J Risk Uncertain (2009) 38:1–25
DOI 10.1007/s11166-008-9058-5

Naive, resolute or sophisticated? A study of dynamic decision making

John D. Hey · Gianna Lotito

Published online: 10 January 2009

Abstract Dynamically inconsistent decision makers have to decide, implicitly or explicitly, what to do about their dynamic inconsistency. Economic theorists have identified three possible responses—to act naively (thus ignoring the dynamic inconsistency), to act resolutely (not letting their inconsistency affect their behaviour) or to act sophisticatedly (hence taking into account their inconsistency). We use data from a unique experiment (which observes both decisions and evaluations) in order to distinguish these three possibilities. We find that the majority of subjects are either naive or resolute (with slightly more being naive) but very few are sophisticated. These results have important implications for predicting the behaviour of people in dynamic situations.

Keywords Dynamic decision making · Naivety · Sophistication · Resoluteness · Dynamic inconsistencies

JEL classifications D90 · D80 · C91

This paper is concerned with dynamic decision-making[1]. An important and recurring issue in this analysis concerns the behaviour of dynamically inconsistent people. Do they know that they are dynamically inconsistent, and, if so, what do they do about it? Economic theory has identified three possible responses (though there are obviously many more): that such decision makers act *naively* (ignoring their inconsistency); that

[1]See, amongst others, Cubitt et al. (2004), Machina (1989), McClennen (1990).

J. D. Hey
LUISS, Rome, Italy

J. D. Hey (✉)
Department of Economics & Related Studies, University of York, Heslington, York YO10 5DD, UK
e-mail: jdh1@york.ac.uk

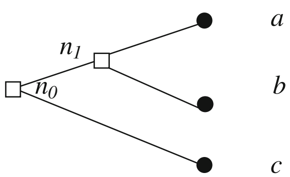
G. Lotito
Università del Piemonte Orientale, Alessandria, Italy

they act *resolutely* (not letting their inconsistency affect their behaviour); that they act *sophisticatedly* (taking their inconsistency into account). We report on an experiment that lets us infer which of these responses describes behaviour better. We have designed the experiment in such a way that we can not only observe choices in dynamic decision problems but also we obtain subjects' evaluations of such problems. Combining these two types of data we can estimate the preferences of the decision makers, and crucially infer whether they are naive, resolute or sophisticated.

Dynamic decision-making has two dimensions: the *sequentiality* of the decision process, and the *passage of real time*. In both dimensions, the issue of dynamic consistency arises: whether decision-makers implement plans that they made earlier. A key element is whether the preferences of the decision maker change through the decision-making process. If preferences do change, the potentiality of dynamic inconsistency arises. This may happen in several ways depending upon the context. In the context of a risk-free problem with no passage of real time, this can only occur if preferences explicitly change during the decision process. In other contexts the reasons are more subtle. In the context of a risky sequential problem, potential dynamic inconsistency may occur if the preferences of the decision-maker are not those of Expected Utility theory. In the context of a risk-free problem with the passage of real time, potential dynamic inconsistency may occur if the preferences of the decision-maker are not those of (exponentially) Discounted Utility theory. Our analysis is relevant to all contexts, though the context which we examine is that of sequential decision-making under risk.

The problem of inconsistent choice in a dynamic decision problem was initially analysed in the literature in a context of certainty, and was related to the problem of preferences changing either endogenously or exogenously through time (see, for example, Strotz 1956). Hammond (1976 and 1977) generalised the analysis, overcoming the distinction between exogenously and endogenously changing tastes, and concentrating the analysis on the essential aspect of the problem—namely that preferences change over time. However he kept it confined to a situation without risk or uncertainty. This was introduced only later: Hammond (1988a, b, 1989), Raiffa (1968), Machina (1989) and McClennen (1990).

The simple example of the drug addict in Hammond (1976) illustrates the problem. Suppose that an individual is considering whether to start taking an addictive drug. The individual[2] would prefer at most to take the drug without consequences. However, he is certain that, if he starts, he will become an addict, with serious consequences for his health. Of course, he can refuse to take the drug in the first place. This agent is facing a simple dynamic decision problem with the following structure (squares representing decision nodes):



[2] We presume that he is male to avoid expositional clumsiness.

Here three options are available to the agent, which lead to the following outcomes: take the drug as long as it is harmless, then stop, which leads to outcome a; become an addict, which leads to outcome b; not take the drug, which leads to outcome c.

At the initial decision node n_0 the agent has to decide whether to take the drug or not, and his preferences are $a \succ c \succ b$. If he gets to the choice node n_1 he has become an addict, and therefore the only relevant preferences are those concerning a and b, and addiction itself means that $b \succ a$. Thus, at n_1 his initial preferences between a and b get reversed. The agent will choose b inconsistently with his previous preference for reaching a.

The requirement of *dynamic consistency* is a requirement of consistency between planned choice and actual choice. In the example, the dynamically inconsistent agent decides initially to take the drug, then stop (option a), but chooses later not to stop (thus actually choosing option b).

Hammond (1976) considers *myopic* (*naive*) and *sophisticated choice* as two possibilities available to an agent in such a situation of dynamic inconsistency. When acting according to a *myopic approach*, the agent selects at each point those strategies which he judges acceptable from the perspective of that point. In the example, the myopic agent ignores that his tastes are changing, and chooses at each stage the option he considers as best at that moment. Therefore, he will choose option a at n_0, but change his mind and choose b at n_1. The final outcome will be b.

When acting according to a *sophisticated approach*, the agent anticipates his future choice and chooses the best plan among those he is ready to follow to the end: he rejects those plans which imply a choice he anticipates he will not make. By doing so, he always ends up choosing *ex post* according to his *ex ante* plans, and avoids violating dynamic consistency. In the example, at n_0 he will forecast that by taking the drug he will become an addict, and realises that his only options are b and c. Therefore he will choose the most preferred option between the two, that is, c.

Hammond's example of dynamically inconsistent choice allows us to consider another possible model of behaviour, which was formalised only later in the literature—McClennen (1990), Machina (1989)—in the context of dynamic inconsistency under risk—*resolute choice*. The agent resolves to act according to a plan judged best from an *ex ante* perspective, and intentionally acts on that resolve when the plan imposes on him *ex post* to make a choice he does not prefer at that point. By so choosing he manages to act in a dynamically consistent manner. In Hammond's example, at n_0 a resolute agent would have resolved to act according to the plan leading to the most preferred outcome a—take the drug while it is harmless, then stop; and he would have acted on that resolve when at n_1 the plan imposed on him to choose the less preferred option—therefore going for a and not for b.

So far we have introduced the problem of dynamic inconsistencies in the context of changing preferences in a world of certainty. As mentioned above, there are however other contexts in which the same considerations apply. One is in a dynamic certain world in which agents discount the future non-exponentially—for example, using quasi-hyperbolic discounting, as in Harris and Laibson (2001). We note that Harris and Laibson remark (p. 939) that "we model an individual as a sequence of autonomous temporal selves", explicitly making clear that preferences are changing through time, in the sense that the *relative* evaluation of consumption in any two periods varies depending upon when the evaluation is carried out. In this paper, and indeed in much of the literature on quasi-hyperbolic consumers, the analysis assumes that the decision maker is sophisticated in the sense that we have used it above. Hence the decision maker takes into account his or her decisions in the future when deciding in the present. In this way, the decision maker resolves his or her dynamic inconsistency. However, it could also be the case that the decision maker acts either naively or resolutely[3]—in which cases the predictions of the quasi-hyperbolic model would be different.

A further dynamic context in which exactly the same considerations apply is that of a risky world in which agents do not have Expected Utility preferences. This is the context in which we carry out our experimental investigation. We begin to give detail in Section 1.

However, we have exactly the same interest in all contexts: how does a dynamically inconsistent agent respond to, and resolve, his or her dynamic inconsistencies: naively, resolutely or sophisticatedly[4]? We attempt to answer this question by adopting an experimental design where the subjects have to take a series of decisions, some of which are dynamic in nature, and in which different kinds of inconsistent agent respond in different ways to the decision problems posed. Subjects are required both to take decisions and also to evaluate decision problems. By observing their behaviour (both over decisions and evaluations) we can infer whether they are (more likely to be) naive, resolute or sophisticated. Interestingly we find that the majority of our subjects appeared to be either naive or resolute (with roughly equal numbers of each), while very few appeared to be sophisticated. This finding has important implications for the modelling of dynamically inconsistent agents (for example, as in the literature on the behaviour of agents who hyperbolically discount the future).

