An Experimental Investigation of the Role of Errors for Explaining Violations of Expected Utility

Ulrich Schmidt[*] and Tibor Neugebauer

University of Hannover
Discussion Paper No. 279
April 2003
ISSN 0949-9962

Abstract:

One possible conclusion from recent experimental research on decision making under risk is that observed behaviour can be reasonable accommodated by expected utility plus an error term. This conclusion implies that the violation rate of expected utility should decrease if errors are excluded. The present paper presents an experiment which investigates this implication. Indeed, the results show that the exclusion of errors leads to a significant reduction of the violation rate for most of the subjects and most of the choice problems under risk. However, it turns out that for decision problems under ambiguity the exclusion of errors in contrast increases the violation rate significantly. In this sense the Ellsberg paradox can be regarded as a more serious challenge of expected utility than the Allais paradox. More general, while expected utility plus error term may be regarded as a reasonable representation for choice under risk this does not seem to be true for ambiguous choice problems.

Keywords: expected utility, choice errors, Allais paradox, Ellsberg paradox.
JEL classification: C9, D8.

---

[*] Corresponding author: Ulrich Schmidt, Lehrstuhl für Finanzmarkttheorie, Universität Hannover, Königsworther Platz 1, 30167 Hannover, Germany, Phone: (+49) (511) 762-5117, Fax: (+49) (511) 762-5240, e-mail: u.schmidt@mbox.vwl.uni-hannover.de .

1 Introduction

The common consequence effect and the common ratio effect both introduced by Allais (1953) as well as the well-known paradox of Ellsberg (1961) are the most prominent and most investigated violations of expected utility (EU). They have motivated the development of alternative theories of choice under risk and uncertainty able to accommodate the observed patterns of behaviour. Nowadays a large number of alternative theories exist (cf. Starmer (2000), Sugden (2002), and Schmidt (2002) for surveys) and naturally the question arises which theory can accommodate observed choice behaviour best.

Many studies have been devoted to this question, most prominent seem to be those of Harless and Camerer (1994) and Hey and Orme (1994). Since EU and its alternatives are deterministic theories but observed choices are stochastic both papers integrated an error term into their estimations. This fact has aroused interest in a general discussion of the role of errors for decision making under risk. The discussion can be disentangled into two issues. The first issue is the best way of modelling the stochastic component. Three different ways are discussed in the literature, the constant error model of Harless and Camerer (1994), the white noise model of Hey and Orme (1994) and Hey (1995), and stochastic utility models revived by Loomes and Sugden (1995). Experimental investigations of these models have been conducted by Carbone (1997), Ballinger and Wilcox (1997), Loomes and Sugden (1998), Carbone and Hey (2000), Buschena and Zilberman (2000), and Loomes, Moffat, and Sugden (2002). While the constant error model performs in general poorly the evidence for comparing the white noise model and stochastic utility models is mixed. The second issue concerns the performance of EU and its alternatives for given specifications of the error term. In this context, Hey (1995), building upon the results of Hey and Orme (1994), arrives at the following conclusion:

> “It may be the case that these further explorations may alter the conclusion to which I am increasingly being drawn: that one can explain experimental analyses of decision making under risk better (and simpler) as EU plus noise – rather than through some higher level functional – as long as one specifies the noise appropriately.”

This conclusion is reinforced by the results of Buschena and Zilberman (2000) which show that, under heteroscedastic error terms, the alternative theories do not offer a “statistically significant improvement in predictive power over EU”.

These conclusions are obviously in conflict with the high violation rates of EU observed in the common consequence and common ratio effects as well as the Ellsberg paradox. In other words, can EU - plus an appropriate error term - be a good explanation of choice behaviour although there exist choice problems for which most subjects behave in contrast to EU? Answering this question with yes obviously implies that the observed violations of EU are, at least partly, due to errors. The goal of the present paper is to analyse whether this is true. More precisely, we consider common consequence and common ratio effects as well as the Ellsberg-paradox and investigate whether the violation rates in the absence of errors is significantly lower than the usual violation rate in the presence of errors. It seems to be indisputable that only if this is indeed the case, EU plus error term can be an acceptable descriptive model.

