Empirical evidence of inconsistency in Standard Gamble choices under direct and indirect elicitation methods

SUSAN CHILTON* and ANNE SPENCER** ‡

1. INTRODUCTION

It is becoming increasingly common for government and other decision-making bodies to take account of the general public's opinions when formulating health, safety and environmental policies. Survey methods are often used to assist this process. One of the more popular methods for collecting data on the public's preferences towards different policies is the Standard Gamble (SG) method which is based upon the VON NEUMANN and MORGENSTERN (1994) axioms of Expected Utility Theory (EUT). To allow respondents to express their preferences for minor health states or injuries, the SG question is often decomposed into two or more intermediate questions (DRUMMOND, O'BRIEN, STODDART and TORRANCE, 1997). EUT predicts equivalence for directly and indirectly elicited preferences under the same method. This equivalence is termed procedural invariance and is defined more formally below.

... that the relation of preference should not depend on the description of the options or on the method of elicitation. (TVERSKY, SATTATH and SLOVIC, 1988).

If procedural invariance fails and both methods generate quite different results, then it is arguably difficult to recommend to policy makers whether to use results from the direct or indirect method.

Procedural invariance is likely to fail if respondents value changes in outcome relative to a reference point rather than valuing absolute outcomes (CAMERER, 1995, p. 625). Procedural invariance is also likely to fail if respondents assign non-linear weights to probabilities (MACHINA, 1987). Systematic failures of procedural invariance, therefore, bring into question the appropriateness of EUT as a model of decision making in applied studies. It is the aim of this paper to test whether procedural invariance holds.

The results reported here arise from two studies that utilise SG questions. At the outset we note the seemingly atheoretic nature of the results from the two studies. Due to the methodological objectives of these studies, it is impossible to say whether the SG per se fails procedural invariance. In particular, procedural invariance might fail be-

* Department of Economics, University of Newcastle-upon-Tyne, UK.
** Department of Economics, Queen Mary, University of London, UK. The authors would like to acknowledge the work of Mike Jones-Lee, Graham Loomes, Judith Covey, Nick Pidgeon and Lorraine Hopkins in the Health and Safety study reported in this paper. Any errors in this further paper are the responsibiity of the current authors alone.
‡ Authors in alphabetical order. No seniority implied.

cause: 1) the SG method produces results that are inconsistent with EUT, 2) the indirect method produces results that are inconsistent with EUT, or 3) both produce results that are inconsistent with EUT. Nevertheless, we feel that the results call into question the robustness of the method and outline a future research agenda to investigate the fundamental properties of the SG method.

The two studies (CARTHY, CHILTON, COVEY, HOPKINS, JONES-LEE, LOOMES, PIDGEON and SPENCER, 1999; SPENCER, 1998) utilise a modified SG question.[1] The modified SG question offers the choice between two risky outcomes whereas conventional SG question offers the choice between a certain and risky outcome.[2] The first study is a value of statistical life study, VOSL (BEATTIE, CHILTON, COOKSON, COVEY, HOPKINS, JONES-LEE, LOOMES, PIDGEON, ROBINSON and SPENCER, 1998a). Here the VOSL is estimated through a combination of Contingent Valuation (CV), described in detail later, and modified SG questions. Both the CV and modified SG questions are based on EUT and should comply with procedural invariance. The second study is a SG laboratory experiment designed to measure preferences over monetary lotteries and again should comply with procedural invariance.

We test for procedural invariance of the SG question using both direct and indirect methods. For example, in the VOSL study reported below, the direct method requires respondents to set one indifference probability between two risky options, based on the continuity axiom of EUT.[3] Meanwhile, the indirect method requires respondents to set two indifference probabilities which are then linked together to derive a VOSL that should be statistically equivalent to the VOSL derived under the direct method.

The two studies produce results which are not in accord with EUT and, further, the discrepancies are not in the same direction. The direction of the discrepancies, therefore, appears to be crucially dependent upon framing of the questions. Failures of procedural invariance have already been reported for SG questions that are decomposed into two or more intermediate questions (LLEWELLYN-THOMAS et al., 1982; RUTTEN-VAN MÖLKEN et al., 1995; SPENCER, 1998). Moreover, these failures have been attributed to theories other than EUT, such as Prospect Theory. Our paper adds to this debate by showing that the direction of discrepancy observed in these studies is also implied by Prospect Theory and to our knowledge this is the first time that Prospect Theory has been used to explain such discrepancies.

The remainder of the paper is set out as follows. In Sections 2 and 3, we present the

1. CARTHY et al. (1999) coined the phrase *modified SG* but these types of questions are also referred to as paired gambles by FARQUHAR (1984) and lottery equivalent questions by MCCORD and DE NEUFVILLE (1986).
2. Conventional SG questions offer the choice between a certain or risky outcome. For health care decisions, these SG questions involve the choice between remaining in a given health state or a more risky treatment with the potential for a better health state. For financial decisions, these questions involve the choice between a certain monetary outcome or a more risky monetary lottery with the potential for a larger financial return. The modified SG method used here asks respondents to compare two risky treatments.
3. A SG question has a similar format, but one of the treatments involves no risk (i. e. $\theta = 0$).

studies and note that the results imply that procedural invariance does not hold. In Section 4 we consider alternative models that appear to explain the data. Section 5 concludes by reviewing the main policy and research implications.

2. STUDY 1: SAFETY VALUATION

2.1. Methods

The main aim of the study was to estimate the VOSL for road risks to update the figure used by the UK Department of the Environment, Transport and the Regions in cost-benefit analysis of road safety schemes.[4]

The VOSL is the value that individuals place on the avoidance of one statistical death (JONES-LEE, HAMMERTON and PHILIPS, 1985). Jones-Lee proves the standard result that the VOSL is given by the population mean marginal rate of substitution (MRS) of wealth for the probability of death (JONES-LEE, 1989). The CV method (MITCHELL and CARSON, 1989) has become an increasingly accepted and well-used questionnaire method to elicit the VOSL in studies utilising a "stated preference" approach (as opposed to relying on "revealed preference"). A representative sample of the population are asked more or less directly about the amount of money they would be willing to pay or to accept as compensation for pre-specified variations in their own safety. Respondent values are then simply summed and the mean taken, to derive the aggregate VOSL.

However, BEATTIE et al. (1998a) report serious problems in using the direct CV method to estimate the VOSL. Based on this, CARTHY et al. (1999) developed an alternative, arguably cognitively less complex, methodology to help respondents arrive at their trade-off between money and risk of death. In this "CV/SG Chained" method, respondents were asked CV questions to elicit their willingness-to-pay (WTP) for a certain cure for a non-fatal injury and their willingness-to-accept (WTA) compensation for the certainty of "putting up" with the same injury. These were then combined with a modified SG question to estimate the VOSL both directly and indirectly.