1 Dynamically inconsistent behaviour in a sequential risky context

In the context of risk and uncertainty, the problem of dynamic inconsistency is crucially linked to the question of whether the decision maker has Expected Utility preferences or not. As is well known, if the decision maker's preferences satisfy Expected Utility theory, then the problem of dynamic inconsistency does not arise. For a non-Expected-Utility decision maker, however, the problem may arise[5]. We illustrate this in Fig. 1, where three of the decision problems played in the experiment are shown. In this figure, the squares (green in the experiment) represent decision nodes; and the circles (red in the experiment) represent chance nodes (where Nature moves with the probabilities indicated in the figure). The letters on the various choice branches denote

[3] Or in some other way—these three ways of reacting to the potential dynamic inconsistency are not the only ones.

[4] We should note that the issue could also be regarded as one of *framing*. As we shall see later when we discuss the actual decision problems used in the experiment, different ways of framing the problem might lead to different solutions, for a potentially dynamically-inconsistent person. Indeed, it could be argued that dynamic-consistency and immunity-to-framing-effects are one and the same thing.

[5] See Machina (1989) for a detailed description of the "argument for the dynamic inconsistency of the non-expected utility maximisers" (p. 1636). Other references include Raiffa (1968) and McClennen (1990).

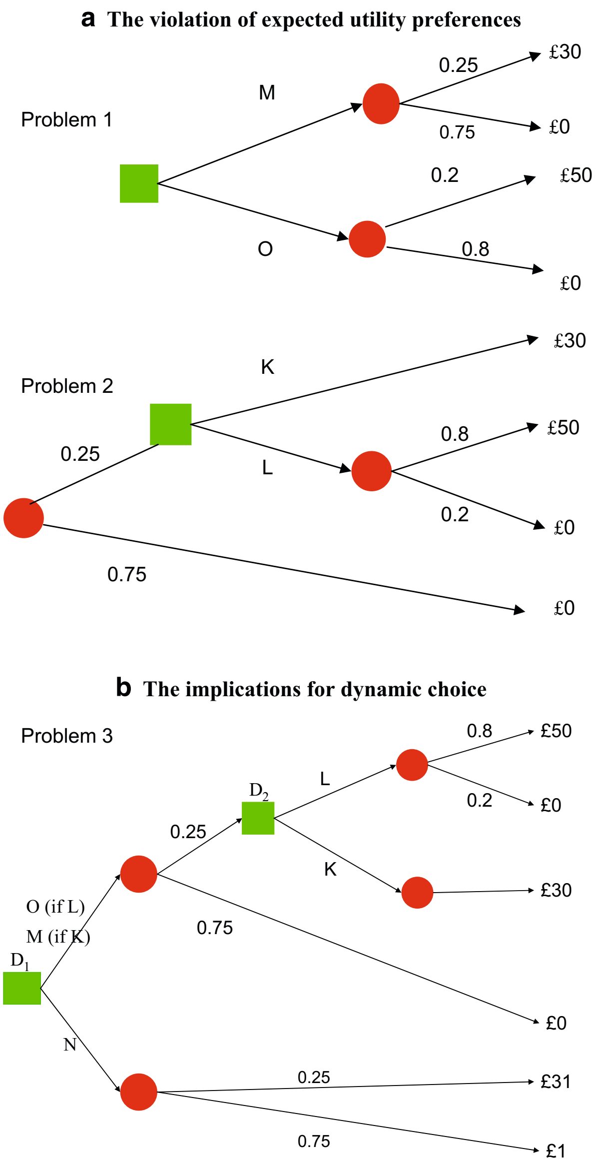


Fig. 1 The problem of dynamic inconsistency. **a** The violation of expected utility preferences. **b** The implications for dynamic choice

the implied lottery: for example, in Problem 1, *M* indicates the lottery obtained choosing Up and *O* the lottery obtained by choosing Down; at the decision node of Problem 2, *K* is the certainty of £30, while *L* is the lottery which gives £50 with probability 0.8 and £0 with probability 0.2. In Problem 3[6], *O* denotes the lottery obtained by playing Up at both decision nodes; *M* denotes the lottery obtained by

[6] A tree with a similar structure is found in McClennen (1990).

playing Up at the first and Down at the second; while N denotes the lottery obtained by playing Down at the first decision node.

Consider first Fig. 1a and consider an individual[7] who prefers O to M in Problem 1 and prefers K to L in Problem 2. These preferences imply a violation of Expected Utility theory in the form of a Common Ratio Effect (see Starmer (2000) for references concerning empirical evidence of this effect)[8]. Now consider how such an individual would tackle Problem 3, in which there are two decision nodes D_1 and D_2.

Suppose this individual is at decision node D_2. Then her preferences indicate that she would choose to move Down at that node, because she prefers K to L. Now look at the situation as viewed from node D_1. As viewed from there, if she moves Up at that first node, then she is faced with either getting O (by moving Up at D_2) or getting M (by moving Down at D_2); if she moves Down she is faced with getting N. Assume that for this individual O is preferred to N, which is preferred to M[9]. Then as viewed from node D_1 she prefers O, and that, by assumption, is preferred to the lottery obtained by moving Down at D_1 namely N. Hence she sets off by choosing Up at D_1. A problem arises, however, if she arrives at node D_2. As we have already argued, at this node Down is preferred, and so, if the preferences at that point are followed, the individual will choose Down at node D_2. Hence the individual, in arguing in this way, plans, at node D_1, to choose Up at node D_2 but, if she arrives there, actually chooses Down. This is dynamic inconsistency: adopting a plan but then not implementing it at a later node. Moreover, it is a problem to the individual when at node D_1. If she is aware of this dynamic inconsistency, then she realises that by choosing Up at D_1 and then choosing Down at D_2 she therefore ends up with the lottery M which is dominated by the lottery N—which she could get by choosing Down at D_1.

What might the individual do about this dynamic inconsistency? Well, she might simply be unaware of it, or ignore it, acting *naively*, and choose at each point the strategy most preferred from the perspective of *that* point—Up at D_1 and then Down at D_2. However, if she is aware of this dynamic inconsistency, she might want to do something about it. One possibility is that the individual, anticipating the fact that she will choose Down at D_2, if she reaches it, realises that it would be better to choose Down at node D_1. This, in the literature, is termed acting *sophisticatedly*—see, for example, McClennen (1990) and Machina (1989)[10]. This sophisticated individual anticipates her future behaviour, and avoids the inconsistency by making a choice at the initial decision node which is constrained by her anticipated choice at each following node.

However, as viewed from the individual at node D_1 this behaviour implies the lottery N whereas, as viewed from D_1, a better lottery is O which is achievable by choosing Up at D_1 and Up at D_2. The individual at D_1 might accordingly decide to

[7] Whom we assume here is female to avoid expositional clumsiness.

[8] The Independence Axiom of Expected Utility theory implies that £30 for sure is preferred to a 80:20 lottery over £50 and £0 if and only if a 25:75 lottery over £30 and £0 is preferred to a 20:80 lottery over £50 and £0.

[9] Note that lottery N stochastically dominates lottery M.

[10] Machina though does not use directly the term sophisticated choice for this kind of approach. Also Karni and Safra's (1989) model of "behavioural consistency" represents a way of implementing the sophisticated choice approach, and represents a solution to the problem of dynamic inconsistency with non-linear preferences.

be *resolute* and choose Up at D_1 and Up at D_2. This resolute individual avoids inconsistency by making the choice of a plan most preferred at the initial node to constrain future behaviour. She resolves to implement the plan originally adopted, despite the fact that this implies, at some future node, making a choice that she would not have liked to make once arrived at that node. Again a discussion of this kind of behaviour can be found in Machina (1989)[11] and McClennen (1990)[12].

Such problems do not arise with Expected Utility individuals. In the context of the above decision problem, EU preferences could be such that K is preferred to L and M to O, or the reverse. In the first case, the individual chooses Down at node D_1 and in the reverse case chooses Up at D_1 and Up at D_2.

2 Related experiments

The purpose of the study reported in this paper is to try and find out whether subjects are naive, resolute or sophisticated. To the best of our knowledge, this is the first such study. However, there are some related experiments, which also study behaviour in dynamic contexts. We shall confine ourselves to those in economics which have an appropriate incentive mechanism[13]. One such paper is that by Cubitt et al. (1998), which is primarily concerned with trying to discover, in the context of dynamic decision making, which of a set of dynamic choice principles is apparently violated by individuals with non-EU preferences. Cubitt et al. presented (to different groups of subjects) different decision trees and studied the behaviour of subjects in such trees. These trees show some similarities with those in Problems 1, 2 and 3 shown in Fig. 1. A sub-set of the trees in Cubitt et al. was used in Hey and Paradiso (2006) who, instead of looking at the *decisions* of subjects, collected data on subjects' *evaluations* of decision problems. The two studies are closely related though they had different objectives. These three sets of trees are relevant to the objectives of this present paper since naive, resolute and sophisticated subjects could have different behaviours in the trees, as well as having different evaluations of them. Cubitt et al. found that subjects behave differently in the different trees, while Hey and Paradiso find that the temporal frame also affects the evaluations the subjects make of the different trees.

