.67	-65.74	-72.22	-89.23

Table 2.1: Bayesian Information Criteria - Number of observations used for estimation 162

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	53.14	48.42	52.51	53.51	47.15*	76.78	53.72	54.16	75.44	161.49	165.56	157.01
2	84.21	90.83	78.05*	95.92	90.46	86.25	78.05	82.49	145.01	101.81	114.79	203.06
3	72.74	73.02	70.20*	77.34	71.62	77.42	70.20	75.27	149.62	149.62	163.65	220.00
4	68.50	69.11	73.55	74.20	72.02	73.55	58.17*	63.26	147.36	89.40	111.11	167.64
5	32.88	48.61	37.97	53.69	32.82*	37.97	37.97	42.50	109.35	175.38	180.81	164.98
6	54.70	58.40	51.09	63.42	61.63	51.09	58.07	56.18	25.72*	140.06	140.20	155.24
7	96.34	101.06	87.62	104.68	95.90	91.58	87.62	90.83	112.92	80.28*	103.10	203.80
8	75.24	82.85	69.51*	86.91	77.90	78.55	69.51	72.83	147.36	155.91	169.16	219.05
9	73.39	65.00	57.79*	70.08	78.23	78.48	57.79	62.87	134.74	164.83	172.48	210.45
10	86.99	101.78	90.90	106.85	83.64*	91.86	90.90	94.29	116.38	145.01	150.57	201.00
11	86.77	87.89	76.95*	87.76	89.85	83.89	76.95	77.78	112.92	142.58	155.25	211.86
12	103.88	105.80	99.20	105.90	104.80	99.07*	99.20	100.59	122.93	157.84	170.86	209.33
13	53.95	72.65	59.05	79.40	38.08*	58.57	59.05	62.94	109.05	174.45	178.73	164.98
14	81.31	81.16	77.42	85.49	81.43	74.21*	77.44	74.85	159.70	131.94	145.56	215.94
15	80.86	92.36	85.06	93.89	84.96	96.59	83.86	88.95	80.28*	163.20	167.40	168.70
16	158.38*	171.99	162.58	177.08	160.37	163.47	163.46	168.55	181.67	173.16	179.59	229.05
17	56.16	62.45	50.92*	67.54	54.67	55.89	50.92	52.66	101.81	149.62	157.46	200.29
18	79.53	72.88	60.99*	76.62	79.93	80.55	60.99	63.62	142.58	137.45	150.57	219.41
19	63.63	76.30	68.68	81.39	59.26*	63.91	68.72	69.00	97.58	155.27	161.04	188.06
20	69.61	86.73	74.12	91.10	66.36*	71.92	74.12	76.91	80.28	155.91	161.66	181.35
21	114.45	112.10	108.57	116.31	119.90	105.38*	108.57	108.22	129.04	149.62	163.65	218.84
22	127.81	139.20	125.44*	143.57	129.97	130.58	125.44	129.59	136.98	136.98	152.41	218.94
23	77.01	95.64	81.53	100.73	81.30	81.11	81.53	86.10	40.69*	136.98	136.95	176.01
24	129.65	129.62	125.01	134.71	135.05	124.58	125.01	122.88*	136.98	139.57	154.68	223.87
25	20.35*	40.70	25.44	45.79	20.35	25.44	28.21	30.53	173.16	101.81	125.12	164.60
26	65.19	83.82	70.11	85.55	64.18*	69.11	70.11	74.19	97.82	147.36	152.95	187.19
27	91.40	99.13	91.07	102.98	87.74*	91.12	91.07	93.10	126.04	169.31	176.95	207.99
28	84.66	98.65	85.76	99.22	79.26*	87.01	85.76	90.30	122.93	175.38	179.59	193.50
29	92.00	94.96	91.48	100.05	92.88	89.38*	90.73	92.31	164.09	144.48	156.88	219.39
30	126.61	137.00	131.19	142.09	132.92	131.70	131.19	135.40	138.57	97.07*	117.37	189.81
31	82.87	93.98	81.08*	99.07	88.40	85.89	81.08	83.94	129.04	157.84	165.56	213.30
32	95.06	79.70	66.37*	84.71	93.81	95.39	67.04	70.23	131.94	112.92	131.45	213.25
33	65.04*	74.85	69.55	78.36	74.78	86.81	69.55	70.91	122.93	149.62	157.46	196.35
34	79.93	88.64	80.06	93.72	77.45	84.20	85.01	85.15	70.37*	155.91	161.66	167.90
35	77.95	86.86	82.99	91.95	78.30	80.51	82.99	85.60	116.38	75.44*	103.36	187.23
36	92.52	85.55	89.94	88.41	93.72	97.61	93.44	98.52	70.23*	116.04	118.03	181.95
37	75.15	77.94	68.24	83.02	68.00*	96.44	78.77	72.74	105.65	169.31	172.48	160.23
38	71.07*	89.30	76.05	94.39	71.75	76.05	75.92	80.91	161.49	173.16	180.81	214.75
39	73.31	81.84	68.09	86.92	65.37	68.05	68.09	72.41	145.01	5.09*	41.12	150.40
40	47.95	65.15	53.04	68.97	39.13*	51.60	53.04	56.69	40.75	153.89	157.46	157.01
41	33.65	48.67	38.37	50.17	35.80	245.52	38.37	43.46	145.01	5.09*	47.86	150.40
42	74.31	72.57	59.08*	77.66	70.04	76.84	59.08	62.91	129.04	164.83	172.48	208.59
43	87.25	65.37	52.00*	63.06	86.74	78.72	52.00	54.84	109.35	84.93	103.36	209.42
44	70.49*	90.19	75.53	95.13	71.81	75.53	75.53	80.59	166.39	105.65	125.12	179.09
45	29.12	48.50	34.20	53.43	25.35*	34.20	33.80	38.89	105.65	174.30	179.59	164.98
46	100.14*	115.47	101.05	117.56	101.10	103.68	101.05	105.33	153.89	122.93	136.51	208.18
47	73.87	77.07	66.59*	82.15	71.52	72.03	66.59	69.76	89.40	137.45	150.57	201.21
48	65.70	74.46	69.84	80.17	64.90	61.40*	69.84	65.16	159.70	163.20	174.04	220.00

Table 2.2: Bayesian Information Criteria - Number of observations used for estimation 152