The main argument of the analysis is as follows. It is assumed that subjects have a deterministic preference ordering over lotteries which does not change fundamentally during the experiment. However, due to errors, subjects may not always choose in accordance with this preference ordering. In the experiment preferences for a given choice problem are assessed three times, i.e. the same choice questions are asked in three different rounds. Suppose that an individual made the same choice in all three rounds. Then, if the probability of errors is not too high, it is very probable that this choice reflects the true preferences of the individual, i.e. no error occurred. On the other hand, if the choice varied in the single rounds, errors must have been involved in the choice process. Suppose that EU plus error term is the correct representation of preferences. Then violations of EU must be less frequent in the choice problems where no errors occurred than in the choice problems with errors. This is the hypothesis we focus on in our investigation. The experimental design is presented in the next section. Section 3 discusses our theoretical framework and explains our hypothesis more detailed. The results are presented in Section 4 and Section 5 contains some concluding remarks.

2 Experimental Design

The experiment was conducted at the Centre of Experimental Economics at the University of York with 24 participants. Each participant had to attend five separate rounds, A, B, C, D, and E, on five different days. After a subject had completed all five rounds one question of one round was selected randomly and played out for real. The average payment to the subjects was £34.17 with £80 being the highest and £0 being the lowest payment.

Table 1: The lottery pairs

	Safe lottery				Risky lottery			
No.	£0	£10	£30	£40	£0	£10	£30	£40
1	.000	.000	1.000	.000	.200	.000	.000	.800
2	.750	.000	.250	.000	.800	.000	.000	.200
3	.300	.600	.100	.000	.320	.600	.000	.080
4	.000	.600	.100	.300	.020	.600	.000	.380
5	.000	1.000	.000	.000	.700	.000	.000	.300
6	.000	.500	.500	.000	.350	.000	.500	.150
7	.500	.500	.000	.000	.850	.000	.000	.150
8	.000	.000	.700	.300	.150	.000	.000	.850
9	.800	.000	.140	.060	.830	.000	.000	.170
10	.200	.000	.740	.060	.230	.000	.600	.170
11	.000	.200	.800	.000	.000	.500	.000	.500
12	.500	.100	.400	.000	.500	.250	.000	.250
13	.000	.200	.600	.200	.200	.000	.400	.400
14	.000	.100	.300	.600	.100	.000	.200	.700
15	.200	.800	.000	.000	.800	.000	.000	.200
16	.100	.400	.500	.000	.400	.000	.500	.100
17	.000	.400	.600	.000	.400	.000	.000	.600
18	.500	.200	.300	.000	.700	.000	.000	.300
19	.000	.200	.300	.500	.200	.000	.000	.800
20	.000	.200	.700	.100	.200	.000	.400	.400
21	.000	.000	.500	.500	.100	.000	.000	.900
22	.500	.000	.500	.000	.600	.000	.000	.400
23	.250	.500	.250	.000	.300	.500	.000	.200
24	.000	.500	.000	.500	.200	.200	.000	.600
25	.500	.250	.000	.250	.600	.100	.000	.300
26	.000	.250	.500	.250	.000	.350	.000	.650
27	.000	.000	.750	.250	.000	.100	.250	.650
28	.250	.250	.500	.000	.250	.350	.000	.400

In each of the five rounds the subjects were presented with the same 30 lottery pairs, 28 risky and two ambiguous ones. All the risky lotteries were composed of four different consequences, £0, £10, £30, £40. The probabilities of the 28 risky lottery pairs over these consequences are given in Table 1. Note that in each pair in Table 1 the left lottery is safer than the right lottery, though in the experiment the left-right juxtaposition was randomised. The lotteries were presented as segmented circles on the computer screen. Figure 1 presents an example in which there is a 50% chance of getting £10, a 20% chance of getting £30, and a 30% chance of getting £40. If a subject received a particular lottery as reward he or she had to spin a wheel on the corresponding circle. The amount won was then determined by the segment of the circle in which the arrow on the wheel stopped.