Our experiment differs from both these previous studies, most particularly because this present paper has a different agenda. However, we build on the framework set by these two papers by using similar decision trees to those used in Cubitt et al. and in Hey and Paradiso. We also add a fourth tree which helps us in our task of distinguishing between naive, resolute and sophisticated subjects. In addition, our experiment differs from both these other two in that we observe not only choices but also gather data on valuations[14]. Therefore in our experiment we combine these

[11] Machina's model of choice is equivalent even if differently formalised to McClennen's model of resolute choice. According to Machina, resolute choice represents one of the "antecedents of the formal approach" presented in his paper.

[12] An interesting point is how this person can force him- or herself to behave resolutely.

[13] Busemeyer and his associates in psychology have carried out some related experiments (see, for example, Busemeyer et al. 2000) but without such incentives.

[14] We note that valuation data is potentially more informative than choice data, as the latter only tells us which choice is preferred and nothing about the strength of preference.

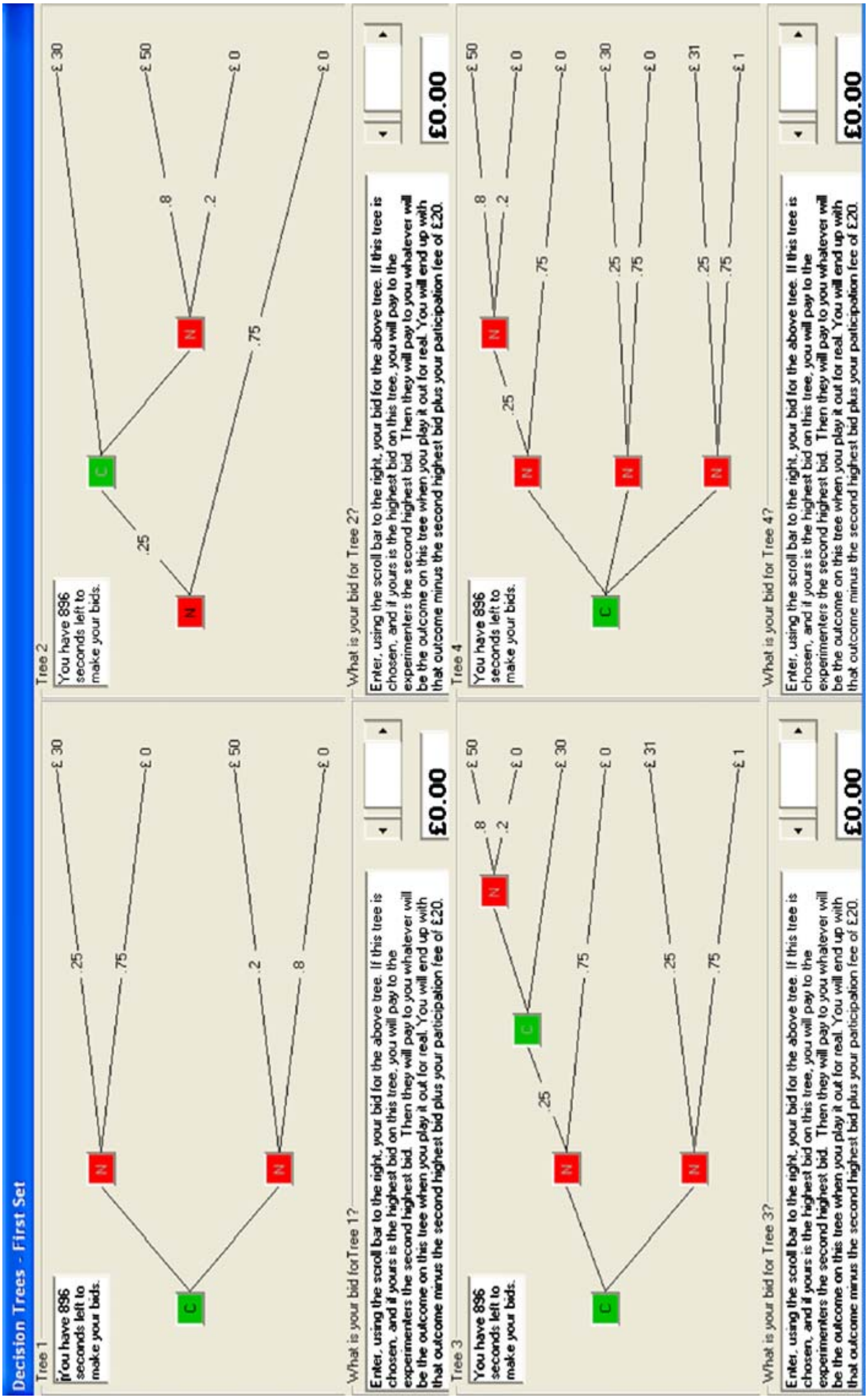

Decision Trees - First Set
Tree 1
You have 896 seconds left to make your bids.
What is your bid for Tree 1?
Enter, using the scroll bar to the right, your bid for the above tree. If this tree is chosen, and if yours is the highest bid on this tree, you will pay to the experimenters the second highest bid. Then they will pay to you whatever will be the outcome on this tree when you play it out for real. You will end up with that outcome minus the second highest bid plus your participation fee of £20.
£0.00
Tree 2
You have 896 seconds left to make your bids.
What is your bid for Tree 2?
Enter, using the scroll bar to the right, your bid for the above tree. If this tree is chosen, and if yours is the highest bid on this tree, you will pay to the experimenters the second highest bid. Then they will pay to you whatever will be the outcome on this tree when you play it out for real. You will end up with that outcome minus the second highest bid plus your participation fee of £20.
£0.00
Tree 3
You have 896 seconds left to make your bids.
What is your bid for Tree 3?
Enter, using the scroll bar to the right, your bid for the above tree. If this tree is chosen, and if yours is the highest bid on this tree, you will pay to the experimenters the second highest bid. Then they will pay to you whatever will be the outcome on this tree when you play it out for real. You will end up with that outcome minus the second highest bid plus your participation fee of £20.
£0.00
Tree 4
You have 896 seconds left to make your bids.
What is your bid for Tree 4?
Enter, using the scroll bar to the right, your bid for the above tree. If this tree is chosen, and if yours is the highest bid on this tree, you will pay to the experimenters the second highest bid. Then they will pay to you whatever will be the outcome on this tree when you play it out for real. You will end up with that outcome minus the second highest bid plus your participation fee of £20.
£0.00

Fig. 2 A screen shot from the experiment—*squares labelled C* (*green* in the experiment) are choice nodes; *squares labelled N* (*red* in the experiment) are chance nodes

two other approaches but with a differently directed research agenda—that of detecting whether subjects are naive, resolute or sophisticated.

Our experiment is built around the four trees shown in Fig. 2 (which is a screen shot from our experiment—the colour version of which can be found in the Web Technical Appendix). Trees 1, 2 and 3 are identical to Problems 1, 2 and 3 shown in Fig. 1. Tree 4 is the static version of Tree 3, where the decision problem is reduced to that of a single decision—rather than that of a sequence of decisions. We note that Expected Utility theory has a strong prediction about relative behaviour in Trees 1 and 2; given that, as we have discussed above, we have predictions about the behaviour in Tree 3 of naive, resolute or sophisticated subjects whose preferences do not respect Expected Utility theory.

In essence, observing behaviour in these trees should enable us to detect whether people are naive, resolute or sophisticated. However, there are two problems. The first is that we can only hope to detect differences in type if the agent has non-EU preferences—since the distinction only has sense for non-EU agents. Moreover, it may be the case that a subject has non-EU preferences but our trees do not enable us to detect this fact: a subject who chooses M in Problem 1 and K in Problem 2 (or O in Problem 1 and L in Problem 2) is not necessarily a person with EU preferences. A second problem is that, inevitably in experiments, there is *noise* in the subjects' responses (see Hey 2005). It is well known from many previous experiments that this noise can be quite large; to give some idea of this, we note that, when asked the same question on two separate occasions, subjects give different answers roughly 25–35% of the time. Accordingly we cannot be sure that any stated decision (or any stated valuation) is *exactly* in accordance with the subject's preferences.