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	53.11	48.42	52.45	53.51	47.11*	76.26	53.70	54.05	74.10	150.65	152.03	146.40
2	80.02	88.18	75.93*	93.19	86.54	80.89	75.93	79.64	138.36	95.78	111.96	195.07
3	66.28	69.01	62.45*	73.76	64.44	70.96	62.45	67.50	145.50	145.50	156.52	209.92
4	65.49	61.95	70.44	67.08	68.16	70.58	56.24*	61.33	136.00	78.81	101.27	157.68
5	32.88	48.60	37.97	53.69	32.82*	37.97	37.42	42.51	103.00	164.85	168.44	158.05
6	54.61	58.32	51.09	63.41	61.43	51.09	58.06	56.18	25.40*	130.03	132.91	149.36
7	93.95	96.03	82.76	99.91	93.55	88.46	82.76	87.85	103.00	78.62*	96.56	195.98
8	74.17	82.73	69.27*	86.84	76.91	77.55	69.27	74.07	138.36	144.88	157.65	203.75
9	70.67	65.00	55.08*	70.08	75.02	75.76	61.80	60.61	125.63	154.08	161.00	200.81
10	80.04	89.06	80.40	94.15	75.95*	80.79	80.40	80.83	99.61	133.55	139.07	188.15
11	85.47	86.94	75.76*	86.80	88.53	83.02	75.76	76.84	106.86	131.01	143.40	199.55
12	100.93	102.01	94.84*	103.13	102.17	98.18	94.84	96.79	120.23	145.50	159.02	201.42
13	52.16	70.90	57.25	76.31	37.73*	56.86	57.25	61.05	103.31	164.80	168.97	156.66
14	80.14	81.04	76.76	85.43	80.25	73.86*	76.77	74.55	139.47	119.02	127.37	201.56
15	77.00	90.28	81.38	92.19	79.76	91.37	81.22	86.31	69.33*	153.11	157.18	159.35
16	155.06*	169.12	159.45	174.30	157.42	160.15	160.07	165.16	173.77	165.50	171.94	216.49
17	53.04	62.30	50.59	67.39	50.17*	54.15	78.70	52.49	95.78	136.00	143.95	187.68
18	78.52	72.09	60.09*	76.08	79.03	79.42	60.09	62.89	134.06	134.06	146.81	209.63
19	58.74	72.22	63.67	77.31	54.92*	60.80	63.83	65.89	87.64	144.88	150.60	173.98
20	69.56	86.68	74.07	91.03	66.32*	71.88	74.07	76.87	79.00	147.52	150.60	172.23
21	111.25	100.48*	104.35	105.56	115.99	101.01	104.35	104.48	122.78	142.80	156.52	209.51
22	121.20	130.92	117.00*	136.00	123.83	124.73	117.00	121.85	127.44	127.44	142.86	206.82
23	74.18	91.89	78.51	96.53	78.34	77.63	78.51	82.64	40.16*	130.53	130.19	170.43
24	119.23	119.45	111.50	124.53	123.22	117.31	106.84*	110.14	130.53	130.53	145.69	210.85
25	20.35*	40.70	25.44	45.79	20.35	25.44	28.21	30.53	163.71	99.61	121.79	156.56
26	64.52	82.82	69.46	85.06	63.45*	68.47	69.46	73.55	95.78	138.36	146.25	180.54
27	90.03	98.70	90.59	102.65	86.25*	89.77	90.59	92.42	119.02	155.64	163.28	196.42
28	76.17	92.26	79.10	87.15	72.00*	79.13	79.10	83.57	110.29	163.71	167.67	176.07
29	90.35	91.71	89.71	96.80	90.19	87.67*	88.68	90.78	157.21	136.00	150.60	207.52
30	114.59	127.13	117.02	132.22	120.32	119.68	117.02	120.44	130.97	86.48*	107.18	184.15
31	80.78	89.04	78.59*	94.13	86.36	82.28	81.99	79.93	125.63	148.81	153.98	204.96
32	91.67	78.91	65.64*	83.88	91.02	92.69	66.42	69.54	126.07	103.61	118.99	200.91
33	63.01*	74.77	67.47	78.08	72.45	82.95	67.47	70.45	116.76	138.36	146.25	187.27
34	77.60	86.51	77.91	91.60	75.70	81.89	82.69	83.00	69.33*	149.46	155.27	158.44
35	74.79	83.71	79.83	88.80	75.11	77.38	79.83	82.47	107.81	74.61*	97.82	172.42
36	85.14	75.39	82.59	77.45	86.65	90.23	82.59	87.68	69.48*	114.30	116.38	169.15
37	70.33	77.27	65.95	82.35	59.91*	65.95	74.38	70.76	95.78	160.03	163.33	149.47
38	68.21*	87.26	73.30	92.35	68.90	73.30	73.30	78.38	149.96	161.68	169.33	204.39
39	62.53	81.30	67.49	86.39	64.27	67.45	67.48	71.94	133.05	5.09*	33.44	142.63
40	43.65	62.38	48.74	65.96	36.29*	48.42	48.74	53.51	40.09	141.60	147.30	148.33
41	32.34	48.67	37.31	50.57	33.60	101.34	37.31	42.40	135.47	5.09*	40.49	143.57
42	73.69	72.30	58.74*	77.73	69.62	76.38	58.74	62.52	126.50	153.80	163.20	196.23
43	84.95	56.96	46.25*	57.83	84.76	196.81	46.25	51.34	106.86	83.32	101.00	199.30
44	69.06*	88.57	74.02	93.66	70.59	74.02	74.01	79.01	158.75	100.18	119.35	167.87
45	29.12	47.88	34.20	52.35	25.36*	34.20	33.27	38.36	99.61	161.33	166.30	153.47
46	82.92	98.68	84.62	101.21	79.42*	87.50	84.62	89.57	147.52	110.63	124.84	191.00
47	70.57	76.49	64.82*	81.57	67.93	70.49	64.82	69.04	78.81	128.37	141.55	187.16
48	62.78	72.67	67.02	78.02	62.10	60.29*	67.02	63.88	148.15	148.15	161.00	206.65

Table 2.3: Bayesian Information Criteria - Number of observations used for estimation 142

subject	EU	CEU	PT	CPT	DFT	GS MIN	GS MAX	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
1	51.53	47.97	51.36	53.06	46.30*	71.74	52.02	52.97	73.22	147.15	146.25	134.69
2	76.46	86.47	74.43*	91.36	82.90	77.17	74.43	77.78	133.20	89.48	105.21	183.66
3	61.39	65.09	56.38*	69.63	58.37	66.12	56.38	61.37	135.26	135.26	146.12	197.38
4	63.07	61.75	68.03	66.88	66.33	68.16	55.41*	60.50	128.81	72.49	94.97	153.45
5	31.75*	48.60	36.84	54.92	32.78	36.84	36.35	41.44	100.79	155.22	159.88	148.10
6	53.32	57.78	50.36	62.87	59.70	50.36	56.97	55.44	25.15*	126.46	129.54	141.01
7	85.92	89.63	76.59	93.82	85.27	81.45	76.59	79.77	101.12	72.86*	86.48	182.98
8	69.39	80.76	67.20*	85.40	71.97	72.43	67.20	71.78	124.52	135.88	146.80	186.83
9	68.61	63.83	52.94*	68.92	73.33	78.05	52.94	58.03	123.36	148.78	154.93	188.30
10	78.40	87.94	78.97	93.03	73.76*	79.65	78.97	79.79	97.56	129.90	134.94	178.04
11	79.86	84.75	71.81*	85.19	83.47	78.78	71.81	74.60	97.56	125.02	136.69	186.52
12	92.97	95.08	85.70*	96.15	94.60	89.42	85.70	88.26	107.82	134.39	148.15	189.76
13	49.02	67.99	54.11	74.46	37.40*	54.11	54.11	57.81	101.12	157.14	161.68	148.35
14	78.89	77.34	75.67	82.55	79.45	70.81*	75.70	72.04	131.99	115.18	123.36	191.58
15	76.37	89.33	80.81	92.01	78.67	88.19	80.71	85.80	68.39*	143.62	146.89	149.81
16	145.63*	157.29	149.26	161.94	146.78	150.72	150.56	155.65	163.04	158.11	162.92	205.11
17	53.00	62.30	50.58	67.38	50.16*	54.13	78.36	52.44	93.30	128.81	136.69	178.03
18	75.39	68.36	57.72*	72.99	75.76	76.48	57.72	59.50	125.02	127.51	140.05	197.78
19	48.34	66.90	53.27	71.98	46.47*	52.35	53.43	57.44	81.59	131.62	138.89	162.95
20	63.45	80.78	68.15	85.87	60.24*	66.73	68.15	71.74	62.73	131.05	136.12	161.40
21	102.58	97.60	97.06	102.69	106.99	96.55*	97.06	98.23	110.19	131.05	144.95	195.01
22	117.20	128.23	113.67*	133.32	119.67	121.43	113.67	118.72	118.82	121.47	136.69	191.84
23	72.62	90.39	77.23	95.04	76.46	75.92	77.23	81.00	39.60*	121.47	117.78	161.44
24	110.13	114.08	105.40	119.16	112.95	109.92	105.40	104.41*	125.51	120.18	136.00	195.23
25	20.35*	40.70	25.44	45.79	20.35	25.44	25.44	30.53	149.87	89.48	112.11	149.49
26	59.99	75.87	64.89	79.96	58.28*	62.19	64.89	67.28	93.87	134.39	139.47	168.07
27	84.75	91.92	85.45	96.08	81.93*	83.92	85.45	86.59	106.71	141.86	147.86	184.38
28	74.81	90.91	77.64	84.52	71.33*	77.90	77.64	82.20	108.18	154.91	157.85	166.10
29	83.23	80.82	83.22	85.91	84.67	80.10*	82.36	83.71	147.20	124.02	138.30	195.66
30	87.32	94.19	82.28	99.28	83.00	92.40	82.67	87.76	125.97	75.96*	97.26	171.82
31	74.04	81.04	70.21*	85.64	80.97	77.39	71.88	71.91	122.90	142.95	147.48	192.40
32	82.75	61.13	53.40	66.22	82.49	73.81	53.43	53.34*	107.45	89.75	105.88	189.07
33	62.61*	74.07	66.94	77.28	70.80	82.36	64.88	69.96	110.19	128.81	136.69	176.96
34	76.14	85.51	76.45	90.60	74.23	80.24	81.23	81.54	68.06*	142.27	147.48	150.34
35	70.59	79.88	75.49	84.57	70.53	74.33	75.68	79.41	101.12	68.06*	91.29	157.21
36	77.43	70.73	75.79	75.15	78.55	82.51	77.29	80.88	57.96*	111.74	110.29	157.75
37	66.46	75.60	63.61	80.69	57.55*	63.61	70.71	68.39	86.48	154.54	155.68	138.10
38	63.48*	82.52	68.42	87.60	64.07	68.06	68.57	73.15	136.57	149.01	156.65	190.35
39	60.65	79.87	65.71	84.94	61.87	65.70	65.71	70.66	128.81	5.09*	33.04	135.25
40	43.65	62.28	48.74	65.96	36.29*	48.42	48.74	53.51	39.45	131.99	137.70	142.01
41	31.72	47.63	36.72	47.87	33.21	179.59	36.72	41.81	126.46	5.09*	39.96	133.65
42	71.93	70.99	58.01*	76.08	67.78	74.49	58.01	61.59	120.18	144.73	154.08	185.09
43	84.07	56.96	46.25*	57.74	83.59	266.65	46.25	49.25	104.19	81.59	94.42	190.32
44	64.25*	83.54	69.05	88.56	66.69	69.05	68.88	73.80	144.73	90.27	106.54	155.89
45	29.12	47.94	34.20	52.32	25.35*	34.20	33.28	38.37	97.56	152.83	157.85	143.93
46	79.98	96.82	82.81	100.86	76.42*	84.90	82.81	87.87	137.22	100.47	114.81	176.50
47	63.47	70.66	57.32*	75.75	61.26	63.10	57.32	60.91	68.06	116.93	130.53	174.48
48	53.43	65.23	56.32	69.39	50.76*	55.23	56.32	60.32	138.42	136.57	149.52	193.73