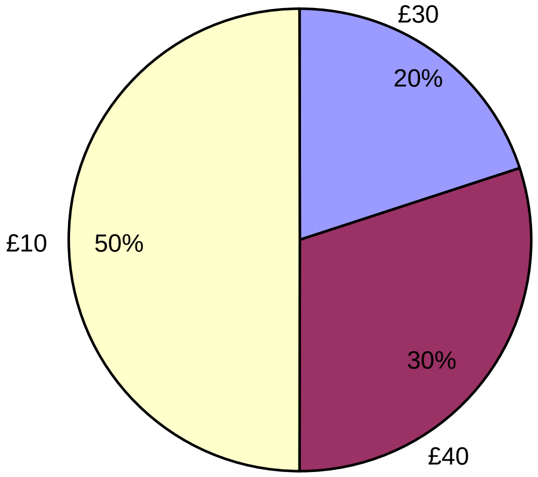

Figure 1: The Presentation of Lotteries

The lottery pairs in Table 1 contain altogether 28 common consequence or common ratio effects. An example for the common consequence effect is given by the lottery pairs 6 and 7. The left lottery in pair 7 is constructed form the left lottery in pair 6 by taking away a probability mass of 50% from the outcome £30 and assigning it to the outcome £0. By the same construction the right lottery in pair 7 is obtained from the right lottery in pair 6. Preferences are only consistent with EU if one prefers the left lottery in both pairs 6 and 7 or the right lottery in both pairs 6 and 7. However, the commonly observed pattern is that subjects choose the left lottery in pair 6 and the right one in pair 7 which violates EU.

An example for the common ratio effect is given by the lottery pairs 1 and 2. The left lottery in pair 2 is constructed form the left lottery in pair 1 by multiplying all probabilities with 0.25 and assigning the remaining probability of 75% to the outcome £0. By the same construction the right lottery in pair 2 is obtained from the right lottery in pair 1. Again, preferences are only consistent with EU if one prefers the left lottery in both pairs or the right lottery in both pairs. However, the commonly observed pattern is that subjects choose the left lottery in pair 1 and the right one in pair 2 which violates EU.

The two ambiguous lottery pairs were expressed as in the Ellsberg Paradox with £30 as the possible prize. More, precisely, for both pairs first the following text appeared on screen: “Consider an urn which contains 90 balls, 30 of them are red, the others are either blue or yellow in an unknown proportion”. Then, the two lottery pairs were described as follows.

Pair 29:

You win £30 if a red ball is drawn from the urn and nothing if a blue or yellow ball is drawn	You win £30 if a blue ball is drawn from the urn and nothing if a red or yellow ball is drawn.

Pair 30:

You win £30 if a red or a yellow ball is drawn from the urn and nothing if a blue ball is drawn.	You win £30 if a blue or a yellow ball is drawn from the urn and nothing if a red ball is drawn .

Preferences are consistent with EU if one prefers the left lottery in both pairs or the right lottery in both pairs. However, the commonly observed pattern is that subjects choose the left lottery in pair 29 and the right one in pair 30.

In last two of our five rounds, i.e. rounds D and E we elicited certainty equivalents which are irrelevant for the present paper. Consequently, our analysis will only rely on the data of rounds A, B, and C. In these rounds the single lottery pairs appeared in randomised order on screen and subjects had to indicate whether they prefer the left lottery or the right lottery or whether they are indifferent between both lotteries. After pressing the corresponding key the choice had to be confirmed by pressing the return key. If a question of rounds A, B, or C was selected for the reward of a subject, she or he could simply play out the chosen lottery. In the case of indifference, one lottery of the pair was chosen by the experimenter.

Altogether, the design allowed us to elicit for each subject three times the preferences for the 30 lottery pairs with an incentive compatible payment mechanism.

3 The Hypothesis

Our theoretical framework is based on the theory of errors developed in Hey (1995). In this theory individuals are assumed to have deterministic preferences between lotteries which can be represented by a functional V where $V(P, Q) > 0$ (< 0) indicates that lottery P is strictly preferred (strictly not preferred) to lottery Q. However, individuals sometimes make errors such that the actual choice may not correspond to the given preference relation. Formally, there is an error term ε such that in practice the value of $V(P, Q) + \varepsilon$ determines the choice between P and Q. More precisely, the individual will choose P (Q) if and only if $V(P, Q) + \varepsilon > 0$ (< 0). It is assumed that ε has a mean of zero which implies that the actual choice is contrary to the true preferences with a probability of less then 50%. In the following a lottery P will be represented by a vector $P = (x_1, p_1; x_2, p_2; \ldots; x_n, p_n)$ indicating that consequence x_i is received with probability p_i. Contrary to Hey (1995) we assume that preferences are always in accordance with EU. Hence, there exists a von Neuman-Morgenstern utility function u such that V(P, Q) has to equal the EU of P minus the EU of Q, i.e. $V(P, Q) = \Sigma_i u(x_i)(p_i - q_i)$.