To help us to get over these two problems, we used *three* sets of four trees—all with the same structure as the four trees in Fig. 2 but with different probabilities and payoffs. We give detail in the next section (Section 3). Moreover, in analysing the data from the experiment, after a descriptive analysis (Section 4), we explicitly assumed the existence of noise in the responses of the subjects and we use all the data on each subject to analyse the preferences of that subject. We give detail in Section 5. For all the technical material we refer to our Web Technical Appendices at http://www.luiss.it/hey/NRS/NRS_Technical_Appendix.pdf.

3 The experimental design and implementation

The experiment was built on the four decision problems introduced in the previous section. We used three sets each with four trees, with different values for the payoffs and the probabilities, giving a total of 12 decision trees. Table 1 gives the details; we note that the possible payoff varies from a minimum of £0 to a maximum of £150. We designed the experiment in such a way that we could observe subjects' choices in all 12 trees as well as their stated evaluations of the 12 trees. The amounts of money and the probabilities in the three sets of trees were chosen to satisfy various criteria. First, we wanted one set of trees (Set 1) that was the same as used in

Table 1 The parameters used in the experiment

Parameter	Set 1	Set 2	Set 3
a	£50	£150	£60
b	£31	£51	£41
c	£30	£50	£40
d	£1	£1	£1
e	£0	£0	£0
q	0.8	0.5	0.75
r	0.25	0.1	0.20

previous experiments (Cubitt et al. 1998 and Hey and Paradiso 2006). Second, we wanted different amounts of money, and different probabilities in the different sets, but with the same properties. Hence the structure of the payoffs and probabilities are identical in the three sets—they are always such that we should be able to detect non-EU behaviour from the choices, and then to distinguish the naive, resolute and sophisticated types. Finally we wanted payoffs and probabilities that gave an incentive to the subjects that was roughly similar in the three different sets (the maximum expected payoff ranges from £7.50 to £10 in the different sets) as well as being reasonable. Here there is a problem in that we preferred that subjects would not leave the experiment with less money than that with which they came. Obviously there is no way to guarantee this, but giving a relatively high participation fee reduces this possibility—and also increases the amount of money that they can bid for the various trees. For a risk-neutral subject, the value of a randomly chosen tree in the different sets is £8.83, so while the value of bidding for a tree was not high, it was also not negligible. Moreover subjects were encouraged to think seriously about their bids by having to spend 15 min deciding on them, which turned out to be a more than sufficient length of time.

The experiment was conducted at *EXEC*, the Centre for Experimental Economics at the University of York. A total of 50 students, both graduate and undergraduate, took part in the experiment. They were given written instructions (see Web Technical Appendix 5). When all participants had finished reading the instructions, a PowerPoint presentation was played at a predetermined speed on their individual screens. After this, they could ask questions. The experiment then started, using a Visual Basic program (available on request). The computer screen showing the four decision trees is that shown in Fig. 2. In order to elicit the subjects' evaluations for each of the 12 trees, we used the second-price sealed-bid auction method as in Hey and Paradiso (2006). This was implemented as follows. Subjects performed the experiment in groups of five. They were sat at individual computer terminals and were not allowed to communicate with each other. They individually made bids for each of the three sets of four trees (12 trees overall) and were given 15 min to bid for each set of four trees in the set. During the bidding period the subjects were allowed to practice playing out the decision trees as much as they wanted in the time allowed. It was made clear that the outcomes of the practice did not affect their payments in any way. The number of seconds left for the practice and bidding was shown in the box at the top left-hand side of each decision tree. When the bidding time was over, the subjects played out all the 12 problems for real. We displayed on each subject's screen the results of his or her playing out plus the bids of all the five subjects in the group for each of the 12 trees. Then we invited one of the five subjects to select a ball at random from a bag containing 12 balls numbered from 1 to 12. This determined the problem on which the auction was held. We then consulted the bids made by the group members, and the subject with the highest bid for the problem paid us the bid of the second-highest bidder. As all subjects were given a £20 participation fee, four of the five members earned £20, while the fifth earned £20 minus the bid of the second highest bidder plus the outcome. Both in the instructions and in the presentation, it had been emphasised, through different examples, that the highest bidder (in the auction corresponding to the randomly selected decision problem) would have to pay the bid of the second-highest bidder, and it was made clear that the bid for each of the 12 problems should be equal to the willingness to pay for the decision problem. It was emphasised that in the case that the subject's bids were all £0, he or she was unlikely to be sold one of the decision problems and thus will definitely end up with no less than the participation fee. Moreover, in the case that the bid was higher than £0 and the subject's bid was the highest for the chosen tree, he or she would end up with the participation fee, minus the bid of the second-highest bidder, plus the outcome. This could—depending on the bid—be less than the participation fee and the subject could end up losing money. Subjects were warned in advance that they had to bring enough cash to the experiment to allow for this possibility, which however never occurred[15].

Before proceeding to the analysis of the data from the experiment, we should comment on one feature of the design: the (implicit) use of the *random lottery incentive mechanism*: while subjects took decisions with respect to, and placed bids on, 12 trees, the payment mechanism implied that there was an average 1/60 probability[16] that the decisions and bids on any tree would actually influence their payment. This kind of random lottery incentive mechanism has a long tradition in experimental economics, but implies that the subjects treat each tree as independent from the other trees—that is, that they *separate* the trees one from the other when they are taking decisions. In the context of dynamic choice, this assumption might appear to be unrealistic, but the opposite assumption—that the subjects take into account all 12 trees when taking their decisions on any one—might be considered as even more unrealistic. However, there is considerable evidence that this random lottery mechanism does have the desired property: a direct test was first carried out by Starmer and Sugden (1991) with just two choice problems; an indirect test with 100 choice problems was implemented by Hey and Lee (2005a); a further test of the proposition that subjects may not take into account all the questions when answering any one, but instead take into account the recently answered questions was carried out by Hey and Lee (2005b). All three of these studies seem to confirm that subjects do separate and hence that they answer to each question (each tree in our context) as if that were the only question (tree) in the experiment.

[15] We take this as a sign that the subjects understood the second-price sealed-bid auction method. Besides, the raw data on the subjects' bids (see Web Technical Appendix 1) show that no case of under or over-bidding occurred.

[16] Recall that there were five subjects in each group, implying an average probability of 1/5 that any subject would be paid from playing out a particular tree, and that there were 12 trees, each chosen with probability 1/12.

4 Description of the data from the experiment

We recall from Section 1 that the experiment design allows us to classify subjects according to whether their preferences obey EU or not by observing their behaviour in trees T_1 and T_2. Four patterns of behaviour can be observed on these two trees:

1. K preferred to L, and M to O
2. L preferred to K, and O to M
3. K preferred to L, and O to M
4. L preferred to K, and M to O

The first two patterns are subjects consistent with EU preferences. The last two classify subjects who violate EU; pattern (3) represents the more common violation. Table 2 gives the number of subjects who behaved according to the above patterns in the different sets.

For all the four patterns we can predict the subjects' behaviour in T_3 and T_4. In cases (1), (2) and (4) there should be no differences in behaviour for the sophisticated, naive and resolute subjects, while in case (3) there should be a difference. Therefore, for pattern (3) we can also classify how many subjects are consistent with the predictions of the different dynamic choice models (assuming no error). Table 3 gives the number of subjects who behave consistently with the predictions in T_3 and T_4 in each set.

We note that the rate of consistency to predicted behaviour—particularly of non-Expected-Utility subjects—is low. It is interesting to note also that nearly all inconsistent behaviour in case (3) followed a common pattern: subjects chose prospect N in both T_3 and T_4. Choice of the N gamble is consistent with sophistication in T_3, but not in T_4, where all subjects with this non-EU preference pattern should have preferred gamble O.

The experiment design allows also the prediction of how subjects with different preference patterns should evaluate the different trees, by assigning higher bids to the trees which they value more. Table 4 gives—for each set and for each different preference pattern—the predictions in terms of bids and the number of subjects who were consistent with such predictions. Again we note that also consistency to the predictions on subjects' evaluations appears to be very low. This is almost certainly because of error in the subjects' responses, which this descriptive analysis does not take into account.