Table 3.1: Summary of 'Best' Model using BIC

Subject	**162**	**152**	**142**	**Overall (see text)**
1	*DFT*	*DFT*	*DFT*	***DFT***
2	*PT*	*PT*	*PT*	***PT***
3	*PT*	*PT*	*PT*	***PT***
4	*GS Max*	*GS Max*	*GS Max*	***GS Max***
5	*DFT*	*DFT*	*EU*	*DFT/EU*
6	*MaxMin*	*MaxMin*	*MaxMin*	***MaxMin***
7	*MaxMax*	*MaxMax*	*MaxMax*	***MaxMax***
8	*PT*	*PT*	*PT*	***PT***
9	*PT*	*PT*	*PT*	***PT***
10	*DFT*	*DFT*	*DFT*	***DFT***
11	*PT*	*PT*	*PT*	***PT***
12	*GS Min*	*PT*	*PT*	*GS Min*
13	*DFT*	*DFT*	*DFT*	***DFT***
14	*GS Min*	*GS Min*	*GS Min*	***GS Min***
15	*MaxMin*	*MaxMin*	*MaxMin*	***MaxMin***
16	EU	EU	EU	**EU**
17	PT	DFT	DFT	PT
18	PT	PT	PT	**PT**
19	DFT	DFT	DFT	**DFT**
20	DFT	DFT	DFT	**DFT**
21	GS Min	CEU	GS Min	GS Min
22	PT	PT	PT	**PT**
23	MaxMin	MaxMin	MaxMin	**MaxMin**
24	Alpha	GS Max	Alpha	Alpha
25	EU	EU	EU	**EU**
26	DFT	DFT	DFT	**DFT**
27	DFT	DFT	DFT	**DFT**
28	DFT	DFT	DFT	**DFT**
29	GS Min	GS Min	GS Min	**GS Min**
30	MaxMax	MaxMax	MaxMax	**MaxMax**
31	PT	PT	PT	**PT**
32	PT	PT	Alpha	PT/Alpha
33	*EU*	*EU*	*EU*	***EU***
34	*MaxMin*	*MaxMin*	*MaxMin*	***MaxMin***
35	*MaxMax*	*MaxMax*	*MaxMax*	***MaxMax***
36	*MaxMin*	*MaxMin*	*MaxMin*	***MaxMin***
37	*DFT*	*DFT*	*DFT*	***DFT***
38	*EU*	*EU*	*EU*	***EU***
39	*MaxMax*	*MaxMax*	*MaxMax*	***Maxmax***
40	*DFT*	*DFT*	*DFT*	***DFT***
41	*MaxMax*	*MaxMax*	*MaxMax*	***MaxMax***
42	*PT*	*PT*	*PT*	***PT***
43	*PT*	*PT*	*PT*	***PT***
44	*EU*	*EU*	*EU*	***EU***
45	*DFT*	*DFT*	*DFT*	***DFT***
46	*EU*	*DFT*	*DFT*	*EU/DFT*
47	*PT*	*PT*	*PT*	***PT***
48	*GS Min*	*GS Min*	*DFT*	*GS Min*

Treatment 1: subjects 1 through 15; Treatment 2: subjects 16 through 32; Treatment 3; subjects 33 through 48. In the final column, **bold** indicates no conflict through the data sets.

Figure 3.2: An overall summary

	Treatment 1	**Treatment 2**	**Treatment 3**	**All treatments**
EU[22]	½	2	3½	6
CEU	0	0	0	0
PT	5	4½	3	12½
CPT	0	0	0	0
DFT	3½	5	3½	12
GS Min	2	2	1	5
GS Max	1	0	0	1
Alpha	0	1½	0	1½
MaxMin	2	1	2	5
MaxMax	1	1	3	5
MinReg	0	0	0	0
Hurwicz	0	0	0	0
Totals	**15**	**17**	**16**	**48**

[22] We note that Expected Value Maximisation is a special case of EU. Subject 25 clearly was simply maximising the Expected Value (albeit with the wrong probabilities) while subject 16 was close to doing so (and with probabilities close to the correct ones). We thank Edi Karni for pointing out these special cases.

Table 4.1: Predictive Ability in and out of sample (number is % correct) - Number of observations used for estimation 162

subject	EU	CEU	PT	CPT	DFT	GS Min	GS Max	Alpha	MaxMin	MaxMax	MinReg	Hurwicz
	In	in	in	in	in	in	in	in	in	in	in	in
1	96	99	98	99	96	95	98	98	94	77	77	72
2	92	94	93	94	91	93	93	94	81	90	88	67
3	93	95	95	96	93	93	95	95	80	80	77	62
4	94	96	95	96	94	95	97	97	81	92	89	73
5	98	98	98	98	99	98	98	98	89	70	70	67
6	98	98	99	98	98	99	98	99	99	83	83	74
7	85	87	89	90	85	87	89	89	87	92	89	67
8	93	94	94	94	93	92	94	94	81	78	75	62
9	92	98	98	98	92	92	98	98	84	75	74	64
10	94	92	93	93	93	91	93	93	88	81	81	67
11	90	95	94	95	89	91	94	95	88	82	80	65
12	89	91	91	93	90	90	91	91	86	78	75	63
13	99	99	99	97	99	99	99	99	89	70	71	67
14	94	94	93	95	93	93	93	96	77	85	82	63
15	90	93	90	92	89	89	91	91	93	76	76	70
16	81	84	81	84	79	81	81	81	65	72	71	60
17	94	98	96	98	95	96	96	98	90	80	79	67
18	91	95	93	95	92	93	93	94	82	83	81	62
19	94	94	94	94	94	94	94	94	90	78	77	68
20	94	94	94	95	94	96	94	95	93	78	78	69
21	86	88	88	90	86	88	88	88	85	80	77	62
22	89	87	85	88	89	87	85	85	83	83	80	63
23	95	95	95	95	95	95	95	95	97	83	84	72
24	84	88	87	88	84	86	87	87	83	82	79	61
25	100	100	100	100	100	100	99	100	72	90	86	68
26	93	93	94	94	93	93	94	93	91	81	80	69
27	91	94	92	93	91	92	92	92	86	73	72	65
28	95	96	96	94	94	96	96	96	86	70	71	64
29	91	92	91	92	90	90	92	92	75	81	79	63
30	91	90	92	90	91	91	92	90	82	90	87	69
31	93	96	94	96	93	92	94	95	85	78	77	64
32	90	95	95	95	91	92	95	95	85	88	85	65
33	95	96	94	96	94	93	94	94	86	80	79	69
34	94	93	93	93	93	95	94	93	94	78	78	72
35	93	94	93	94	93	96	93	96	88	94	90	69
36	89	91	89	93	88	89	88	88	94	87	87	68
37	94	94	94	94	94	93	94	94	90	73	74	71
38	91	93	92	93	92	92	91	92	77	72	70	64
39	92	91	90	91	90	90	90	90	81	100	98	74
40	97	98	97	96	97	98	97	98	98	79	79	72
41	99	99	99	99	98	98	99	99	81	100	97	74
42	94	97	96	97	92	93	96	96	85	75	74	63
43	91	97	97	98	91	93	97	97	89	93	90	67
44	93	93	93	93	93	93	93	93	75	90	86	69
45	99	99	99	99	99	99	99	99	90	71	71	67
46	90	89	90	90	90	90	90	89	78	86	83	65
47	93	95	94	95	93	94	94	94	92	83	81	68
48	93	94	93	94	93	94	93	94	77	76	73	62
average	92.63	94.06	93.52	94.25	92.40	92.90	93.54	93.81	85.46	81.58	80.00	66.83