Recall from the preceding section that each of the three problems, i.e. the common consequence effect, the common ratio effect, as well as the Ellsberg-paradox, can be described by four lotteries, P, Q, P*, and Q*, such that each EU maximiser prefers P over Q in the first lottery pair if and only if she or he prefers P* to Q* in the second one. In terms of our model this gives $V(P, Q) > 0$ if and only if $V(P^*, Q^*) > 0$ for all functions u. In the experiment we have six observations for each problem: the three choices from the first lottery pair in rounds A, B, and C and the three choices from the second lottery pair in rounds A, B, and C. In the following these observations will be represented by a vector where the first three entries report the choices from the first lottery pair in rounds A, B, and C respectively while the last three entries report the choices from the second lottery pair in rounds A, B, and C respectively. Hence, for instance (P, Q, P, Q*, Q*, P*) indicates that a subject chose P over Q in rounds A and C, Q over P in round B, Q* over P* in rounds A and B, and P* over Q* in round C. We say in the following that no *observed error* occurred for a given problem if the first three entries are identical and the last three entries are identical for this problem. Hence an observed error occurred if the individual made contradictory choices in the single rounds for at least one of the two lottery pairs. In contrast we say that a *real error* occurs if one given choice in any round and for any lottery pair deviates from true preferences, i.e. a subject chose Q over P also V(P, Q) > 0 holds. Note that the probability of a real error is always less than 50% since the mean of ε has been assumed to be equal to zero.

Let us ignore the possibility of indifference at the moment and suppose first that no observed errors occur, i.e. for a given lottery pair the choices in all three rounds are identical. In this case there are only four possible response patterns: (P, P, P, P*, P*, P*), (Q, Q, Q, Q*, Q*, Q*), (P, P, P, Q*, Q*, Q*), and (Q, Q, Q, P*, P*, P*), where the latter two patterns violate EU. Suppose that according to true preferences P is better than Q and, consequently, also P* better than Q* since true preferences are assumed to be consistent with EU. Moreover, assume that the probability of a real error is α when choosing between P and Q and β when choosing between P* and Q*. This implies that (P, P, P, P*, P*, P*) is observed with a probability of $(1-\alpha)^3(1-\beta)^3$, (Q, Q, Q, Q*, Q*, Q*) with a probability of $\alpha^3\beta^3$, (P, P, P, Q*, Q*, Q*) with a probability of $(1-\alpha)^3\beta^3$ and (Q, Q, Q, P*, P*, P*) with a probability of $\alpha^3(1-\beta)^3$. Hence we observe a violation of EU with a probability of $[(1-\alpha)^3\beta^3 + \alpha^3(1-\beta)^3] / [(1-\alpha)^3(1-\beta)^3 + \alpha^3\beta^3 + (1-\alpha)^3\beta^3 + \alpha^3(1-\beta)^3]$ if no observed error occurs. Analogously one can calculate the probability a violation of EU if an observed error occurs. It turns out that according to our model violations of EU have a higher probability in the cases where a observed error occurs then in the cases where no observed error occurs. Precisely this implication of the model will be tested in the experiment.

Note that in the case observed errors occur we may observe violations of EU in some rounds while in others we do not. Suppose that we have the response pattern (P, P, Q, P*, Q*, Q*). Here, we have a violation of EU in round B while the choices in rounds A and C are consistent with EU. Consequently, this pattern will be treated only as a 33.33% violation of EU in the statistical analysis of the next section. Analogously, (P, P, P, P*, Q*, Q*) has to be treated as a 66.67% violation, (P, P, P, Q*, Q*, Q*) as a 100% violation, and (P, P, P, P*, P*, P*) as a 0% violation.