Table 2 Number of subjects with EU and non-EU preferences in each set

		Set 1	Set 2	Set 3
Expected utility	K preferred to L, hence M to O	12	12	15
	L preferred to K, hence O to M	19	16	11
	Total number	31	28	26
Non-expected utility	K preferred to L, but O to M (*common violation*)	15	18	18
	L preferred to K, but M to O	4	4	6
	Total number	19	22	24

Table 3 Number of subjects with consistent behaviour, and total number, in each set

			Set 1		Set 2		Set 3	
			consistent	**total**	**consistent**	**total**	**consistent**	**total**
Expected Utility	***K*** and ***M***		9	12	8	12	8	15
	L and ***O***		11	19	5	16	7	11
Non-Expected utility	***K*** but ***O*** (common violation)	**Naïve**	1	15	2	18	1	18
		Resolute	2		4		2	
		Sophisticated	1		3		1	
	L but ***M***		0	4	2	4	2	6

5 A formal analysis

There are two problems with the above description of the data. First, it is partial and does not use all the data from each subject in a systematic fashion. Second, as we have already noted, it ignores the existence of noise in the subjects' responses. However, there is noise in experimental data. This creates an immediate problem in that subjects' behaviour might suggest that they were naive on Set 1, resolute on Set 2 and sophisticated on Set 3[17]. What would we conclude then? More seriously, because of their noise, we cannot be sure whether behaviour in Trees 1 and 2 actually does reveal EU or non-EU preferences. They may be EU but choose Up in Tree 1 and Down in Tree 2 because of this error; similarly they may be non-EU and yet choose Up in both Trees 1 and 2 (either with or without error). We need to take into account of this noise. What we do then, rather than carry out a series of individual tests on the data which require the assumption of zero noise, we use all the data to estimate models of behaviour. Doing so enables us to use all the data on each subject (24 observations), and, in particular, to try and discover the type (naive, resolute or sophisticated) of each subject. Methodologically, it seems better to use all the data on each subject to try and fit the various types rather than to carry out a series of individual tests on bits of the data.

At this point it is necessary to note that we deliberately fixed the random number generator in the real playing-out of the Trees 2 and 3, so that we could have observations on all the subjects' choices later in the tree. This is deception, and we feel that we should comment on our use of it. Although we are, like all experimental economists, averse to practicing deception, we note two things[18]: first, that the

[17] Recall that subjects were presented with three sets of trees with the structure as in Fig. 2.

[18] We should also note that one of the referees asked us to record that he or she "judged that this feature of the design was both unethical and unnecessary". We agree with the 'unethical' but would argue that it was necessary to increase the number of useful observations and save expense on paying the subjects: if we had not used this deception, we would have had to have many more subjects in the experiment.

Table 4 Predicted bids and number of subjects consistent with the predictions, and total number, in each set

		Predicted bids	**Set 1**		**Set 2**		**Set 3**	
Expected Utility			**consistent**	**total**	**consistent**	**total**	**consistent**	**total**
	K **and** ***M***	*T1=T2 <T3=T4*	*4*	*12*	*5*	*12*	*4*	*15*
	L **and** ***O***	*T1=T2 =T3=T4*	*3*	*19*	*1*	*16*	*3*	*11*
Non-Expected Utility	***K*** **but** ***O*** (common violation)	Naïve or Resolute *T1=T2 =T3=T4*	*1*	*15*	*2*	*18*	*3*	*18*
		Sophisticated *T1=T4 >T3>T2*	*0*		*1*		*0*	
	L **but** ***M***	*T1=T2 <T3=T4*	*0*	*4*	*0*	*4*	*2*	*6*

deception was in the subjects' interests in that they were no worse off with the deception than without it; second, that without this deception we would have had significantly fewer observations, crucially those which enable us to distinguish between the various types (naive, resolute and sophisticated). We did not, however, fix the generator during the practice playing out.

Our data analysis strategy is the following. We assume that subjects are different, both in terms of their preferences and in terms of their *type* (naive, resolute or sophisticated). We fit types to the data from each subject individually. We use all the data (24 observations) on each subject with our ultimate goal of telling whether the subject is (more likely[19] to be) naive, resolute or sophisticated.

While the analysis of the data on decisions only requires us to know preferences over a relatively small number of lotteries, the analysis of the data on evaluations of the trees requires us to know preferences not only over these lotteries but also over all the evaluations of these lotteries (in monetary terms by the subject). For each subject, there are 12 of these evaluations (the bids made for each of the 12 trees). In

[19] We note that the presence of this error means that inevitably our inferences can not be certain.

each set of four trees, there are ten lotteries, involving five amounts of money. We neither know which type the subject is, nor the preference functional of the subject. This has to be estimated. Given that we have just 24 observations per subject, it is clear that we cannot estimate the evaluation of each of the 30 lotteries (ten lotteries in each set of trees) involved in the experiment; we have to make some restrictions. Accordingly we assume a particular form of the preference functional—to be precise that of the Rank Dependent Expected Utility model. This seems to be well accepted in the literature as the empirically most-valid generalisation of EU; moreover, it contains EU as a special case.

The Rank Dependent functional is composed of a utility function and a probability weighting function; we denote the former by $u(.)$ and the latter by $w(.)$. The Rank Dependent Expected Utility, $U(G)$, of a gamble $G=(x_1, x_2,\ldots,x_I; p_1, p_2, \ldots, p_I)$, where the prospects are indexed in order from the worst x_1 to the best x_I, is given by

$$U(G) = u(x_1) + \sum_{i=2}^{I} [u(x_i) - u(x_{i-1})]w(p_i + p_{i+1} + \ldots + p_I) \tag{1}$$

We note that Rank Dependent Expected Utility preferences reduce to Expected Utility preferences when the weighting function is given by $w(p)=p$.

To fully characterise the preferences of a subject obeying the Rank Dependent Expected Utility model, we need to know the utility function $u(.)$ and the weighting function $w(.)$. As we describe later, we assume particular functional forms for these two functions and estimate the parameters of the functions. We choose the functional forms which best fit the responses of each subject—in a way that we will describe shortly.

We now need to have a story about the noise in the subjects' responses—more technically, we need to specify the stochastic structure of the data. There are various stories that one can use and we choose to follow a Fechnerian measurement error story[20]. To be precise, we assume that, when evaluating any lottery (whether certain or risky), the subject makes a measurement error. More specifically, if $u(.)$ is the utility function of the individual, and $u^{-1}(.)$ its inverse, then we assume that the evaluated certainty equivalent of any gamble G is given by $u^{-1}(U(G))+e$, where $U(G)$ is the Rank Dependent Expected Utility of the gamble (using equation 1) and

[20] Before arriving at this particular specification, we tried several others, of varying degrees of sophistication. One simple alternative was that subjects made all evaluations with error but then 'trembled' (see Moffatt and Peters 2001) when taking decisions and when making bids; the trouble with this story (in addition to the fact that it does not seem empirically valid) is that, while the tremble story is simple to apply to decisions, it is not obvious how to interpret it with respect to bids. There are also other variations that we have tried on the basic story that we report in this paper; in particular, we explored the hypothesis that subjects made no mistakes when evaluating certainties—this performed worse than the variant reported in the paper; and also a variant that takes into account that the extreme value distribution incorporates a bias (the expected value of a variable with an extreme value distribution with parameters m and $1/s$ is not m but rather $m+\gamma s$ where γ is the Euler-Mascheroni constant 0.5772156649), by exploring the notion that the subjects corrected for this bias when making their bids; this also performed worse than the variant we have used in the paper.

where e is an error—a measurement error. To complete the story, we need to specify the distribution of e. We do this in a way that is tractable and not unreasonable—we assume that e has an Extreme Value distribution with parameters 0 and $1/s$. Thus the cumulative distribution function, $F(.)$ of e is given by:

$$F(e) = \exp(-\exp(-es)) \quad (2)$$

It follows that the probability density function $f(.)$ is given by:

$$f(e) = \exp(-\exp(-es))\exp(-es)s \quad (3)$$

To summarise: we assume that the subjects each have a well-defined (Rank-Dependent) preference function and that each subject is either naive, resolute or sophisticated. For each subject we find the best-fitting preference function and the best-fitting type (naive, resolute or sophisticated) to represent the subject's responses on the experiment. The responses are the bids (for the four trees in each of the three sets) and the decisions (in four trees on each of the three sets)—a total of 24 observations for each subject. We note that the nature of the data is different—for the bids we have a number, while for the decisions we have their choice. The former is essentially a continuous variable while the latter is a discrete variable (taking one of either two or three values). The analysis of the two kinds of data has to be different. The details are given in the next section, where we are more specific about how we have interpreted and applied our stochastic specification.