Table 4.2: Predictive Ability in and out of sample (number is % correct)
Number of observations used for estimation 152

subject	EU		CEU		PT		CPT		DFT		GS Min		GS Max		Alpha		MaxMin		MaxMax		MinReg		Hurwicz	
	in	out	in	out	in	out	in	out	in	out	in	out	in	out	In	out	in	out	in	out	in	out	in	out
1	96	100	99	100	97	100	99	90	96	100	95	100	97	100	97	100	93	100	77	70	78	60	73	65
2	91	90	94	90	93	100	94	90	91	90	92	90	93	100	94	90	81	90	90	90	88	100	66	80
3	93	80	96	80	97	80	96	80	93	80	93	80	97	80	97	80	79	100	79	100	76	90	61	75
4	95	90	97	70	95	90	97	70	94	90	95	90	97	100	97	100	82	70	93	80	89	80	74	70
5	98	100	98	100	98	100	98	100	99	100	98	100	98	100	98	100	89	90	70	80	70	70	67	75
6	98	100	98	100	99	100	98	100	98	100	99	100	98	100	99	100	99	100	83	80	83	90	73	85
7	85	90	90	80	89	80	90	80	85	90	86	80	89	80	89	80	88	80	92	100	89	90	66	85
8	92	100	93	100	94	100	93	100	93	100	91	100	94	100	94	100	81	80	79	70	76	70	63	40
9	92	90	98	100	98	90	98	100	91	90	92	90	95	90	98	80	84	80	76	70	74	70	63	70
10	94	90	91	80	95	80	91	80	93	90	93	80	95	80	93	80	89	60	82	70	82	70	67	60
11	89	100	95	100	93	100	95	100	89	100	91	100	93	100	95	100	88	90	83	70	80	70	65	65
12	88	100	92	90	93	90	93	90	89	100	89	95	93	90	91	90	86	100	79	60	76	60	62	80
13	99	100	99	100	99	100	97	100	99	100	99	100	99	100	99	100	88	90	70	70	71	70	66	69
14	93	100	94	100	93	100	95	100	93	100	92	100	93	100	95	100	80	30	86	70	84	50	64	45
15	89	90	91	100	89	90	92	100	89	90	88	90	89	90	89	90	94	80	76	70	76	70	70	75
16	80	90	83	100	80	90	82	90	79	90	80	90	80	90	80	90	64	80	71	80	70	80	60	60
17	95	90	97	100	96	100	97	100	95	90	97	90	92	100	97	100	90	90	82	60	80	60	67	55
18	91	100	95	100	93	100	95	100	91	100	92	100	93	100	95	100	82	80	82	100	80	100	61	80
19	94	90	95	90	94	90	95	90	94	90	94	90	94	90	94	90	91	80	79	70	78	70	69	50
20	94	100	94	100	94	100	95	100	93	100	95	100	94	100	95	100	93	100	78	80	78	70	69	70
21	86	100	90	80	88	90	90	80	86	100	88	90	88	90	88	100	85	90	80	90	76	90	61	80
22	88	90	88	70	87	70	88	70	88	90	87	90	87	70	87	70	83	80	83	80	80	80	63	60
23	94	100	94	90	94	100	95	80	94	100	94	100	94	100	94	100	97	100	83	90	83	90	71	85
24	85	80	87	80	87	80	88	80	85	80	87	80	88	80	87	80	83	90	83	80	79	80	62	55
25	100	100	100	100	100	100	100	100	100	100	100	100	99	100	100	100	71	80	89	100	86	100	67	80
26	93	100	93	100	93	100	94	100	93	100	93	100	93	100	93	100	90	100	81	80	80	90	68	85
27	90	100	94	100	91	100	93	100	91	100	91	100	91	100	91	100	86	90	74	60	73	60	65	65
28	97	80	97	80	97	80	95	80	95	80	97	80	97	80	97	80	88	70	71	60	72	60	64	65
29	90	100	92	90	90	100	92	90	89	90	90	90	90	100	91	100	74	90	81	80	78	90	63	70
30	92	60	92	70	94	60	92	70	92	70	92	60	94	60	94	60	81	90	90	80	87	80	67	95
31	93	100	95	80	94	90	95	80	92	100	92	80	94	90	95	90	84	100	78	80	77	70	63	80
32	90	90	94	100	95	100	95	100	90	90	91	90	95	100	95	100	84	90	89	80	86	70	65	65
33	94	90	96	100	94	90	96	100	94	90	93	80	94	90	93	100	86	90	81	70	80	70	68	75
34	93	100	93	100	91	100	93	100	93	100	95	100	93	100	91	100	94	100	78	90	77	90	72	70
35	93	100	94	90	92	100	94	90	92	100	94	100	92	100	94	100	88	80	93	100	90	90	70	55
36	90	75	91	75	90	75	94	85	90	75	90	75	90	75	90	75	94	95	87	95	87	95	69	60
37	95	90	94	100	94	90	94	100	95	90	94	90	95	90	94	100	90	80	73	80	74	80	71	65
38	92	90	93	90	91	90	93	90	94	90	92	90	91	90	91	90	77	70	72	70	70	70	63	75
39	89	100	91	100	89	100	91	100	91	100	89	100	89	100	89	100	82	70	100	100	98	90	74	75
40	97	90	97	90	97	90	96	90	97	90	97	90	97	90	97	90	97	100	80	70	79	80	72	75
41	99	90	99	100	99	90	99	100	99	90	99	90	99	90	99	90	82	80	100	100	97	90	74	80
42	93	100	97	100	95	100	97	100	91	100	93	100	95	100	96	100	84	100	76	60	74	70	63	55
43	90	90	97	90	97	90	97	90	90	100	74	85	97	90	97	90	88	100	92	100	89	100	66	70
44	93	100	93	100	93	100	93	100	92	100	93	100	93	100	93	100	74	80	89	90	86	90	69	55
45	99	100	99	90	99	100	99	90	99	100	99	100	99	90	99	90	89	90	72	50	72	50	68	50
46	90	80	92	80	91	80	91	80	91	80	91	80	91	80	91	80	78	90	87	70	84	70	67	45
47	93	90	95	100	95	90	95	100	93	90	94	100	95	90	94	100	93	80	84	80	81	80	68	60
48	93	90	94	90	93	90	95	90	93	90	93	100	93	90	94	100	78	70	78	50	74	60	62	60
average	92.44	93.02	94.17	91.98	93.52	92.19	94.25	91.56	92.35	93.23	92.35	91.77	93.46	92.19	93.75	92.81	85.44	85.73	81.90	78.65	80.15	77.60	66.69	68.00