Before the results are presented it should be explained how indifferences have been treated in the analysis. If for a given lottery pair and a given subject there was one indifference in the three rounds this indifference has been treated as missing observation. If there were two or three indifferences, the complete choice problem for this subject has been removed from the analysis. This procedure allowed us to exclude violations of EU which simply result from imprecise preference.

4 Results

Let us first give an overview of our results and consider the complete sample of all subjects and all lottery pairs. In this sample the overall violation rate of EU is given by 25.51%. In the cases without observed error this violation rate decreases to16.95, whereas it increases to 43.21% in the cases with observed error. This shows that many observed violations of EU are due to errors since excluding observed errors reduces the violation rate substantially. Note that the violation rate without observed error is less than both, the overall violation rate and the violation rate with observed error at a significance level of 1%.

As we will argue below it may be interesting to consider only the risky lottery pairs and exclude the ambiguous ones from the analysis. Then we observe the following violation rates: 23.44% for all cases, 13.89% for cases without observed error, and 42.99% for cases with observed error. Again, the violation rate without observed error is less than both, the overall violation rate and the violation rate with observed error at a significance level of 1%. We can conclude that for decision making under risk the violation rate of EU can be reduced to such a low number as 13.89% by excluding observed errors. It seems to be questionable whether this number justifies the effort which has been exerted in order to develop alternative models to EU.

Let us now turn to the between-subject analysis. Table 2 gives for each choice problem the average violation rate of EU for the cases in which an observed error occurred (second column) and for the cases in which no observed error was made (third column). The last column of table 2 reports the difference of these violation rates. A “+” means a positive difference, a “-“ a negative difference. The number of characters in this last column characterises the significance level of the difference: one character indicates an insignificant difference, two characters a significance level of 10%, three characters of 5%, and four characters of 1%.

Table 2: Violation rates for the single lottery pairs

Choice problem	Error	No Error	Difference
1/2	28.57	15.38	+
1/3	53.33	56.25	-
1/4	47.92	26.67	++++
2/3	26.67	23.08	+
2/4	43.75	30.00	+++
3/4	63.89	26.67	++++
5/6	33.33	5.55	+++
5/7	30.56	0.00	++++
6/7	39.58	7.14	++++
8/9	50.00	55.56	-
8/10	48.61	60.00	---
9/10	50.00	-	
11/12	53.33	0.00	++++
13/14	30.56	33.33	-
15/16	45.83	0.00	++++
17/18	22.22	5.55	+++
17/19	27.78	0.00	++++
17/20	29.17	0.00	++++
18/19	46.67	5.26	+++
18/20	47.62	6.25	++++
19/20	33.33	0.00	+++
21/22	37.50	7.69	++++
21/23	41.67	15.38	++++
22/23	28.57	21.43	+
24/25	38.89	6.25	++++
26/27	48.81	0.00	++++
26/28	54.17	14.29	++++
27/28	58.97	0.00	++++
29/30	58.33	100.00	----

Choosing a significance level of 5%, table 2 shows that the exclusion of the cases with observed error leads to a significant reduction of the violation rate for 20 of the 28 choice problems without missing observations, whereas it leads to a significant increase for only two choice problems. One of these two choice problems deserves further attention, namely the Ellsberg paradox consisting of the ambiguous lottery pairs 29/30. Here the violation rate in the cases without observed error amounts to noteworthy 100% while it is significantly lower for the cases with observed error. Therefore, we can conclude that EU plus error term does not seem to be suitable model for accommodating the Ellsberg paradox. Altogether, the between-subject analysis confirms the inferences from the overall data for the common consequence and the common ratio effect common ratio effect but not for the Ellsberg paradox.