6 The stochastic and functional specifications

We have assumed that subjects evaluate lotteries with error. More specifically we have assumed that the expressed valuation, VG, of a lottery G with true certainty equivalent $u^{-1}(U(G))$ is given by $VG=u^{-1}(U(G))+e$, where e has an extreme value distribution with parameters 0 and $1/s$. When taking a decision between lottery A or lottery B, we assume that the subject evaluates both A and B with error and chooses the lottery with the highest evaluation. Given the properties of the extreme value distribution, it follows that the probability that A is chosen is $\frac{e^{VAs}}{e^{VAs}+e^{VBs}}$ and the probability that B is chosen is $\frac{e^{VBs}}{e^{VAs}+e^{VBs}}$ where VA is the true valuation of lottery A and VB the true valuation of lottery B. We note the importance of the parameter s: when s is zero then both lotteries are chosen with probability one-half, and when s is infinite then the lottery with the highest true valuation is chosen with certainty. We can therefore interpret s as indicating the precision of the subject's decision—the higher is s the more precise is the subject, and the more likely he or she is to choose the lottery with the truly highest value.

The above story can easily be extended to a choice between three lotteries A, B and C: given our error story, lottery Z (=A, B or C) is chosen with probability $\frac{e^{VZs}}{e^{VAs}+e^{VBs}+e^{VCs}}$.

This error story also enables us to determine the probability density of the bids. We assume that if a tree offers two choices—between lotteries A and B—then the bid is put equal to the *maximum* of the evaluations of A and B. Using the properties of the extreme value distribution, and assuming that the errors in the valuations are independent, it follows that the cumulative distribution function of the bid x is given by

$$\exp\{-[\exp(-(x-VA)s)+\exp(-(x-VB)s)]\} \tag{4}$$

From equation 4 we can find the probability density of the bid x based on two lotteries.

An obvious extension leads us to the cumulative distribution function of the bid when the tree offers three lotteries A, B and C and the bid is put equal to the maximum of the evaluations of these three lotteries:

$$\exp\{-[\exp(-(x-VA)s)+\exp(-(x-VB)s)+\exp(-(x-VC)s)]\} \tag{5}$$

From this we can find the probability density of the bid x based on three lotteries.

We now need to say something about how the different types are assumed to behave in the four trees. Tree 1 is simple and all types do the same thing. The decisions are based on the evaluations of M (Up) and O (Down), and the bids are made on the basis of the maximum of these two evaluations.

In Tree 2 different types do different things. The naive subject takes the decision at the decision node and this decision is based on the evaluations of K (Up) and L (Down); however the bid is based on the evaluations as viewed from the beginning of the tree—from which perspective the choice is between M and O—and hence the bid is based on the maximum of the evaluations of M and O. The resolute subject, however, bases both the decision and the bid as viewed from the beginning of the tree—that is on the basis of the evaluations of M and O. The sophisticated works backwards: he or she anticipates that the decision at the decision node will be based on the relative evaluations of K and L; if K is chosen he evaluates the tree as being worth the evaluation of M; if L is chosen the tree is worth the evaluation of O. He or she can work out the probability of choosing K and L and hence work out the expected evaluation of the tree—on which he or she bases his bid.

In Tree 3 again the different types do different things. The actual decision of a naive subject at the second decision node is based on the evaluations of L and K, but the naive agent does not anticipate this when taking the decision at the first node and in making the bid for the tree. These are determined by the evaluations of M, N and O. If either the evaluations of M or O are bigger than the evaluation of N the naive agent chooses Up at the first node; otherwise he or she chooses Down. The bid of a naive is simply determined by the maximum of the evaluations of M, N and O. In contrast a resolute subject bases both his decisions at both nodes on the evaluations of M, N and O: if the evaluation of M is the highest he or she chooses Up and then Down; if the evaluation of O is the highest he or she chooses Up and then Up; and if the evaluation of N is the highest, he or she chooses Down; the bid of a resolute subject is based on the maximum of the evaluations of M, N and O. The sophisticated subject once again anticipates his or her decision at the second node—this will be Up if L is evaluated more highly than K and Down otherwise. If L is evaluated more highly than K then his or her decision at the first node (and the bid for the tree) is based on the evaluations of O and N; if, however, K is evaluated more highly than L then his or her decision at the first node (and the bid for the tree) is based on the evaluations of M and N.

Tree 4 is simple and all do the same: the decisions and the bids are all based on the evaluations of M, N and O.

There is one slight complication that we have so far ignored: and that concerns what exactly a sophisticated person does when working backwards—when backwardly inducting. We have assumed so far that this agent eliminates the branches of the tree that he or she will not be following in the future, and then simplifies the tree by using *reduction*—that is, reducing the remaining compound lottery to a simple lottery using the usual probability rules. But there is an alternative—that this agent simplifies the tree by substituting in *certainty equivalents* of the remaining bits of the tree. Segal (1999) notes and discusses this distinction. Accordingly we should distinguish between two types of sophisticated agents—those who work backwards using reduction and those who work backwards using certainty equivalents. We refer to these as Type 1 sophisticated and Type 2 sophisticated respectively. This distinction does not affect the story we have told above about how sophisticated agents process a dynamic tree, but it does affect the specification of the appropriate likelihood functions[21].

We confine all the technical details to Web Technical Appendix 2 (with the GAUSS program for one of the combinations in Web Technical Appendix 3). Suffice it to say here that we estimate the parameter of the utility function, the parameter of the weighting function and the precision parameter s using maximum likelihood (implemented in GAUSS), subject by subject and type by type. To do this we need to specify the likelihood of the observations. This is different for the decision and for the bids: in essence the likelihood of a decision is the probability that that decision is taken given the parameters and the stochastic specification; the likelihood of a bid is the probability density at the bid given the parameters and the stochastic specification.

There is one further problem that we need to consider—the specifications of the utility function and of the weighting function. Ideally one would not specify functional forms but estimate the functions at all possible values. The problem with this is, as we have already noted, that there are too many values to make this procedure possible: given the number of observations that we have (24 for each subject), we would lose too many degrees of freedom. So we are forced to assume functional forms. The most obvious ones are the CARA and CRRA specifications for the utility function ($u(x) \propto -\exp(-rx)$ and $u(x) \propto x^r$ respectively) and the Quiggin (1982)[22] and Power specification for the weighting function ($w(p) = \frac{p^g}{(p^g+(1-p)^g)^{1/g}}$ and $w(p) = p^g$ respectively). The Quiggin specification allows for an S-shaped weighting function while the Power specification does not. In order to ensure the robustness of our results we use all four possible combinations: CRRA with Power; CRRA with Quiggin; CARA with Power; and CARA with Quiggin, and for each subject we choose the

[21] It should be noted that we do not consider reduction by substitution of certainty-equivalents for the naive and resolute types. In this we follow McClennen (1990) which implies simplification by reduction to hold for these models.

[22] To be strictly correct we should attribute this to Tversky and Kahneman (1992), who proposed this variation on the original specification proposed by Quiggin, namely: $w(p)=p^g/[p^g+(1-p)^g]$

combination which gives the highest maximised log-likelihood averaged over all four (naive, resolute, Type 1 and Type 2 sophisticated) types[23].

7 The results

To summarise: we have, for each subject, fitted the naive, resolute and (the two versions of) sophisticated types to the 24 observations for each subject for each of the four combinations of utility and weighting function. The detailed estimates are available on request. An illustration of some of the results for one of the four combinations (CARA with Quiggin) is provided in Web Technical Appendix 4. For each type and for each combination we have the following information from our estimations:

- The value of the maximised log-likelihood;
- Information as to whether the maximum likelihood converged correctly;
- The estimates and standard errors of (the transformed values[24] of) the parameters—*r*, *g* and *s*.
- (Hence) The estimates of the parameters.

We should note that for some of the specifications there were occasional problems with the convergence of the maximum likelihood process, but there were combinations (particularly CARA with Quiggin) that seemed to be particularly robust. We should note that there is no guarantee that the likelihood function is smoothly concave everywhere; accordingly we tried several starting values for the maximum likelihood software. We are reasonably convinced that we have found the overall maximum in essentially all cases.

The bottom line is the categorisation of subjects as to whether they are naive, resolute or sophisticated. As we have already noted, we begin by choosing, for each subject, the combination of utility function and weighting function that best explains the data on the subject. In some cases this was simple: irrespective of the type (naive, resolute or sophisticated) the same combination yielded the maximum of the maximised log-likelihoods—there were 25 subjects for whom the data had this property. For the rest we had to select the best combination in some manner; our criterion was to select the combination for which the average maximised log-likelihood (across all types) was maximised. This may seem somewhat arbitrary but we have carried out an independent check by seeing if the chosen type might have been different if we had chosen an alternative (but possible) combination. We found that, for all except nine of the 50 subjects, the maximised log-likelihood

[23] We should note that for 25 subjects there was one combination which fitted best on all four specifications; there were 20 subjects for whom one combination fitted best on three of the four specifications; and there were five subjects for whom one combination fitted best on two of the four specifications. The conclusion seems to be clear—for virtually all subjects, the data seems to be telling us that one combination (of utility function and weighting function) fits best *independently* of the specification.