Table 4.3: Predictive Ability in and out of sample (number is % correct)
Number of observations used for estimation 142

subject	EU		CEU		PT		CPT		DFT		GS Min		GS Max		Alpha		MaxMin		MaxMax		MinReg			Hurwicz
	in	out	in	out	in	out	in	out	in	out	in	out	in	out	in	out	in	out	in	out	in	out	in	out
1	96	100	98	100	96	100	98	95	96	100	93	92	97	100	97	100	93	100	75	85	77	75	74	62
2	92	85	94	95	92	95	94	95	92	85	92	85	92	95	94	95	80	95	90	90	87	95	66	77
3	94	80	97	80	98	80	96	80	94	80	94	80	98	80	98	80	79	90	79	90	76	85	61	68
4	94	95	97	85	94	90	97	85	94	90	94	95	96	100	96	100	81	80	93	85	89	85	73	80
5	98	100	98	100	98	100	97	95	99	100	98	100	98	100	98	100	88	95	70	75	70	75	67	68
6	98	100	98	100	99	100	98	100	98	100	99	100	98	100	99	100	99	100	82	90	82	95	73	80
7	86	75	90	75	90	75	90	85	86	75	87	70	90	75	90	75	87	90	92	95	90	85	67	72
8	94	95	93	100	94	100	93	100	93	95	92	95	94	100	94	100	82	70	79	75	76	70	65	40
9	92	95	98	100	98	90	98	100	91	95	94	95	98	90	98	90	83	90	75	80	73	80	63	65
10	94	95	90	90	95	90	90	90	94	95	92	90	95	90	93	90	89	80	81	85	80	85	67	65
11	89	90	94	95	94	95	94	100	89	90	91	90	94	95	94	95	89	85	82	80	80	80	65	60
12	92	80	94	85	94	80	94	80	92	80	92	90	94	80	95	80	87	85	80	65	76	65	62	68
13	99	100	99	100	99	100	96	95	99	100	99	100	99	100	99	100	88	95	69	80	69	80	66	70
14	92	100	94	95	92	100	95	95	92	100	93	95	92	100	94	95	80	60	85	85	83	75	63	60
15	88	95	91	100	88	95	93	100	88	95	88	85	89	95	89	95	94	90	77	70	77	70	70	72
16	81	80	82	75	82	75	83	75	81	80	81	80	81	80	81	80	64	75	69	90	68	90	58	70
17	94	95	97	100	96	100	97	100	95	95	96	95	92	100	97	100	89	95	81	75	80	75	67	63
18	92	95	96	90	94	95	96	95	91	95	92	95	94	95	95	90	82	80	82	95	79	95	61	72
19	95	75	95	75	95	75	95	75	95	75	96	80	95	75	96	80	91	85	80	65	79	70	70	58
20	94	95	95	95	94	95	95	95	94	95	95	95	94	95	94	95	94	85	80	65	80	65	70	65
21	87	90	89	85	88	85	89	85	86	90	87	90	88	85	89	95	86	80	80	80	77	80	62	67
22	88	90	88	80	86	80	88	80	88	90	87	90	86	80	87	80	83	80	83	85	79	85	64	55
23	94	100	94	95	94	100	95	90	94	100	93	100	94	100	93	100	97	100	83	85	84	80	71	80
24	86	80	87	85	87	80	87	85	85	80	86	80	87	80	87	85	82	90	83	75	80	75	63	48
25	100	100	100	95	100	95	100	100	100	100	100	100	100	95	100	95	73	65	90	90	86	90	67	75
26	93	90	94	95	94	90	92	95	93	90	94	90	94	90	94	90	89	100	80	90	79	90	69	72
27	90	90	96	90	91	90	93	90	91	90	92	90	91	90	92	90	87	80	76	55	75	50	64	67
28	96	90	96	90	96	90	94	85	95	90	96	90	96	90	96	90	87	85	71	65	72	65	64	65
29	90	95	93	85	89	90	93	85	89	90	91	80	90	95	91	80	74	85	82	75	79	80	62	70
30	93	65	93	65	95	65	93	65	93	65	93	65	95	70	95	70	80	95	91	80	88	80	67	82
31	93	95	96	90	94	90	96	90	93	95	93	85	94	90	96	90	83	100	77	85	76	80	63	70
32	92	80	95	75	94	80	95	75	92	80	92	75	96	85	95	80	87	70	90	75	87	70	65	65
33	94	95	96	100	94	95	96	100	94	95	92	90	93	100	93	100	86	90	81	75	80	75	68	72
34	93	100	93	100	91	100	93	100	92	100	94	100	93	100	91	100	94	100	77	90	76	90	71	72
35	92	95	94	90	92	95	93	90	92	95	94	95	92	95	94	95	88	85	94	95	90	90	71	57
36	91	73	93	77	91	73	92	73	89	73	91	73	90	73	91	73	95	88	86	98	87	93	70	58
37	94	95	94	100	94	95	94	100	95	90	94	95	95	95	93	95	91	80	72	85	73	80	72	65
38	92	85	94	85	93	85	94	85	94	90	94	85	92	85	94	85	78	65	73	65	71	65	63	67
39	89	100	90	100	89	100	90	100	91	100	89	100	89	100	90	95	81	85	100	100	98	95	74	75
40	96	95	97	95	96	95	96	95	97	95	97	95	96	95	97	95	97	100	80	75	79	80	71	80
41	99	95	99	100	99	100	99	95	99	95	87	90	99	100	99	100	82	80	100	100	97	95	74	75
42	93	100	96	95	96	100	96	95	92	95	92	100	96	100	96	100	84	95	76	70	74	75	64	58
43	89	95	97	95	97	95	97	95	89	100	62	62	97	95	97	95	87	100	92	100	89	95	65	75
44	92	95	94	95	94	95	94	95	92	95	94	95	93	95	94	95	76	65	90	85	87	80	70	55
45	99	100	99	95	99	100	99	95	99	100	99	100	99	95	99	95	89	95	72	65	72	65	68	60
46	90	85	92	90	91	90	91	90	92	85	90	85	91	90	90	90	77	85	87	75	85	75	67	53
47	95	85	96	90	96	90	96	90	95	85	94	90	96	90	96	90	94	80	85	75	82	75	69	60
48	94	90	96	90	94	90	96	90	94	90	93	90	94	90	93	90	77	75	78	60	75	65	62	58
average	92.67	91.10	94.40	90.88	93.67	90.90	94.17	90.58	92.67	90.90	92.04	89.31	93.67	91.52	94.02	91.21	85.48	85.90	81.88	80.69	80.17	79.44	66.83	66.48

Table 5: Doubtful Subjects

Subject	first model	second model	BIC 162		BIC 152		BIC 142		predictions 152		predictions 142	
			first	**second**	**first**	**second**	**first**	**second**	**first**	**second**	**first**	**second**
5	DFT	EU	32.82	32.88	32.82	32.88	32.78	31.75	100	100	100	100
12	GS Min	PT	99.07	99.20	98.18	94.84	89.42	85.70	95	90	90	80
17	PT	DFT	50.92	54.67	50.59	50.17	50.58	50.16	100	90	100	95
21	GS Min	CEU	105.38	112.10	101.01	100.48	96.55	97.60	90	80	90	85
24	Alpha	GS Max	122.88	125.01	110.14	106.84	104.41	105.40	80	80	85	80
32	PT	Alpha	66.37	70.23	65.64	69.54	53.40	53.34	100	100	80	80
46	EU	DFT	100.14	101.10	82.92	79.42	79.98	76.42	80	80	85	85
48	GS Min	DFT	61.40	64.90	60.29	62.10	55.23	50.76	100	90	90	90

Table 6.1: Models that outperform the 'best' model in terms of predictive ability
Number of observations used for estimation 152

subject	'best model'	other models								
3	**PT 80**	MaxMin 100	MaxMax 100	MinReg 90						
9	**PT 90**	CEU 100	CPT 100							
12	**PT 90**	EU 100	DFT 100	GS Min 95	MaxMin 100					
15	**MaxMin 80**	EU 90	CEU 100	PT 90	CPT 100	DFT 90	GS Min 90	GS Max 90	Alpha 90	
16	**EU 90**	CEU 100								
17	**DFT 90**	CEU 100	PT 100	CPT 100	GS Max 100	Alpha 100				
21	**CEU 80**	EU 100	PT 90	DFT 100	GS Min 90	GS Max 90	Alpha 100	MaxMin 90	MaxMax 90	MinReg 90
22	**PT 70**	EU 90	DFT 90	GS Min 90	MaxMin 80	MaxMax 80	MinReg 80			
24	**GS Max 80**	MaxMin 90								
29	**GS Min 90**	EU 100	PT 100	GS Max 100	Alpha 100					
30	**MaxMax 80**	MaxMin 90	Hurwicz 95							
31	**PT 90**	EU 100	DFT 100	MaxMin 100						
33	**EU 90**	CEU 100	CPT 100	Alpha 100						
37	**DFT 90**	CEU 100	CPT 100	Alpha 100						
40	**DFT 90**	MaxMin 100								
43	**PT 90**	DFT 100	MaxMin 100	MaxMax 100	MinReg 100					
46	**DFT 80**	MaxMin 90								
47	**PT 90**	CEU 100	CPT 100	GS Min 100	Alpha 100					

The numbers are the percentage of correct predictions out of (estimation) sample.