It is not the intention here to report the theoretical basis for this new methodology nor the derivation of the various equations used below. These can be accessed in the original paper. The emphasis instead is on the results of the study and the implications of using the modified SG question directly and indirectly. From a theoretical perspective, if respondents are a) acting as EUT maximisers and b) truthful and accurate, then the direct VOSL will be an appropriate measure of utility. In this study, the VOSL is defined as the respondent's implicit MRS of wealth for risk of death as a result of a road accident (m_D). The respondent's MRS of wealth for risk of the non-fatal injury (m_I, where I = injury W or injury X) is obtained from the WTP and WTA questions. CARTHY et al.

4. Project sponsors: UK Health and Safety Executive; the Department of the Environment, Transport and the Regions; the Home Office; and the Treasury.

(1999) show that m_D and m_I can be then be combined with the modified SG question to estimate the VOSL directly and indirectly. In particular, it is shown that $m_D = Km_i$, where K is the implied relativity of MRS for death against the MRS for the non-fatal injury from the modified SG question. It is K that is of interest in this paper. This can be elicited either directly, using the first of two modified SG questions outlined below (equation 1) or indirectly, linking together the responses from two of the modified SG questions (equations 1 and 2 to derive equation 3).

The risky choices posed in the questionnaire are described below. Figure 1 is for illustrative purposes only and summarises the survey instrument that was used.[5] Respondents were faced with two separate modified SG questions.[6] In the first question (question 1), respondents were asked to consider Treatment A, that if successful led to a health state associated with a non-fatal Injury X: *2 weeks hospitalisation; full recovery after 18 months* and if failed led to *death*. Treatment A is depicted in the left-hand-side of Figure 1.[7, 8] This was compared against Treatment B that if successful led to full recovery in 3–4 days (denoted by normal health in right-hand-side of Figure 1) and if failed led to *death*.

Figure 1: Question 1, the modified SG question for Injury X

Treatment A	*Treatment B*
Success: $(1-\theta_x = 1-10^{-3})$ - hospitalisation, prognosis Injury X	Success: $(1-\Pi_x)$ - normal health in 3-4 days
Failure: $(\theta_x = 10^{-3})$ - immediate unconsciouness and death	Failure: (Π_x) - immediate unconsciouness and death

Respondents were asked to set the probability of failure, or their "indifference probability" (Π_X, where $\Pi_X > 10^{-3}$), for Treatment B so that the treatments appeared equally attractive. The failure rate for Treatment A, θ_X, was fixed at 10^{-3}. The respondent's *direct VOSL* can then be calculated by inserting θ_X and their indifference probability into equation (1), where m_X is the respondent's MRS of wealth for the risk of injury X.[9]

5. Details of the questionnaire can be obtained from the authors.
6. If the probability of failure was the same for Treatments A and B, respondents will strictly prefer Treatment B since it leads to normal health rather than injury X when it succeeds. As the probability of failure is increased in Treatment B, respondents will move from having a strict preference for Treatment B to having a strict preference for Treatment A. The continuity axiom of expected utility theory states that there will exist a probability of failure at which point the respondents are indifferent between Treatments A or B.
7. Respondents were also asked to suppose that, if untreated, the injury would result in death.
8. For full details of the administration procedures see CARTHY et al. (1999).
9. At the point of indifference Treatment A and B are equivalent. Treatment A can be expressed algebraically as $(1-\theta_X)m_X + \theta_X m_D$ and Treatment B can be expressed as $(1-\Pi_X)m_N + \Pi_X m_D$. The marginal rate of wealth for risk of normal health is zero, since this represents no loss of health status, so m_N is zero. Setting the two expressions for Treatment A and B equal and rearranging gives equation (1).

$$m_D = \left[\frac{1-\theta_X}{\Pi_X - \theta_X}\right] m_X \tag{1}$$

In the second modified SG question (question 2), the failure was changed from death to injury X. Respondents were asked to consider Treatment C that if successful led to a health state[10] associated with a non-fatal injury W (where W is less severe than X, i. e. W < X) : *2–3 days hospitalisation; full recovery after 3–4 months* and if failed led to *hospitalisation, prognosis Injury* X (rather than death). This was compared against Treatment D as illustrated in Figure 2.

Figure 2: Question 2, the modified SG question for Injury W

Treatment C	*Treatment D*
Success: $(1-\theta_w = 1-10^{-2})$ - hospitalisation, prognosis Injury W Failure: $(\theta_w = 10^{-2})$ - hospitalisation, prognosis Injury X	Success: $(1-\Pi_w)$ - normal health in 3-4 days Failure: (Π_w) - hospitalisation, prognosis Injury X

In this question, the probability of failure was set at $\theta_W = 10^{-2}$ for Treatment C and respondents were asked to set their indifference probability, Π_W, for Treatment D. This gives the ratio $\frac{m_X}{m_W}$, where m_W is the respondent's MRS of wealth for risk of injury W.

$$\frac{m_X}{m_W} = \frac{1-\theta_W}{\Pi_W - \theta_W} \tag{2}$$

Question 1 gives the relationship between m_D and m_X whilst question 2 gives the relationship between m_X and m_W. These two questions can therefore be linked through injury X to estimate the relationship between m_D and m_W. The indirect VOSL is calculated by linking the responses for $\frac{m_D}{m_X}$ and $\frac{m_X}{m_W}$ from equations (1) and (2). This is combined with the estimate of m_W obtained from the WTP and WTA responses to derive the indirect VOSL as shown in equation (3).

$$m_D = \left[\frac{1-\theta_X}{\Pi_X - \theta_X}\right]\left[\frac{1-\theta_W}{\Pi_W - \theta_W}\right] m_W \tag{3}$$

EUT predicts that the utilities assigned to injuries/death or health states through the direct and indirect methods should be equivalent. For instance, the independence axiom allows us to think of a unique utility for death, or other injuries, that is unaffected by the comparisons which led to this value.

10. Respondents were also asked to suppose that, if untreated, the injury would result in the prognosis associated with Injury X.

The empirical study was carried out over 4 weeks (October–November, 1997) and involved a quota sample of 167 respondents specified on the basis of gender, age and social class to reflect national breakdowns for Great Britain. The sample was also geographically spread, with 45 respondents from Newcastle, 43 respondents from York, 54 respondents from Brighton and 25 respondents from Bangor. The data was collected by members of the research team using face-to-face interviews.

CARTHY et al. (1999) report direct and indirect VOSLs for a range of cardinal utility functions. However, for the purposes of brevity, we concentrate here only on the set of results relating to the negative exponential specification.[11,12] In this we transform the CV results using logarithmic transformations to normalise the skewed data and to allow meaningful comparisons.