[24] To avoid problems with the maximum likelihood algorithm we transformed our parameters to restrain their range (for example to stop the *s* parameter becoming negative). The returned estimates and variance-covariance matrix are thus those of the transformed parameters—and they need to be transformed back before they can be interpreted.

Table 5 Examples of log-likelihoods: subject for whom best combination is the same irrespective of the type

Subject number 8	aq	ap	rq	rp	Best	Highest log-likelihood
n	−1.85533	−0.50035	−13.95100	−15.35401	ap	−0.50035
r	−1.98967	−0.54727	−18.18127	−18.18191	ap	−0.54727
s1	−1.98032	−0.54233	−14.72916	−15.66046	ap	−0.54233
s2	−8.73564	−0.54233	−14.71678	−15.66046	ap	−0.54233
Mean	−3.64024	−0.53307	−15.39455	−16.21421	ap	−0.53307

Winner is naive
n naïve, *r* resolute, *s1* Type 1 sophisticated, *s2* Type 2 sophisticated, *aq* CARA and Quiggin, *ap* CARA and Power, *rq* CRRA and Quiggin, *rp* CRRA and Power

with our chosen 'best' combination and chosen 'best' type was, in fact, the highest of all the 16 maximised log-likelihoods that we computed; for seven of these nine, the maximised log-likelihood of our chosen 'best' type was higher than for any other type in all the other combinations. Just in two cases (subjects 11 and 30) would our chosen combination change our chosen type—subject 11, whom we have classified as more likely to be naive, could be classified as being sophisticated; and subject 30, whom we have also classified as being naive, could possibly be reclassified as resolute. We feel that our classification of our types is reasonably robust.

We give some examples of what we have done. Consider Tables 5, 6 and 7 in which we give examples of three subjects which show different characteristics. All the entries in the tables are maximised log-likelihoods. The rows indicate the *type* (naive, resolute or the two types of sophisticated) and the columns indicate the *combination* (of utility function and weighting function) used in the estimation.

Table 5 shows the log-likelihoods for a subject for whom the best type is the same irrespective of the combination. (As we have already noted 25 of the 50 subjects had this property.) Indeed, for this subject the CARA/Power combination performs better than the CARA/Quiggin combination and those two significantly better than either of the combinations involving the CRRA utility function. Moreover, independently of the combination, the naive type fits the data better than the others. We therefore select the CARA/Power combination as representing the subject's preferences and conclude that this subject is more likely to be naive than any of the other types.

Table 6 Examples of log-likelihoods: subject for whom best combination is the same for three of the types

Subject number 3	aq	ap	rq	rp	Best	Highest log-likelihood
n	−15.62822	−16.13782	−15.88433	−17.09272	aq	−15.62822
r	−14.03244	−13.21857	−14.03246	−13.68696	ap	−13.21857
s1	−15.69680	−16.00562	−15.85117	−16.59095	aq	−15.69680
s2	−15.75262	−16.00562	−15.94487	−16.59095	aq	−15.75262
Mean	−15.27752	−15.34191	−15.42821	−15.99040	aq	−15.27752

Winner is resolute. For abbreviations, please refer to Table 5 legend

Table 7 Examples of log-likelihoods: subject for whom best combination is the same for two of the types

Subject number 46	aq	ap	rq	rp	Best	Highest log-likelihood
n	−21.64725	−21.13685	−22.07941	−20.74374	rp	−20.74374
r	−22.56506	−21.83636	−22.96630	−20.81319	rp	−20.81319
s1	−21.28023	−20.36172	−22.01664	−21.27399	ap	−20.36172
s2	−20.82796	−20.36172	−21.37605	−21.27399	ap	−20.36172
Mean	−21.58013	−20.92416	−22.10960	−21.02623	ap	−20.92416

Winner is sophisticated (either Type). For abbreviations, please refer to Table 5 legend

Subject 3, shown in Table 6, is somewhat less clear cut. For this subject the best fitting combination depends upon the type; only by averaging across the types are we able to declare the CARA/Quiggin combination 'best' for this subject. There were 20 subjects for whom the estimates had this property. However, we do note that, whatever the combination, the log-likelihood for the resolute type is higher than for the other types. We therefore conclude that this subject is resolute.

Subject 46, shown in Table 7, is even less clear cut. Only five of the 50 subjects were similar to this one. It will be seen that for two of the types (naive and resolute) the CRRA/Power combination is best, whereas for the two sophisticated types the CARA/Power combination is best. However, both sophisticated types are best for three of the combinations and it is only for the CRRA/Power combination does the sophisticated type come out worse than the other two types. To decide on the best type for this subject, we have averaged the log-likelihoods over all types for each combination and have chosen the combination for which this average is highest—this leads us to the choice of the CARA/Power combination as best representing this subject's preferences. Given that, we note that the sophisticated likelihoods are the largest in the CARA/Power combination—and hence we declare 'sophisticated' as the winner. Note too, that of all 16 maximised log-likelihoods in the table (excluding the means) those for the sophisticated with the CARA/Power combination are the largest.

At this point we have chosen the best combination for each subject. At the same time we have also chosen the best type for each subject (that is, the type for which the maximised log-likelihood for the chosen combination is greatest). However we also want to give an indication of how much better the best-fitting type is relative to the others. To this end we proceeded as follows. For each subject we have three maximised log-likelihoods (where we take just the better of the two sophisticated specifications—that with the highest maximised log-likelihood). Let us denote these by $ll(n)$, $ll(r)$ and $ll(s)$ for the naive, resolute and sophisticated types respectively. If we adopt a Bayesian interpretation of the results, and if we start with equal priors on the three types, then the posterior probabilities of the naive, resolute and sophisticated types being the correct ones are respectively

$$P(i) = \frac{\exp(ll(i))}{\exp(ll(n)) + \exp(ll(r)) + \exp(ll(s))} i = n, r, s \tag{6}$$

Hence, for example, for Subject 1[25] on the CRRA plus Quiggin combination we have:

$$ll(n) = -11.78101, ll(r) = -14.15611 \text{ and } ll(s) = -12.49711$$

and hence, applying formula 6, we have that the posterior probabilities for the three types are

$$P(n) = 0.632, P(r) = 0.059 \text{ and } P(s) = 0.309.$$

Hence, for this subject the naive type is over twice as likely as the sophisticated type, and the resolute type is relatively very unlikely.

We have applied this analysis to each of the subjects (using the best fitting combination of utility function and weighting function for each subject individually) and present the results graphically in Fig. 3. In this triangle we represent the probability of the naive type being correct on the horizontal axis, and the probability of the sophisticated type being correct on the vertical axis. The probability of the resolute type being correct is the residual. In the triangle subjects are indicated by a number. The triangle is divided into three areas—the one to the top being where the sophisticated type is most probable, the one to the right being where the naive type is most probable and the one nearest the origin being where the resolute is most probable. We note that there are 25 subjects in the "naive most likely area", 20 subjects in the "resolute most likely area" and just five subjects in the "sophisticated most likely" area.

Of course, for subjects whose preferences are those of Expected Utility, there is no possibility of dynamic inconsistency and hence no meaning to the distinction between the different types (naive, resolute and sophisticated). In principle, therefore, we should exclude such subjects from our analysis. However, given the stochastic nature of our data we cannot be sure whether a subject is EU or not, though we do have estimates of the parameter g of the weighting function. If g is equal to 1, the weighting function reduces to $w(p)=p$ and the Rank Dependent model reduces to Expected Utility. We can carry out formal tests (based on our estimates) of the hypothesis that the g parameter is significantly different from 1. Out of the 50 subjects there are 28 subjects for whom the g parameter is significantly different from 1 at the 1% level—and hence probably not subjects with Expected Utility preferences. If we restrict our analysis to these 28 subjects we get Fig. 4. It will be seen that there are 15 in the "naive most likely" area, 12 in the "resolute most likely" and just 1 in the "sophisticated most likely". These results strengthen our comments above.

We conclude that the sophisticated type performs consistently worse than the other two, and that the naive type performs marginally better than the resolute type. We comment further on these findings in the next and concluding section.

[25] In this case the maximised log-likelihoods for the Type 1 and Type 2 sophisticated specifications were −13.32295 and −12.49711 and we accordingly prefer Type 2 to represent sophistication.