Table 6.2: Models that outperform the 'best' model in terms of predictive ability
Number of observations used for estimation 142

subject	**'best model'**	**other models**						
3	**PT** **80**	MaxMin 90	MaxMax 90	MinReg 85				
9	**PT** **90**	EU 95	CEU 100	CPT 100	DFT 95	GS Min 95		
11	**PT** **95**	CPT 100						
12	**PT** **80**	CEU 85	GS Min 90	MaxMin 85				
14	**GS Min** **95**	EU 100	PT 100	DFT 100	GS Max 100			
15	**MaxMin** **90**	EU 95	CEU 100	PT 95	CPT 100	DFT 95	GS Max 95	Alpha 95
16	**EU** **80**	MaxMax 90	MinReg 90					
17	**DFT** **95**	CEU 100	PT 100	CPT 100	GS Max 100	Alpha 100		
19	**DFT** **75**	GS Min 80	Alpha 80	MaxMin 85				
21	**GS Min** **90**	Alpha 95						
22	**PT** **80**	EU 90	DFT 90	GS Min 90	MaxMax 85	MinReg 85		
24	**Alpha** **85**	MaxMin 90						
26	**DFT** **90**	CEU 95	CPT 95	MaxMin 100				
29	**GS Min** **80**	EU 95	CEU 85	PT 90	CPT 85	DFT 90	GS Max 95	MaxMin 85
30	**MaxMax** **80**	MaxMin 95	Hurwicz 82					
31	**PT** **90**	EU 95	DFT 95	MaxMin 100				
32	**Alpha** **80**	GS Max 85						
33	**EU** **95**	CEU 100	CPT 100	GS Max 100	Alpha 100			
36	**MaxMin** **88**	MaxMax 98	MinReg 93					
37	**DFT** **90**	EU 95	CEU 100	PT 95	CPT 100	GS Min 95	GS Max 95	Alpha 95
38	**EU** **85**	DFT 90						
40	**DFT** **95**	MaxMin 100						
43	**PT** **95**	DFT 100	MaxMin 100	MaxMax 100				
46	**DFT** **85**	CEU 90	PT 90	CPT 90	GS Max 90	Alpha 90		

The numbers are the percentage of correct predictions out of (estimation) sample.

Table 7.1: Explanatory and Predictive ability
Number of observations used for estimation: 152

Pred → Fit ↓	EU	CEU	PT	CPT	DFT	GS Min	GS Max	Alpha	Max Min	Max Max	Min Reg	Hurwicz	*Total*	Index*
EU	**3**	1.33		0.33				0.33					*5*	0.60
CEU	0.33	**0**			0.33			0.33					*1*	0.00
PT	1.00	0.75	**6**	1.08	0.92	0.58		0.25	1.42	0.75	0.25		*13*	0.46
CPT				**0**									*0*	0.00
DFT		0.53	0.20	0.53	**10**		0.20	0.53	2.00				*14*	0.71
GS Min	0.25		0.25			**2**	0.25	0.25					*3*	0.67
GS Max							**1**		1				*2*	0.50
Alpha								**0**					*0*	0.00
MaxMin		0.50		0.50					**4**	1.00			*6*	0.67
MaxMax										**3**		1.00	*4*	0.75
MinReg											**0**		*0*	0.00
Hurwicz												**0**	*0*	0.00
Total	*4.58*	*3.12*	*6.45*	*2.45*	*11.25*	*2.58*	*1.45*	*1.70*	*8.42*	*4.75*	*0.25*	*1*	*48*	

*Index: ratio of the diagonal entry to the row total.

Table 7.2: Explanatory and Predictive ability
Number of observations used for estimation: 142

Pred → Fit ↓	EU	CEU	PT	CPT	DFT	GS Min	GS Max	Alpha	Max Min	Max Max	Min Reg	Hurwicz	*Total*	Index*
EU	**3**	0.25		0.25	1.00		0.25	0.25		0.50	0.50		*6*	0.50
CEU		**0**											*0*	0.00
PT	0.33	0.50	**5**	1.50	0.67	1.33			1.83	0.83			*12*	0.42
CPT				**0**									*0*	0.00
DFT		0.90	0.40	0.90	**8**		0.40	0.40	3.00				*14*	0.57
GS Min	0.75		0.25		0.25	**0**	0.75	1.00					*3*	0.00
GS Max							**1**						*1*	1.00
Alpha							1.00	**0**	1.00				*2*	0.00
MaxMin		0.50		0.50					**3**	1.00			*5*	0.60
MaxMax									1.00	**4**			*5*	0.80
MinReg											**0**		*0*	0.00
Hurwicz												**0**	*0*	0.00
Total	*4.08*	*2.15*	*5.65*	*3.15*	*9.92*	*1.33*	*3.40*	*1.65*	*9.83*	*6.33*	*0.50*	*0.00*	*48*	

*Index: ratio of the diagonal entry to the row total.

Figure 1: Ilustration of the DFT model

Appendix 1: The Experimental Instructions

EXEC

Centre for Experimental Economics at the University of York

Welcome to this experiment. MIUR (the Italian ministry for the universities) has provided the funds to finance this research. Depending on your decisions you may earn a considerable amount of money which will be paid to you in cash immediately after the end of the experiment. This sum will be composed of the £10 participation fee plus your 'earnings' from a lottery. This latter could be a loss of £10, a gain of £10 or gain of £100. You cannot walk away from this experiment with less money than that with which you arrived, though you might walk away with £20 more or with £110 more.

There are no right or wrong ways to complete the experiment, but the decisions that you take will have implications for what you are paid at the end of the experiment. This depends partly on the decisions that you take during the experiment and partly on chance. So you will need to read these instructions carefully.

At the end of the experiment you will be asked to complete a brief questionnaire and to sign a receipt for the payment that you received, and to acknowledge that you participated voluntarily in the experiment. The results of the experiment will be used for the purpose of academic research and will be published and used in such a way that your anonymity will be preserved.

Outline of the experiment

You will be asked 162 questions. Each will be of the same type. You will be presented with two lotteries and you will be asked which you prefer. After you have answered all 162 questions, one of them will be selected at random, the lottery that you said that you preferred on that question will be played out, and you will be paid the outcome: if the outcome is a loss of £10 you will leave the experiment with the same as when you came; if the outcome is a gain of £10 you will leave the experiment with £20 more than when you came; if the outcome is a gain of £100 you will leave the experiment with £110 more than when you came. If you did not express a preference on the selected question then one of the two lotteries will be selected at random and played out. It is clearly in your interests to answer each question as if that were the question to be played out.

The Bingo Blower

You will have noticed a Bingo Blower in the laboratory. In this Blower there are balls of three different colours: pink, blue and yellow. The balls are constantly being blown about in the Blower. At the end of the experiment, when we come to play out your preferred choice on one of the questions, we will use this Bingo Blower to determine a colour: we will allow you to open the exit chute – this will lead to one ball being expelled. Obviously this expulsion will be done at random as there is no way that you can control the colour of the ball that emerges. The colour of the ball and the lottery that you chose on the question that was selected will determine your payment.