2.2. Results

The null hypothesis (based on the predictions of EUT) to be tested is H_0: direct VOSL = indirect VOSL. Of the 167 respondents, sixteen respondents indicated that they would not take any additional risk of death under Treatment B in question 1. For these respondents, $\Pi_X = \theta_X$, and thus $\frac{m_D}{m_X} = \infty$ (from equation 1). Two of these respondents also stated they were unwilling to take any further risk for Treatment D i. e., $\Pi_W = \theta_W$. One further respondent stated that they were willing to take an additional risk of death in question 1 but not for the prognosis associated with Injury X in question 2. These seventeen respondents are omitted from the computation of the sample mean of m_D in Table 1. Also omitted are two extreme upper tail outliers, one with an estimate that exceeds £ 235 × 10^6 and another that exceeds £ 15 × 10^6, due to some doubts about their reliability and the extreme impact that they have on estimated VOSLs.[13] This is termed "Level 1" Trimming.

At this stage we note that a pervasive feature of CV studies, "scope insensitivity" (BEATTIE et al., 1998a; KAHNEMAN and KNETSCH, 1992; DIAMOND, HAUSMAN, LEONARD and DENNING, 1993; BOYLE, DESVOUSGES, JOHNSON, DUNFORD and HUDSON, 1994) can lead to a difference between direct and indirect VOSLs, over and above procedural variance. Scope sensitivity refers to the way economically consistent measures of WTP (WTA) might be expected to change with changes in the size of the commodity being valued (SMITH and OSBOURNE, 1996). Economic theory predicts that a positive re-

11. Results from the other specifications follow the same general pattern.
12. All results are dependent on the feasible assumption that, in the case of lesser severity of injury (e.g. those involving no permanent disability), the marginal utility of wealth will not be markedly affected by the injury – in other words, the slope of the respondent's utility function is unaffected by the change from normal health state to an injury and ($\forall$W) I′(W) = U′(W) where U(W) is the respondent's cardinal utility of wealth function conditional on normal health and I(W) is the corresponding utility of wealth function conditional on sustaining a non-fatal injury (see CARTHY et al., 1999).
13. This also allows a closer comparability with any figures reported in CARTHY et al. (1999).

lationship should exist between WTP and an improvement in quality or increase in quantity of a good, assuming non-satiation within the range of investigation. Scope insensitivity leads to WTP (WTA) responses that remain unchanged across differing quantity or quality of the good (i. e. the values for different goods are not statistically significant).

Due to the potential impacts of this feature there are strong grounds for arguing that, for a "fairer" and less biased comparison of the different methods, scope insensitive respondents should be removed prior to any calculations. To simplify the notion let the WTPX and WTAX denote the WTP and WTA for injury X, likewise for injury W. Thus, "Level 2" Trimming excludes Case 1 respondents (WTPX = WTPW and WTAX = WTAW and hence $\frac{m_D}{m_X} = \frac{m_X}{m_W}$); while "Level 3" Trimming further excludes Case 2 (WTPX > WTPW and WTAX = WTAW) and Case 3 (WTPX = WTPW and WTAX > WTAW) respondents. While admittedly not a large number of respondents in this particular study, this may not hold for all studies and the argument for excluding such respondents would become even stronger.

At a first glance, comparing the means, it is clear that the indirect VOSLs are approximately 12 times larger.[14] Table 2 shows the effect on the median of removing scope insensitive respondents from the calculation of the direct VOSL. The direct VOSL increases as we remove the cases where there is scope insensitivity in both WTP and WTA responses (Case 1) i. e. from Level 1 Trimming to Level 2. In this particular case we can see that the net effect of Level 3 Trimming (removal of Case 2 and Case 3 responses) is to increase the direct VOSL and to narrow the gap between it and the indirect VOSL.

14. It could also be argued, of course, that the response mechanism may have had an adverse effect. In this study, respondents identified their indifference probability between the two treatments by sorting cards depicting difference levels of risk into separate piles: definitely reject treatment, definitely accept treatment and hard to choose. Bias may have been introduced, for example, if respondents' were insensitive to the changes of risk on the cards and used a simple rule of thumb to choose the card e.g. "pick the fourth highest card, irrespective of the value represented on its face". In question 2 (W compared X and normal health), choosing the fourth highest card would lead them to taking a higher risk than would choosing the fourth highest card in question 1 (X compared normal health and death). However, this would tend to reduce the indirect value and decrease the difference between the direct and indirect VOSL, and so would work to undermine the effect of linking rather than contribute towards it.

Table 1: Mean Direct and Indirect VOSLs

	Direct VOSL		Indirect VOSL	
Trimming	mean	standard error	mean	standard error
Level 1[1]	£0.93 $\times 10^6$	£0.15 $\times 10^6$	£11.73 $\times 10^6$	£7.73 $\times 10^6$
Level 2[2]	£0.95 $\times 10^6$	£0.16 $\times 10^6$	£11.54 $\times 10^6$	£7.99 $\times 10^6$
Level 3[3]	£0.99 $\times 10^6$	£0.17 $\times 10^6$	£12.13 $\times 10^6$	£8.78 $\times 10^6$

1. n = 148. Excludes 16 respondents where $\Pi_X = \theta_X$ (including two where $\Pi_W = \theta_W$), 1 further respondent where $\Pi_W = \theta_W$, 1 case for which *direct* $m_D >$ £235 $\times 10^6$ and 1 case for which *direct* $m_D >$ £15 $\times 10^6$.
2. n = 143. Further excludes those respondents for whom $\frac{m_D}{m_X} = \frac{m_X}{m_W}$.
3. n = 130. Further excludes those respondents for whom WTPX = WTPW, and those respondents for whom WTAX = WTAW.

Nevertheless, in all cases the indirect value remains greater than the direct, suggesting that linking questions continues to assert an upwardly inflating effect on the indirect VOSL.

Table 2: Median direct and indirect VOSL

Trimming	Direct VOSL median	Indirect VOSL median
Level 1	£0.24 $\times 10^6$	£0.47 $\times 10^6$
Level 2	£0.24 $\times 10^6$	£0.41 $\times 10^6$
Level 3	£0.26 $\times 10^6$	£0.40 $\times 10^6$

In table 3 a similar pattern emerges. This shows that there are a larger number of respondents whose indirect VOSL is greater than the direct VOSL and that the differences between the direct and indirect VOSL are statistically significant.