Table 3: Violation rates for the single subjects

Subject No.	Error	No Error	Difference
1	.6667	.3500	+++
2	-	.1111	
3	.4896	.0833	++++
4	.4167	.2632	++++
5	.4405	.2889	+++
6	.5000	.2424	++++
7	.2222	.3846	----
8	.4242	.0952	++++
9	.3733	.0000	++++
10	.3810	.2857	+++
11	-	.1333	
12	.6000	.3333	++++
13	.2222	.0400	+
14	.5370	.0000	++++
15	.4762	.0166	++++
16	.4167	.0909	++
17	.5714	.1250	++++
18	.4583	.3200	+++
19	.3333	.3529	-
20	-	.0000	
21	.3333	.0769	++++
22	.3333	.0000	++++
23	.4286	.3636	++
24	.4500	.1111	++++

Let us finally turn to the within-subject analysis. Table 3 gives for each subject the average violation rate of EU for the cases in which an observed error occurred (second column) and for the cases in which no observed error was made (third column). The last column gives as in table 2 the sign of the difference of theses violation rates and the significance level (insignificant: one character, significance level of 10%: two characters, 5%: three characters, and 1%: four characters). Table 3 shows that the exclusion of the cases with observed error leads to a reduction of the violation rate for 90.48% (19 from 21) of the subjects without missing observations. At a significance level of 5%, this reduction is significant for 76.19% of the subjects (i.e. 16) whereas there is a significant increase of the violation rate for only one subject. Therefore, also the within-subject analysis supports the main hypothesis of the present paper.

5 Conclusions

Altogether our results show that the main implication of representing preferences by EU plus error term is confirmed by our data. However, this is only true for decision making under risk while the exclusion of errors for ambiguous choice problems increases the violation rate of EU. Insofar the Ellsberg paradox can be regarded as a more serious challenge to EU than the Allais paradox. For future research the question arises in which way the stochastic component of choice behaviour should be modelled for decision making under uncertainty and ambiguity.

Acknowledgements

The experiment was conducted while the author was visiting EXEC at the University of York. Financial support for this visit by the Deutsche Forschungsgemeinschaft is gratefully acknowledged. Further financial support was obtained by the ENDEAR project of the European Union under contract CT98-0238. First of all I would like to thank John D. Hey for his hospitality and his guidance and support in designing and running the experiment. I am indebted to Norman Spivey for writing the software for this experiment. Also thanks to Stefan Traub for help with the data processing. For critical and helpful comments on the instructions and the paper I wish to thank Marie Edith Bissey, Carmela DiMauro, and Anna Maffioletti.

References

Allais, M. (1953), Le Comportement de l'Homme Rationnel devant le Risque, Critique des Postulates et Axiomes de l'Ecole Américaine, Econometrica, Vol. 21, S. 503-546.

Ballinger, T.P., and N. Wilcox (1997), Decisions, Error and Heterogenity, Economic Journal 107, 1090-1105.

Buschena, D., and D. Zilberman (2000), Generalized Expected Utility, Heteroscedastic Error, and Path Dependence in Risky Choice, Journal of Risk and Uncertainty 20, 67-88

Carbone, E. (1997), Investigation of Stochastic Preference Theory Using Experimental Data, Economics Letters 57, 305-312.

Carbone, E. And J.D. Hey (2000), Which Error Story is Best?, Journal of Risk and Uncertainty 20, 161-176.

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

Harless, D.W., and C.F. Camerer (1994), The Predictive Utility of Generalized Expected Utility Theories, Econometrica 62, 1251-1289.

Hey, J.D. (1995), Experimental Investigations of Errors in Decision Making under Risk, European Economic Review 39, 633-640.

Hey, J.D., and Orme, C. (1994), Investigating Generalizations of Expected Utility Theory Using Experimental Data, Econometrica 62, 1291-1326.

Loomes, G., P.G. Moffatt, and R. Sugden (2002), A Microeconometric Test of Alternative Stochastic Theories of Risky Choice, Journal of Risk and Uncertainty 24**,** 103-130.

Loomes, G, and R. Sugden (1995), Incorporating a Stochastic Element into Decision Theories, European Economic Review 39, 641-648.

Loomes G., and Sugden, R. (1998), Testing Alternative Stochastic Specifications for Risky Choice, Economica 65, 581-598.

Schmidt, U. (2001), Alternatives to Expected Utility: Some Formal Theories, in: P.J. Hammond, S. Barberá, and C. Seidl (eds.), Handbook of Utility Theory Vol. II, Kluwer, Boston, forthcoming.

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.

Sugden, R. (2001), Alternatives to Expected Utility: Foundations, in: P.J. Hammond, S. Barberá, and C. Seidl (eds.), Handbook of Utility Theory Vol. II, Kluwer, Boston, forthcoming.