The Questions

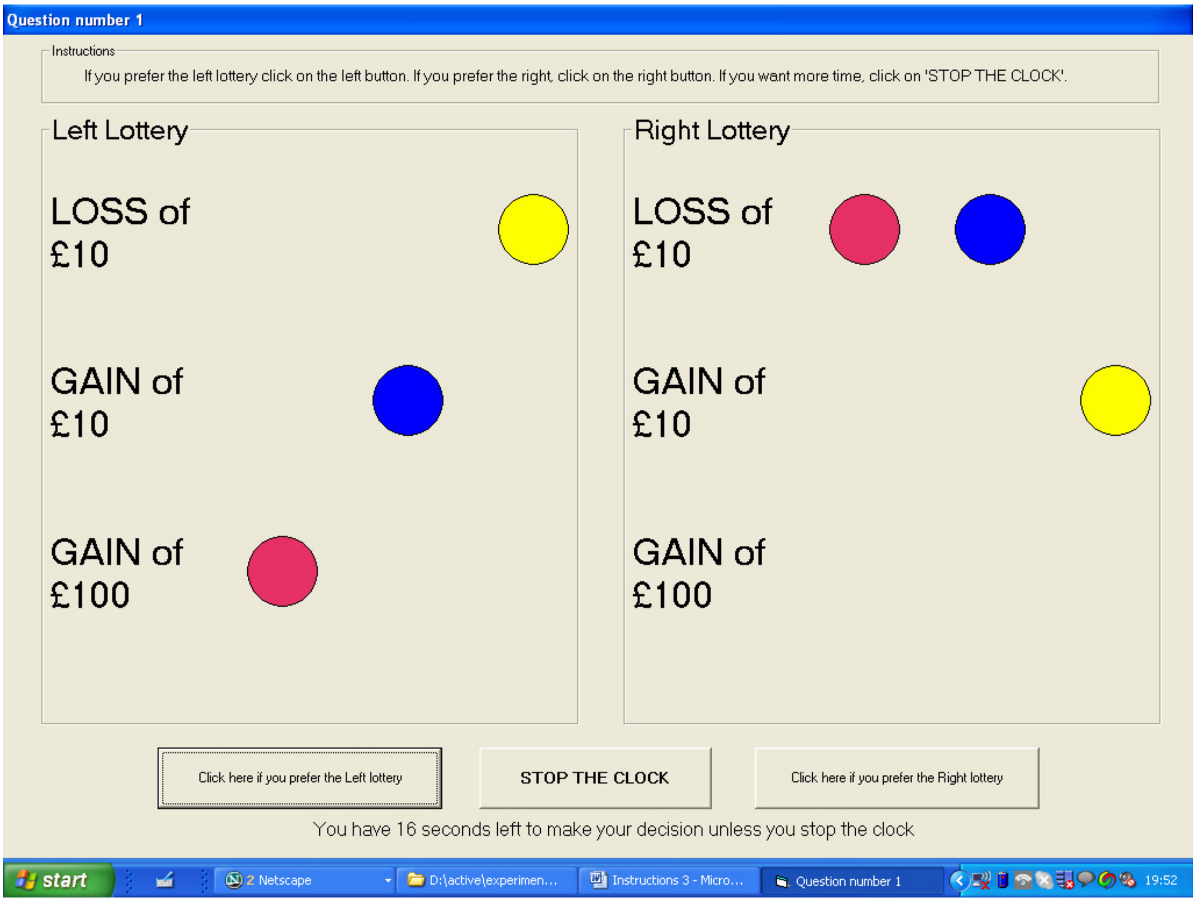
A sample question is illustrated in the Figure attached to these Instructions. In this figure, there are two lotteries – that on the left and that on the right. The lottery on the left would lead to a loss of £10 if the ball expelled was yellow, to a gain of £10 if the ball expelled was blue and to a gain of £100 if the ball expelled was pink. The lottery on the right would lead to a loss of £10 if the ball expelled was pink or blue and to a gain of £10 if the ball expelled was yellow. You have to decide for each question whether you prefer the lottery on the left or that on the right. You should indicate your choice by clicking on the box below the appropriate lottery. You will be given at least 30 seconds to make up your mind and you cannot proceed to the next question until these 30 seconds have elapsed. The number of seconds left to make your decision will be indicated at the bottom of the screen. If you want more time, simply click on 'STOP THE CLOCK'; then click on 'RESTART THE CLOCK' when you are happy to proceed. If the 30 seconds have elapsed and you have not taken a decision then 'no decision' will be recorded for that question.

The end of the experiment

After you have answered all 162 questions you will be asked to call over an experimenter. In front of him or her you will choose at random one of the questions - by picking at random a ticket from a set of cloakroom tickets numbered from 1 to 162. The computer will recall that question and your answer to it, and then you will play out your preferred choice on that question – in the manner described above. If you did not take a decision on that question then you will toss a coin to determine which of the two lotteries will be played out. You will then be asked to fill in a short questionnaire. We will then pay you, you will sign a receipt and then you will be free to go. Note that the experiment will take at least 81 minutes of your time. You can take longer and it is clearly in your interests to be as careful as you can when you are answering the questions.

If you have any questions at any stage, please ask one of the experimenters.

John Hey Gianna Lotito Anna Maffioletti


Question number 1
Instructions
If you prefer the left lottery click on the left button. If you prefer the right, click on the right button. If you want more time, click on 'STOP THE CLOCK'.
Left Lottery
LOSS of £10
GAIN of £10
GAIN of £100
Right Lottery
LOSS of £10
GAIN of £10
GAIN of £100
Click here if you prefer the Left lottery
STOP THE CLOCK
Click here if you prefer the Right lottery
You have 16 seconds left to make your decision unless you stop the clock
start
2 Netscape
Question number 1
19:52

Appendix 2: The Specifications

This appendix provides technical detail on the various specifications that we fitted to the data. We start with some notation. We denote the three possible outcomes in the experiment by x_1, x_2 and x_3. Except for the CPT specification we normalise the highest to have a utility of 1 and the lowest to have a utility of 0; we denote the utility of the middle outcome by *u*. We denote the three colours by *a, b* and *c*. A lottery can be denoted by

$$L = (x_1, S_1; x_2, S_2; x_3, S_3) \tag{1}$$

Here S_i is the state (one of $\varnothing$, a, b, c, $a \cup b$, $a \cup c$, $b \cup c$ and $a \cup b \cup c$) in which the lottery pays out x_i.

1. Expected Utility theory (EU)

In this subjects choose between lotteries on the basis of their expected utility, calculated on the basis of the subject's subjective probabilities attached to the various states. The expected utility of the lottery *L* is given by

$$EU(L) = p_2 u + p_3 \tag{2}$$

where p_i is the (subjective) probability of state *i*. If we use the notation that $p_i = P(S_i)$ where S_i denotes the state in which the lottery pays out x_i, then we have

$$\begin{aligned} &P(\varnothing) = 0 \\ &P(a) = p_a \quad P(b) = p_b \quad P(c) = p_c \\ &P(a \cup b) = p_a + p_b \quad P(a \cup c) = p_a + p_c \quad P(b \cup c) = p_b + p_c \\ &P(a \cup b \cup c) = p_a + p_b + p_c = 1 \end{aligned} \tag{3}$$

where p_a, p_b and p_c are the subject's subjective probabilities for the three colours. In this model we estimate u, p_a, p_b and p_c (subject to the constraint that $p_a + p_b + p_c = 1$).

2. Choquet Expected Utility theory (CEU)

Here the Choquet Expected Utility of the lottery *L* is given by

$$CEU(L) = w_{23} u + w_3 (1-u) = (w_{23} - w_3) u + w_3 \tag{4}$$

where w_i is the Choquet capacity (or weight[23]) of state i. If we use the notation that $w_i = W(S_i)$ where S_i denotes the state in which the lottery pays out x_i, then we have, in order to satisfy the Choquet conditions, that

$$\begin{aligned} &W(\varnothing) = 0 \\ &W(a) = w_a \quad W(b) = w_b \quad W(c) = w_c \\ &W(a \cup b) = w_{ab} \quad W(a \cup c) = w_{ac} \quad W(b \cup c) = w_{bc} \\ &W(a \cup b \cup c) = 1 \end{aligned} \tag{5}$$

Here w_a, w_b, w_c w_{ab}, w_{ac} and w_{bc} are the subject's Choquet capacities (or weights) for the various possible states. In this model we estimate u, w_a, w_b, w_c w_{ab}, w_{ac} and w_{bc}. Note that there is no necessity that $w_{de} = w_d + w_e$ for any d or e. That is, there is no necessity that the weights are additive (probabilities). Indeed this is the main difference between Expected Utility theory and Choquet Expected Utility theory.