Table 3: Comparison of Individual Direct and Indirect VOSL's

Trimming Level	1	2	3
VOSL	no. of cases	no. of cases	no. of cases
Indirect VOSL > Direct VOSL	96	91	81
Indirect VOSL < Direct VOSL	52	52	49
Indirect VOSL = Direct VOSL	0	0	0
N	148	143	130
Wilcoxon Z	−5.66	−5.33	−4.59
2-tailed P	.000	.000	.000

If we assume that the degree of random error introduced by the linking process is small, then in totality, the results suggest that procedural invariance does not hold and, further, that respondents' preferences are not well explained by EUT. In summary, we reject H_0 in favour of H_1 direct VOSL $\neq$ indirect VOSL and we observe that the direct VOSL < indirect VOSL.

3. STUDY 2: MONETARY LOTTERIES

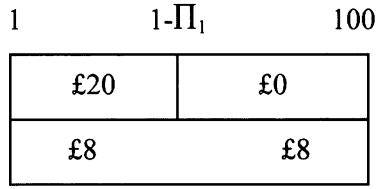
3.1. Methods

The monetary lotteries experiment was carried out at the University of York using a convenience sample of students (n=40, SPENCER, 1998). The SG questions were designed to measure respondents' risk attitude towards monetary lotteries. Here we focus on the extent to which the direct and indirect methods generate consistent estimations of utility.

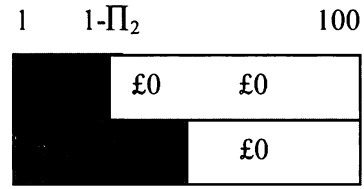
Nine of the twelve questions asked in this study are relevant to this paper.[15] All these questions can be used to estimate the utility of £ 8 through direct or indirect methods, and so the estimation of this utility is the focus of this study. In each of these questions, respondents considered two choices. For example, in question 1 (Figure 3) respondents were asked to consider the choice between a $(1 - \Pi_1)$ probability of £ 20 and the certainty of £ 8. Respondents were then asked to set the probability of £ 20 in the first lottery so that the two choices were equally attractive and they were indifferent between the two of them.[16,17]

15. The nine questions considered here were part of a larger study which is not reported here.
16. The probabilities were set at the following initial values: $(1 - \Pi_2) = (1 - \Pi_3) = (1 - \Pi_5) = (1 - \Pi_7) = 0.25$ and $(1 - \Pi_1) = (1 - \Pi_4) = (1 - \Pi_6) = (1 - \Pi_8) = (1 - \Pi_9) = 0.50$.
17. The majority of the questions fixed the probability of failure in the second lottery, θ_i, and asked respondents to set the probability of failure in the first lottery Π_i so that the two choices were equally attractive. However, in questions 6 and 9 the probabilities in the two lotteries considered were linked. For example in question 6 the probability of £ 30 was $(1 - \Pi_6)$ in the first lottery and the probability of £ 8 was Π_6 in the second. In these questions then respondents increased the probability of £ 30 in the first lottery at the cost of reducing the probability of £ 8 in the second. This should not be confused with linking responses in the indirect method which refers to the way in which the responses are analysed rather than any direct linkages between the question's probability.

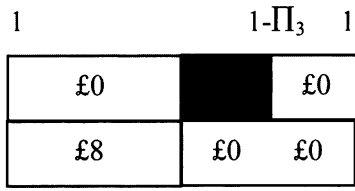
Figure 3: The monetary lottery questions



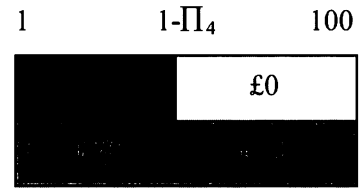
Question 1



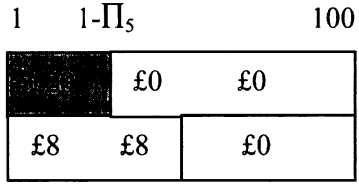
Question 2



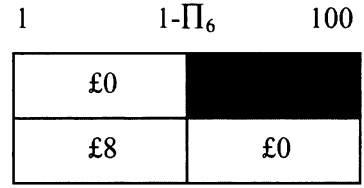
Question 3



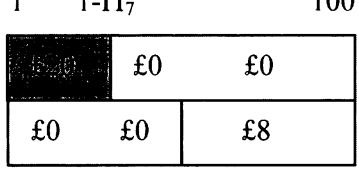
Question 4



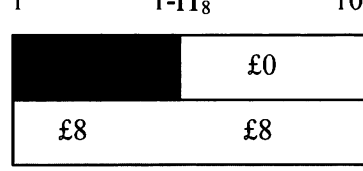
Question 5



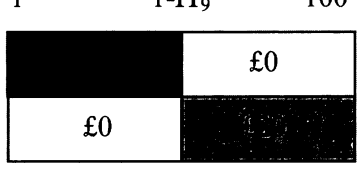
Question 6



Question 7



Question 8



Question 9

Let $U(x)$ denote the utility function of receiving x, where x can take the values £ 30, £ 20, £ 8 and £ 0. The objective is then to estimate the utility of £ 8, U(£ 8), directly or indirectly on a scale between U(£ 30) and U(£ 0). For simplicity, we set U(£ 30) = 1 and U(£ 0) = 0 and calculate the U(£ 8) relative to these two points.

The direct questions compared £ 8 directly against £ 30 and £ 0. For example, in question 3 the probability of £ 30 in the first lottery was $(1 - \Pi_3)$ and the probability of £ 8 in the second lottery was 0.5. Respondents were asked to set $(1 - \Pi_3)$ so that they were indifferent between the two choices. When U(£ 30) = 1 and U(£ 0) = 0, and EUT holds, the direct U(£ 8) = 2 $(1 - \Pi_3)$.[18]

In the indirect method, two questions were used. One question compared £ 8 against £ 20 and £ 0. For example, in question 7, the probability of £ 20 was $(1 - \Pi_7)$ and the probability of £ 8 in the second lottery was 0.5. Respondents set $(1 - \Pi_7)$ so that they were indifferent between the two choices resulting in U(£ 8) = $2(1 - \Pi_7)$ U(£ 20). The other question compared £ 20 against £ 30 and £ 0. For example, in question 2, the probability of £ 30 was $(1 - \Pi_2)$ and the probability of £ 20 was 0.5. Respondents set $(1 - \Pi_2)$ so that they were indifferent between the two choices giving U(£ 20) = $2(1 - \Pi_2)$. Question 7 gives the relationship between U(£ 8) and U(£ 20) and question 2 gives the relationship between U(£ 20) and U(£ 30). These two questions can then be linked together through U(£ 20) to derive the indirect U(£ 8) against U(£ 30).[19] For example, linking questions 7 and 2 through £ 20 lead to the following indirect utility:[20]

$$U(£\,8) = 4(1 - \Pi_2)(1 - \Pi_7) \tag{4}$$

The study included SG and modified SG questions. In the SG questions (questions 1, 4 and 8) one of the lottery tickets was certain to yield a given sum. In the modified SG questions (questions 2, 3, 5, 6, 7 and 9) respondents compared two risky monetary lotteries. For example, in question 2 respondents were asked to compare two lotteries and set $(1 - \Pi_2)$, the probability of £ 30, in the first so that they were indifferent between the two choices.