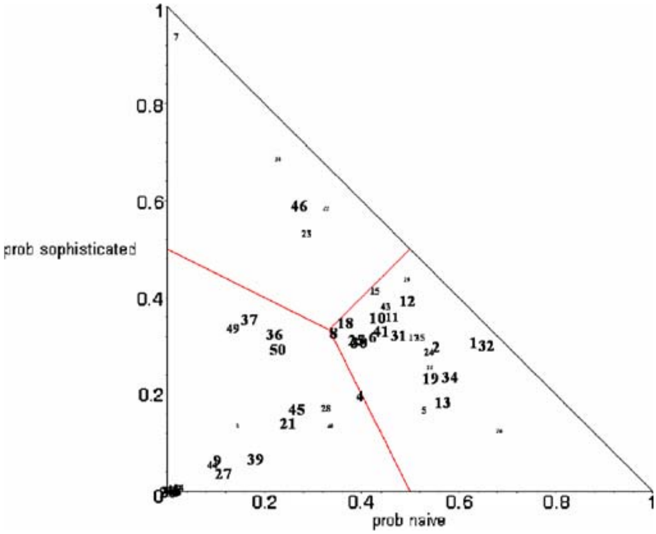
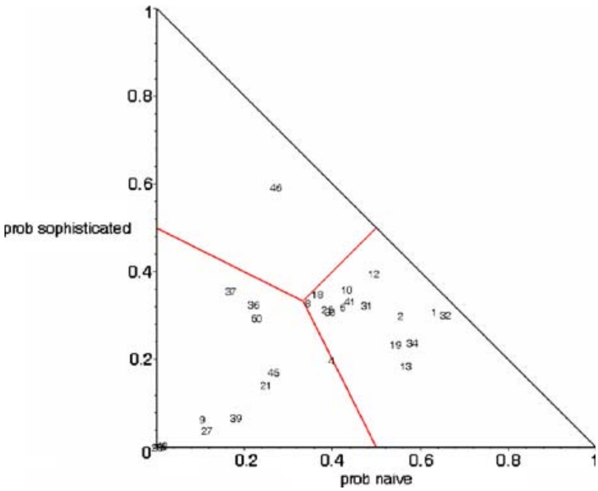


Fig. 3 The *ex post* probabilities of the various types for all subjects. Large number—*g* parameter significantly different from 1 at 1%. Medium sized number—*g* parameter significantly different from 1 at 5%. Small number—*g* parameter insignificantly different from 1. Very small number—*g* parameter insignificantly different from 1 but GAUSS unable to calculate standard errors

8 Conclusions

Dynamically inconsistent economic agents provide a serious challenge to economic theory. In order to predict their future behaviour, one needs to know how they are resolving any dynamic decision problem—in particular, for example, whether they are naive, resolute or sophisticated. In conducting experiments to try to determine whether subjects are naive, resolute or sophisticated, experimenters have a problem in designing the experiment in such a way that they can infer what subjects are doing when processing the dynamic decision problem. Essentially, the issue is concerned

Fig. 4 The *ex post* probabilities for those subjects who appear to have non-EU preferences



with whether subjects are planning their future behaviour and anticipating any possible future inconsistencies. Ideally, one would like to design an experiment where any plans that the subject is making are revealed. However, there are serious (possibly insurmountable[26]) difficulties in observing plans and hence in seeing whether they are implemented. If one simply asks the subject what he or she is planning to do, then, unless one forces the subject to implement that plan, there is no incentive for the subject to report any plans honestly; moreover, asking them such a question raises in the subject's mind the idea that they possibly *ought* to plan. Furthermore, if one insists that the subject implements the plan that he or she has stated, then one changes the nature of the problem from a dynamic problem to a pre-commitment problem.

The experiment surmounts these difficulties with a unique design in which not only behaviour is observed but also reported evaluations of different dynamic decision problems are obtained. Moreover, we have constructed the decision problems in such a way that we can distinguish, using data on both decisions and reported evaluations, the decision process actually followed.

In particular, we can distinguish between naive decision makers (those who ignore any possible future inconsistencies), resolute decision makers (those who are resolute in implementing their *a priori* plans) and sophisticated decision makers (who anticipate their future inconsistencies and who are not sufficiently resolute to overcome them). These different types will have different behaviour in our trees as well as different evaluations of the trees. We have used the data to infer the type of each subject. The picture is somewhat clouded as we do not know *ex ante* the preference functional of each individual, but we have investigated four different combinations for each subject and chosen the best. While the final picture is not totally clear, it seems to be the case that around 50% of our 50 subjects are naive, 40% are resolute and just 10% sophisticated.

The large number of resolute subjects and the small number of sophisticated subjects in our experiment surprised us, as we thought *ex ante* that it would be difficult for subjects to be resolute. However, taking into account that we obliged the subjects to spend 15 min evaluating the trees (and practising playing them out), it seems to have been the case that subjects used this time to work out an *ex ante* strategy and to realise that it was better for them to implement that rather than behaving in a sophisticated way: remember that sophisticated behaviour is not optimal as viewed by the decision maker at the beginning of the tree. It would be interesting to explore whether this finding is sensitive to the length of the time spent contemplating the trees.

The implications for economic theory are significant. If we look at models which incorporate dynamically inconsistent behaviour (such as the literature on quasi-hyperbolic discounting in the context of a life-cycle saving model[27]), it will be seen that most of these models assume sophisticated behaviour. Our results suggest that this might be descriptively implausible. If subjects are indeed naive or resolute rather than sophisticated, then the predictions of these models need to be modified appropriately.

[26] But see Busemeyer et al. (2000) for the psychologists' way round the problem.

[27] See, for example, Harris and Laibson (2001).

References

Busemeyer, J. R., Weg, E., Barkan, R., Li, X., & Ma, Z. (2000). Dynamic and consequential consistency of choices between paths of decision trees. *Journal of Experimental Psychology: General*, *129*, 530–545.

Cubitt, R. P., Starmer, C., & Sugden, R. (1998). Dynamic choice and the common ratio effect: an experimental investigation. *Economic Journal*, *108*, 1362–1380.

Cubitt, R. P., Starmer, C., & Sugden, R. (2004). Dynamic decisions under uncertainty: Some recent evidence from economics and psychology. In I. Brocas and J. D. Carrillo (Eds.) *The psychology of economic decisions, volume 2: Reasons and choices* (pp. 81–107). Oxford: Oxford University Press.

Hammond, P. J. (1976). Changing tastes and coherent dynamic choice. *Review of Economic Studies*, *43*, 159–173.

Hammond, P. J. (1977). Dynamic restrictions on metastatic choice. *Economica*, *44*, 337–350.

Hammond, P. J. (1988a). Consequentialism and the independence axiom. In B. R. Munier (Ed.), *Risk, decision and rationality* (pp. 503–516). Dordrecht: Reidel.

Hammond, P. J. (1988b). Consequentialist foundations for expected utility. *Theory and Decision*, *25*, 25–78.

Hammond, P. J. (1989). Consistent plans, consequentialism, and expected utility. *Econometrica*, *57*, 1445–1449.

Harris, C., & Laibson, D. (2001). Dynamic choices of hyperbolic consumers. *Econometrica*, *69*, 935–957.

Hey, J. D. (2005). Why we should not be silent about noise. *Experimental Economics*, *8*, 325–345.

Hey, J. D., & Lee, J. (2005a). Do subjects separate (or are they sophisticated)? *Experimental Economics*, *8*, 233–265.

Hey, J. D., & Lee, J. (2005b). Do subjects remember the past? *Applied Economics*, *37*, 9–18.

Hey, D. J., & Paradiso, M. (2006). Preferences over temporal frames in dynamic decision problems: An experimental investigation. *The Manchester School*, *74*, 123–137.

Karni, E., & Safra, Z. (1989). Dynamic consistency, revelations in auctions, and the structure of preferences. *Review of Economic Studies*, *56*, 421–433.

Machina, M. J. (1989). Dynamic consistency and non-expected utility models of choice under uncertainty. *Journal of Economic Literature*, *27*, 1622–1668.

McClennen, E. F. (1990). *Rationality and dynamic choice: Foundational explorations*. Cambridge: Cambridge University Press.

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

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

Raiffa, H. (1968). *Decision analysis*. Reading: Addison-Wesley.

Segal, U. (1999). Dynamic consistency and non-expected utility. In L. Luini (Ed.), *Uncertain Decisions. Bridging theory and experiments*. Dordrecht: Kluwer.

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

Starmer, C., & Sugden, R. (1991). Does the random-lottery incentive system elicit true preferences? *American Economic Review*, *81*, 971–979.

Strotz, R. H. (1955–6). Myopia and inconsistency in dynamic utility maximisation. *Review of Economic Studies*, *23*, 165–180.

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

Further Reading

Beggs, S., Cardell, S., & Hausman, J. (1981). Assessing the potential demand for electric cars. *Journal of Econometrics*, *16*, 1–19.