3. Prospect Theory (PT)

This is a preference functional 'between' that of EU and the Choquet Expected Utility functional. We should say at the outset that we are hesitant about the acceptability of this term being used in this context, but it seems appropriate. Prospect Theory (see Kahneman and Tversky 1979) envisages utilities being weighted by some function of the 'true' probabilities. If there are true probabilities of the various colours π_a, π_b and π_c then Prospect Theory envisages them being replaced by $f(\pi_a)$, $f(\pi_b)$ and $f(\pi_c)$. If we denote these respectively by p_a, p_b and p_c then we get this specification. It is precisely the same as the Expected Utility preference functional except for the fact that the 'probabilities' are not additive. In this model we estimate u, p_a, p_b and p_c (but no longer subject to the constraint that $p_a + p_b + p_c = 1$). We note that this preference functional may not satisfy dominance (though it does so in this context), unlike the Choquet preference functional,

[23] We borrow this term from Rank Dependent Expected Utility theory, which has strong affinities with the Choquet Expected Utility theory.

which does. It could also be interpreted as a model in which the decision-maker has several possible probabilities for each of the three colours, and works with the *minimum probability* for each colour.

4. Cumulative Prospect Theory (CPT)

This is almost the same as Choquet Expected Utility theory except for the incorporation of a reference point. We take this to be a gain of £0 in the experiment. We normalise the utility function so that *u(£0) = 0* and *u(£110) = 1* and estimate *u = u(£10)*. We also estimate *t = -u(-£10)*. In other respects the theory is the same as CEU. We note that CEU is nested within EU.

5. Decision Field Theory (DFT)

The Decision Field Model, as proposed by Busemeyer and Townsend (1993) is similar to EU except insofar as the error term is heteroscedastic. So the difference between two lotteries is valued exactly as in EU but the error variance is not constant. To define it we have to introduce some extra notation. Consider a choice between a Left lottery which yields outcomes O_a^L, O_b^L and O_c^L in the states (colours) *a, b* and *c* respectively, and a Right lottery which yields outcomes O_a^R, O_b^R and O_c^R in these states (colours). Then the difference between the two lotteries is evaluated, as in SEU, (using an obvious notation) by

$$V(L)-V(R)+\varepsilon = p_a[U(S_a^L)-U(S_a^R)]+p_b[U(S_b^L)-U(S_b^R)]+p_c[U(S_c^L)-U(S_c^R)]+\varepsilon$$

where the error term ε has a normal distribution with mean 0 and variance given by:

$$s^2\{p_a[U(S_a^L)-U(S_a^R)]^2+p_b[U(S_b^L)-U(S_b^R)]^2+p_c[U(S_c^L)-U(S_c^R)]^2-[V(L)-V(R)]^2\}$$

Here we estimate *u,* the (additive) probabilities and *s*. Note that this model has exactly the same number of parameters as EU but neither is nested inside the other.

6. Maximin Expected Utility Theory (GS Min)

This is a special case of Alpha theory (described below) with the parameter α of that theory equal to 1.

7. Maximax Expected Utility Theory (GS Max)

This is a special case of Alpha theory (described below) with the parameter α of that theory equal to 0.

8. Alpha Theory

In this theory, decision-makers are envisaged as thinking of the probabilities (of the various events) lying in some convex space. Denoting this convex space by *P,* each point in which represents a possible probability for each of the three colours (necessarily additive), then, according to this theory, the decision-maker chooses between lotteries on the basis of the maximum of

$$\alpha \min_{p \in P}[EU(L)] + (1-\alpha) \max_{p \in P}[EU(L)] \qquad (6)$$

The convex set *P* is individual specific. We parameterise that and estimate these parameters, in addition to the utility parameter *u*. However, the theory does not specify how it should be parameterised. We assumed that this convex space can be represented as a convex area within the triangle defined by the vertices (0,0), (1,0) and (0,1) in a space with the probability of colour *a* on the horizontal axis and the probability of colour *b* on the vertical axis. In order to make the convex space symmetrical as far as the treatment of the three probabilities were concerned, we characterised it as bounded by a vertical line at $\overline{p}_a$, a horizontal line at $\overline{p}_b$ and a line parallel to the hypoteneuse (with therefore a slope of -45°) such that $1 - p_a - p_b = \overline{p}_c$. We estimate $\overline{p}_a$, $\overline{p}_b$ and $\overline{p}_c$ along with the other parameters. We note that this convex space can take a variety of different forms. Clearly if it just consists of a single point then the Alpha model reduces to EU.

9. MaxMin

In this, the decision-maker is presumed to follow the rule of choosing the lottery for which the worst outcome is the best. We assume that the rule is followed lexicographically, so that we get the following rule, where l_1, l_2 and l_3 denote the three outcomes on one of the two lotteries, L, *ordered from the worst to the best*, and m_1, m_2 and m_3 denote the outcomes on the other lottery, M, also ordered from the worst to the best:

$$\begin{aligned}
&\text{if } l_1 > m_1 \text{ then } L \succ M \\
&\text{if } l_1 < m_1 \text{ then } L \prec M \\
&\text{if } l_1 = m_1 \text{ and } l_2 > m_2 \text{ then } L \succ M \\
&\text{if } l_1 = m_1 \text{ and } l_2 < m_2 \text{ then } L \prec M \\
&\text{if } l_1 = m_1,\, l_2 = m_2 \text{ and } l_3 > m_3 \text{ then } L \succ M \\
&\text{if } l_1 = m_1,\, l_2 = m_2 \text{ and } l_3 < m_3 \text{ then } L \prec M \\
&\text{if } l_1 = m_1,\, l_2 = m_2 \text{ and } l_3 = m_3 \text{ then } L \sim M
\end{aligned} \tag{7}$$

We note that there are no parameters to be estimated in this model, though we do assume that the decision-maker ranks £100 as the best outcome, £10 as the second best and -£10 as the worst.

10. MaxMax

In this the decision-maker is presumed to follow the rule of choosing the lottery for which the best outcome is the best. We assume that the rule is followed lexicographically, so that we get the following rule, using the same notation as above:

$$\begin{aligned}
&\text{if } l_3 > m_3 \text{ then } L \succ M \\
&\text{if } l_3 < m_3 \text{ then } L \prec M \\
&\text{if } l_3 = m_3 \text{ and } l_2 > m_2 \text{ then } L \succ M \\
&\text{if } l_3 = m_3 \text{ and } l_2 < m_2 \text{ then } L \prec M \\
&\text{if } l_3 = m_3,\, l_2 = m_2 \text{ and } l_1 > m_1 \text{ then } L \succ M \\
&\text{if } l_3 = m_3,\, l_2 = m_2 \text{ and } l_1 < m_1 \text{ then } L \prec M \\
&\text{if } l_3 = m_3,\, l_2 = m_2 \text{ and } l_1 = m_1 \text{ then } L \sim M
\end{aligned} \tag{8}$$

Again there are no parameters to estimate.

11. Minimax Regret[24] (MinReg)

With this preference functional, the decision-maker is envisaged as imagining each possible ball drawn, calculating the regret associated with choosing each of the two lotteries, and choosing the lottery for which the maximum regret is minimized. Again there are no parameters to estimate, though it is assumed that there is a larger regret associated with a larger difference between the outcome on the chosen lottery and the outcome on the non-chosen lottery.

12. ***Hurwicz***

In this the decision-maker is presumed to follow the rule of choosing the lottery for which a weighted average of the worst outcome and the best outcome is the best. As this rule implies some way of comparing outcomes, a utility function (characterised as in EU) is required. We estimate u along with the weight α attached to the worst outcome.

[24] See Luce and Raiffa (1957).