18. In question 3: (0.5) U(£ 8) = $(1 - \Pi_3)$ U(£ 30)+ Π_3 U(£ 0). When U(£ 30) = 1 and U(£ 0) = 0, U(£ 8) = $2(1 - \Pi_3)$ by rearrangement.
19. The experiment is unable to consider what would have occurred had the losing outcome been fixed at other non-zero levels which, in turn, could have been linked back to zero. This was because at least one lottery in each of the questions included the possibility of receiving nothing so the losing outcome was always fixed at zero. Had it been possible to vary the losing outcome this would have been comparable to linking health states indirectly to death.
20. In question 7: $(1 - \Pi_7)$ U(£ 20) + Π_7 U(£ 0) = (0.5) U(£ 8) + (0.5) U(£ 0)
Set U(£ 0) = 0 and rearrange leads to U(£ 8) = $2(1 - \Pi_7)$ U(£ 20)
In question 2: $(1 - \Pi_2)$ U(£ 30) + Π_2 U(£ 0) = (0.5) U(£ 20) + (0.5) U(£ 0)
Set U(£ 30) = 1 and U(£ 0)=0 and rearrange leads to U(£ 20) = $2(1 - \Pi_2)$
Combining these two expressions gives U(£ 8) = $4(1 - \Pi_2)(1 - \Pi_7)$

3.2. *Results*

The null hypothesis (based on the predictions of EUT) that will be tested is H_0: direct utility = indirect utility.[21] In what follows we analyse the differences between direct and indirect utilities for the SG questions (questions 1, 4 and 8) independently from the modified SG questions (questions 2, 3, 5, 6, 7 and 9). Maintaining this distinction in the tables that follow will ensure that the results do not confound the effects arising for SG and for modified SG questions. At first glance, comparing the mean direct and indirect utilities for £ 8 in Table 4, the indirect utilities are less than or equal to the direct utilities. The results are considered further by using the Wilcoxon test in Table 5. Column 1 considers the direct utilities from question 8 compared against the indirect utilities from linking questions 4 and 1. For example, for question 8 row 1 column 1 shows that 16 respondents gave indirect utilities that were greater than the direct and row 2 shows that 24 respondents gave indirect utilities that were less than the direct. In the SG question, question 8, the direct and indirect utilities are not statistically different and lead to an acceptance of the null hypothesis (the 2-tailed P was 0.5816, row 6, column 1 in Table 5). In the modified SG questions, the differences are statistically significant for question 3 with the indirect utilities less than the direct (the 2-tailed P was less than 0.005 for all comparisons, row 6, columns 2 to 5 in Table 5). The differences are statistically significant for two out of the four comparisons for question 6 (row 6, columns 6 to 9 in Table 5) and for these two comparisons the indirect utilities are less than the direct as before.[22] This set of results are further summarised in Table 6. A decrease in the indirect utility suggests that respondents are willing to accept more risk in lotteries for which utilities are elicited indirectly than directly and are more risk-prone than expected. In addition, calculating the indirect utility using questions 9 and 7 appeared to widen the gap between the direct and indirect utility as shown by the falling P-values (in columns 8 and 9 in Table 5).

Table 4: Mean Direct and Indirect Utilities

	Direct utility		Indirect utility	
	mean	standard error	mean	standard error
Question 8	0.39	3.05×10^{-2}	0.38	3.09×10^{-2}
Question 3	0.52	3.92×10^{-2}	0.43–0.48	$3.53 \times 10^{-2} - 3.64 \times 10^{-2}$
Question 6	0.47	3.98×10^{-2}	0.43–0.48	$3.53 \times 10^{-2} - 3.64 \times 10^{-2}$

21. More specifically, the direct U(£ 8) is based on comparisons against £ 30 and £ 0. The indirect U(£ 8) is based on two questions which give the relationship between U(£ 8) and U(£ 20) and between U(£ 20) and U(£ 30). These two questions can then be linked through U(£ 20) to derive an the indirect U(£ 8).
22. It could be argued that it is more appropriate to compare question 6 against question 9, since both link the probabilities in the first and second lotteries. In such cases, the differences are statistically significant.

Table 5: The SG and modified SG in the monetary lottery study

Direct utility vs. indirect utility	Question 8 vs. questions 4 and 1	Question 3 vs. questions 2 and 5	Question 3 vs. questions 2 and 7	Question 3 vs. questions 9 and 5	Question 3 vs. questions 9 and 7	Question 6 vs. questions 2 and 5	Question 6 vs. questions 2 and 7	Question 6 vs. questions 9 and 5	Question 6 vs. questions 9 and 7
Number of respondents where indirect utility > direct utility	16	11	11	6	6	20	19	10	7
Number of respondents where indirect utility < direct utility	24	29	29	33	33	20	21	30	33
Number of respondents where indirect utility = direct utility	0	0	0	1	1	0	0	0	0
Total	40	40	40	40	40	40	40	40	40
Wilcoxon Z	−0.5511	−2.7756	−3.0915	−4.0609	−4.6057	−0.0403	−0.4503	−2.2850	−3.5619
2-tailed P	0.5816	0.0055	0.0020	0.0000	0.0000	0.9678	0.6525	0.0223	0.0004
H_0	accept	reject	reject	reject	reject	accept	accept	reject	reject

Table 6: Hypothesis tests on the H_0 for money lotteries

Question	H_0: Direct utility = indirect utility		Direction of discrepancy
	Accept	Reject	
Question 8	*		none
Question 3		*	Indirect < direct
Question 6 vs. questions 2 and 5	*		none
vs. questions 2 and 7	*		none
vs. questions 9 and 5		*	Indirect < direct
vs. questions 9 and 7		*	Indirect < direct

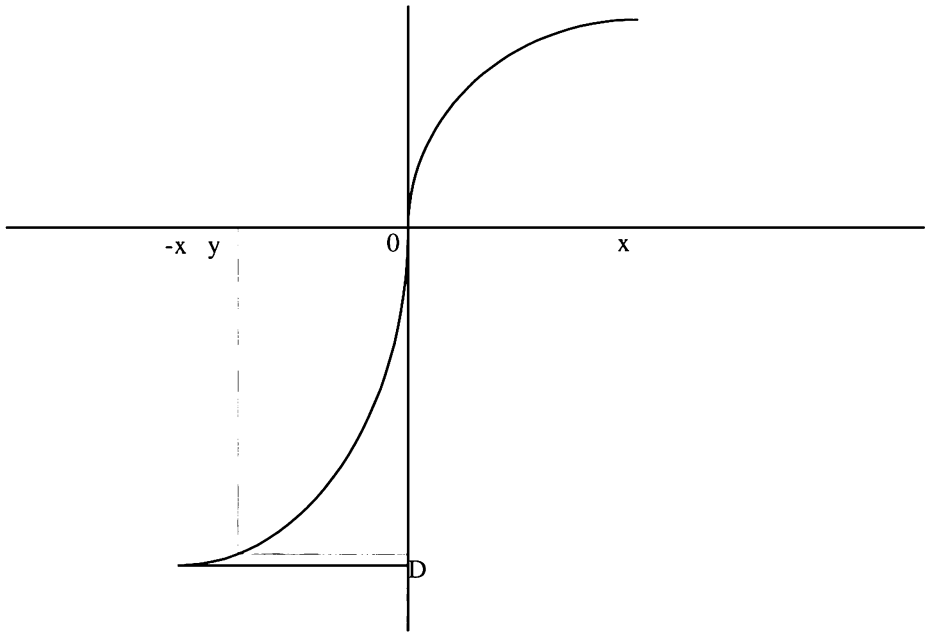
If we again assume that the degree of random error introduced by the linking process is small, then in totality, the results suggest that procedural invariance holds in some instances but not all and, further, that the differences are not in the same direction as the previous study. In summary, we reject H_0 in favour of H_1 in some instances and in these cases we observe that the direct utility > indirect utility.

4. DISCUSSION

Prospect theory (PT) was proposed as far back 1982 to explain differences in direct and indirect SG methods (LLEWELLYN-THOMAS, SUTHERLAND, TIBSHIRANI, CIAMPI, TILL and BOYD, 1982). This will be considered again here but will be expanded to embrace the predicted differences between direct and indirect methods that link through the failure, injury X in study 1, and the success, £ 20, in study 2.

PT incorporates a value function that weights outcomes, and a weighting function that weights probabilities. These introduce two key features. In the value function respondents reformulate mixed gambles of gains and losses into gains and losses relative to fixed a reference point (KAHNEMAN and TVERSKY, 1979). In the weighting function respondents underweight high probabilities and overweight low probabilities. An editing phase is imposed to ensure that stochastic dominance holds. Cumulative PT retains the notion of a reference point in the value functions but assumes a particular functional form for the weighting function based on ranking the probabilities (i. e. a decumulative probability weighting function). This ensures that preferences do not violate stochastic dominance and hence is preferred by economists (TVERSKY and KAHNEMAN, 1992).

Figure 4: Prospect Theory



The shape of the value function $v(\cdot)$ is shown in Figure 4 where x is the outcome (in study 1x denotes injuries and in study 2, x denotes money) and $x = 0$ is the reference point. There is diminishing sensitivity to gains or losses as these increase relative to the reference point. In particular the curve is concave for gains ($v''(x) \leq 0$) and convex for losses ($v''(-x) \geq 0$). Hence the preference function is quasi-concave under PT and respondents give relatively greater emphasis to small-to-medium gains and losses than they do larger gains and losses. In addition, there is a pronounced asymmetry between gains and losses, termed loss aversion. For an equivalent change in money or health status, respondents experience a greater feeling of displeasure from losses than of pleasure from gains of the same magnitude, i. e. $v'(x) < v'(-x)$, where $v(-x) = D$ in Figure 4.

We now show that the combination of diminished sensitivity and loss aversion can account for systematically different predictions for indirect methods linked through the success or failure. In study 1, the indirect method links questions through the failure, injury X. Consider again question 2 in study 1 and let $\prod_X^e$ denote the probability of failure under EUT and $\prod_X^p$ denote the probability under PT. In question 2, if Treatment C or D fail injury X occurs, rather than death, ceteris paribus this should make treatment D appear more attractive. Under EUT a respondent should be willing to bear more risk of failure, but the insensitivity to losses under PT leads to $\prod_X^p < \prod_X^e$. In terms of equation (3) the indirect VOSL would increase relative to the direct VOSL. Similarly, PT would predict that the indirect U(£ 8) would increase relative to the direct U(£ 8) had this question been asked in study 2. This explanation is consistent with the finding that the indirect VOSL is greater than the direct VOSL.

In study 2, the indirect method links questions through the success, £ 20. Consider again question 7 in study 2 and let $\prod_7^e$ denote the probability of failure under EUT and $\prod_7^p$ denote the probability under PT.In question 7 they have the chance of £ 20, rather than £ 30, *ceteris paribus* this should make the lottery appear less attractive. Under EUT a respondent should be less willing to bear risk of failure and should increase the probability of success but the insensitivity of responses to gains under PT suggests that $\prod_7^p > \prod_7^e$ and $1 - \prod_7^p < 1 - \prod_7^e$. This insensitivity has the potential to decrease the indirect U(£ 8) below that of the direct, as in equation (4) when $1 - \prod_7^p < 1 - \prod_7^e$. However, since losses loom greater than gains, the opportunity for this insensitivity to be transmitted into the responses is much less in the case of gains. This would suggest that responses to the indirect methods are less than or equal to the direct. Similarly, PT would predict that the indirect VOSL would be less than or equal to the direct VOSL had this question been asked in study 1. This explanation is consistent with the finding that the indirect utility is less than or equal to the direct utility in study 2.

Given the nature of study 2, it is possible that a further phenomenon is at work.[23] Regret theory predicts that changes in the framing of the questions can change respondents' attitudes towards risk (BELL, 1982; FISHBURN, 1982; LOOMES and SUGDEN, 1982). For instance, regret-aversion implies that respondents give relatively greater em-

23. Study 1 did not vary the visual stimulus given to respondents and so is not affected by these issues.

phasis to large gains and losses than they to do small-to-medium gains and losses. The implication of this is that changing the position of outcomes in the outcome matrix (i. e. changing the juxtaposition of consequences) keeping all else constant (i. e. the size of the outcomes and overall probability) can change preferences. The effect of juxtaposition is denoted by the conditional probability of a non-zero outcome occurring in second lottery given that a non-zero outcome occurs in first lottery, ω.[24,25]

For example, question 5 in Figure 3 is an example where $\omega = P(£8|£20) = 1$, the maximum overlap, and question 7 is an example where $\omega = P(£8|£20) = 0$, the minimum overlap. Regret theory predicts that responses become more risk-prone as ω tends towards zero. For example, the net advantage of choosing the first over the second lottery is greater in question 7 than question 5,[26] and respondents increase the probability of £20 in question 7 as they would if their responses are more risk-prone,[27] and this leads to a decrease in indirect utility.

The results suggest that the modified SG opened up the possibility for juxtapositioning effects in questions 9 and 7. For example, when listing the pairs of questions that form the indirect utilities in terms of the least to most difference between the direct and indirect utilities (with the least difference indicated by a higher P value) they follow the pattern 2 and 5, 2 and 7, 9 and 5, 9 and 7 for question 6 (with the last two pairs being equally ranked for question 3, see Table 5). This emphasises the sensitivity of responses to slight changes in the framing of questions and, as such, violates the principle of procedural invariance.

5. CONCLUSIONS

In this paper, we set out to test the EUT assumption that outcomes will be procedurally invariant in two studies designed to elicit people's preferences for risky choices. The first study utilised the results from the "CV/SG Chained" method developed by CARTHY et al. (1999) to elicit preferences over health states, while the second study used the SG method to measure preferences for monetary lotteries. EUT predicts equivalence for directly and indirectly elicited utilities/values under the same method. We note, however,

24. For instance, P(y|z) is the conditional probability of y given z occurs.
25. Unlike the monetary experiment, the safety study in study 1 (or indeed other health valuation studies using conventional SG) did not vary the juxtapositioning of outcomes and the modified SG question had overlap = 1 in all cases. Regret theory, therefore, cannot be used to explain variations in risk attitude arising from changes in the juxtaposition of outcomes in study 1.
26. In particular, respondents experience greater regret/rejoicing from the difference between £20 and £0 in question 7 than the combined regret/rejoicing from £20 and £8 and £8 and £0 in question 5.
27. It has been questioned whether regret applies to comparisons of health states involving death since there is unlikely to be time to regret when you face immediate death (JONES-LEE, 1989, p. 136). However, ENEMARK (1994) argued that even though the question may involve the chance of immediate death, respondent's fear of regret can have an impact and it follows that provided there are differences between the size of the payoffs regret may have an impact.

that this is not the case in our studies. Perhaps more interestingly, we find that the differences are not in the same direction across the two. In the case of the safety study, the indirect VOSL are greater than the direct VOSL. Meanwhile, the reverse is true for the utilities in the monetary lotteries study. PT appears to explain the patterns observed, predicting that the indirect values/utilities will be greater than the direct for questions linked through the failure and indirect values/utilities will be less than or equal to the direct for questions linked through the success. To our knowledge this distinction between the success or failure has not been discussed in the literature before although past studies have found similar patterns in the responses (MORRISON, 1996; RUTTEN-VAN MÖLKEN, 1995; LLEWELLYN-THOMAS et al., 1982). This would suggest that if we assume the degree of random error introduced by the linking process is small in both studies, respondents' preferences are better explained by theories other than EUT.

Although PT appears to explain the general trends arising in these studies there are a number of issues that need to be addressed in future research. For instance, if respondents perceive there to be a reference point in the modified SG questions what is this? In SG questions that offer the choice between a certain and risky outcome, the certain outcome is assumed to be the reference point. In the modified SG questions that offer the choice between two risky outcomes, the reference point is less clear. In the analysis carried out in this paper, the discussion assumed that reference point to be the expected utility of treatment or lottery whose probability is fixed. Future work is needed to investigate if there is any special decision heuristics used in such questions.

Another issue is the extent to which the viewpoint adopted in the SG responses are affected by decisions involved. Responses to SG questions that offer the choice between a certain and risky outcome are thought to be influenced by the *certainty effect* where high probabilities are underweighted (TVERSKY and KAHNEMAN, 1992). This has been used as an argument for using the modified SG question which involves only risky outcomes and so is less prone to this effect. The impact upon the SG responses depends upon whether the outcomes are viewed as gains or losses. Until recently the empirical evidence suggested that the outcomes were viewed as gains only but more recently this has been questioned. To illustrate this let $\prod_X^p$ denote the probability of failure set by respondents in the SG questions under PT, which incorporates the certainty effect in its weighing function, and let $\prod_X^e$ denote the SG responses set under EUT. If respondents view the outcomes in the SG questions as a gain, they will underweight the probabilities of success in the risky treatment/lottery and will be more risk-averse in their responses, i. e. $\prod_X^p < \prod_X^e$. The VOSL, therefore, will decrease as respondents are unwilling to choose the risky alternative. MCCORD and DE NEUFVILLE (1986) found support that respondents view monetary lotteries as gains, as did WAKKER and DENEFFE (1992) for questions involving life years. A recent paper, however, challenges this view for questions involving health states. LAW, PATHAK and MCCORD (1998) found support that respondents view questions involving health states as losses. If respondents view the outcomes in the SG questions as a loss, they will underweight the probabilities of this loss in the risky treatment/lottery and will be more risk-prone in their responses, i. e.

$\prod_X^p > \prod_X^e$. The VOSL, therefore, will increase as respondents are more willing to choose the risky alternative. The difference between SG questions involving life years or health states has not been discussed in the literature before. Further research is needed on whether these differences are robust across different samples and to investigate the origins of these differences.

To what extent can policy be changed to accommodate the predicted asymmetry between questions that link through the failure or success? Indirect methods that link through the failure are now commonly used to elicit the general public's preferences over minor injuries (DRUMMOND, O'BRIEN, STODDART and TORRANCE, 1997). These indirect questions avoid comparisons with death in the initial stages, which many respondents find too extreme, and allow responses to be linked through death in later questions. The finding that the indirect VOSL is greater than the direct VOSL has led to policy makers taking the lower limit of these values given this is the more conservative measure (CARTHY et al., 1999). Indirect methods that link through the success are less common, mainly because the rationale to avoid comparisons with normal health in the initial stages that this would imply is less compelling. Given the trend to make questions more realistic to the decisions involved (BEATTIE et al., 1998b), it is easy to imagine that questions linked through the success will become more popular as many treatments do not lead to full recovery. Policy makers then have to decide whether again to use the conservative measure, the indirect VOSL in this case, or whether to use the upper limit based on the direct VOSL. In either case, the problem is less acute, since the differences between the direct and indirect VOSL appears much less in questions that link through the success. From a theoretical perspective, if respondents are a) acting as EUT maximisers and b) truthful and accurate, then the direct VOSL will be an appropriate measure of utility. We could then assume the observed differences between the direct and indirect VOSL arise from error introduced simply by the decomposition process itself. However, if both a) and b) do not hold it is questionable if any VOSL measure is a true reflection of a respondent's utility.

Turning now to more general policy implications, if the failure of procedural invariance is a pervasive feature of these (and other) methods, then any elicitation mechanism is likely to generate quite different results depending on whether values are derived directly or indirectly. To date, there exists an insufficient number of studies incorporating both direct and indirect values to determine whether the error is systematic and, if so, in what direction and whether it may be possible to recalibrate responses to account for these differences.

At the moment, it seems that the best way forward would be to build internal consistency checks into all future studies based on EUT assumptions, in order to build up an impression of the degree to which procedural invariance affects outcomes under a variety of elicitation methods. BARON and UBEL (1999) go further and suggest that if procedural invariance fails then subjects should be confronted with the results and asked to reconcile them. This has already been attempted with some success by ROBINSON, DOLAN and WILLIAMS (1997) for reconciling the differences between the Time Trade-Off

and Visual Analogue Scale methods, but care is needed to ensure that respondents do not feel pressured to manufacture answers that they think the interviewer wants to hear. If procedural invariance fails the responses may merit closer inspection to identify an alternative theory that may better fit observed responses. Whether this theory is found within economics or from other disciplines, such as psychology, is unknown. At the very least, though, such findings suggest that the need for closer collaboration between the disciplines is paramount if we are to develop economic models of preferences which are "in tune" with the data.

REFERENCES

BARON, J. and P.A. UBEL (1999), "Inconsistencies in person-tradeoff and difference judgments of health-state utility", *Working Papers*, Pittsburgh University.

BEATTIE, J., J. COVEY, P. DOLAN, L. HOPKINS, M. JONESLEE, G. LOOMES, N. PIDGEON, A. ROBINSON and A. SPENCER (1998a), "On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 1 – Caveat Investigator", *Journal of Risk and Uncertainty*, 17, 5–25.

BEATTIE, J., S. CHILTON, R. COOKSON, J. COVEY, L. HOPKINS, M. JONESLEE, G. LOOMES, N. PIDGEON, A. ROBINSON and A. SPENCER (1998b), *Valuing health and safety controls: a literature review, Contract Research Report*, London, 171.

BELL, D.E. (1982), "Regret in Decision Making under Uncertainty", *Operations Research*, 30, 961–981.

BOYLE, J., W.H. DESVOUSGES, F.R. JOHNSON, R.W. DUNFORD and S.P. HUDSON (1994), "An Investigation of Part-Whole Biases in Contingent Valuation Studies", *Journal of Environmental Economics and Management*, 27, 64–83.

CAMERER, C. (1995), "Choice under risk and uncertainty", in: Kagel, J.H. and A.E. Roth (eds.), *The Handbook of Experimental Economics*, New Jersey.

CARTHY, T., S. CHILTON, J. COVEY, L. HOPKINS, M. JONES-LEE, G. LOOMES, N. PIDGEON and A. SPENCER (1999), "On the Contingent Valuation of Safety and the Safety of Contingent Valuation: Part 2 – The CV/SG 'Chained' Approach", *Journal of Risk and Uncertainty*, 17, 187–213.

DIAMOND, A., J.A. HAUSMAN, G.K. LEONARD and M.A. DENNING (1993), "Does Contingent Valuation Measure Preferences? Experimental Evidence", in: J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*, Amsterdam.

DRUMMOND, M.F., B. O'BRIEN, G.L. STODDART and G.W. TORRANCE (1997), *Methods for the economic evaluation of health care programmes*, second edition, Oxford.

ENEMARK, U. (1994), "Alternatives to expected utility theory in clinical decision-making: a review of prospect theory, regret theory and disappointment theory", *Studies in Health Economics*, 5, Lund University, Lund.

FARQUHAR, P.H. (1984), "Utility Assessment Methods", *Management Science*, 30, 1283–1300.

FISHBURN, P. (1982), "Nontransitive Measurable Utility", *Journal of Mathematical Psychology*, 33, 1–27.

JONES-LEE, M.W. (1989), *The Economics of Safety and Physical Risk*, Oxford.

JONES-LEE, M.W., M. HAMMERTON and P.R. PHILIPS (1985), "The Value of Safety: Results of a National Sample Survey", *Economic Journal*, 95, 49–72.

KAHNEMAN, D. and J.L. KNETSCH (1992), "Valuing Public Goods: The Purchase of Moral Satisfaction", *Journal of Environmental Economics and Management*, 22, 57–70.

KAHNEMAN, D. and A. TVERSKY (1979), "Prospect Theory: An Analysis of Decision Under Risk", *Econometrica*, 47, 263–291.

LAW, A.V., D.S. PATHAK and M.R. MCCORD (1998), "Health Status Utility Assessment by Standard Gamble: a Comparison of the Probability Equivalence and the Lottery Equivalent Approaches", *Pharmaceutical Research*, 15, 105–109.

LLEWELLYN-THOMAS, H., H.J. SUTHERLAND, R. TIBSHIRANI, A. CIAMPI, J.E. TILL and N.F. BOYD (1982), "The Measurement of Patients' Values", *Medical Decision Making*, 2, 449–462.

LOOMES, G. and R. SUGDEN (1982), "Regret Theory: an Alternative Theory of Rational Choice under Uncertainty", *Economic Journal*, 92, 805–824.

MACHINA, M.J. (1987), "Choice under Uncertainty: Problems Solved and Unsolved", *Economic Perspectives*, 121–154.

MCCORD, M. and R. DE NEUFVILLE (1986), "Lottery Equivalents: Reduction of the Certainty Effect Problem in Utility Assessment", *Management Science*, 32, 56–60.

MITCHELL, C. and R.T. CARSON (1989), *Using Surveys to Value Public Goods: The Contingent Valuation Method*, Washington D.C.

MORRISON, G.C. (1996), *An assessment and comparison of selected methods for eliciting quality of life valuations from the public*, York.

ROBINSON, A., P. DOLAN and A. WILLIAMS (1997), "Valuing Health Status using VAS and TTO: What lies Behind the Numbers?" *Social Science and Medicine*, 45, 1289–1297.

RUTTEN-VAN MÖLKEN, M. (1995), *Costs and effects of pharmacotherapy in asthma and COPD*, Ph.D. thesis, University of Maastricht.

SPENCER, A. (1998), *Value elicitation techniques: methodological issues involved in two-stage procedures*, D.Phil. thesis, University of York.

SMITH, K. and L.L. OSBOURNE (1996), "Do Contingent Valuation Estimates Pass a 'Scope' Test? A Meta Analysis", *Journal of Environmental Economics and Management*, 31, 287–301.

TVERSKY, A. and D. KAHNEMAN (1992), "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, 5, 297–323.

TVERSKY, A., S. SATTATH and P. SLOVIC (1988), "Contingent Weighting in Judgement and Choice", *Psychological Review*, 95, 371–384.

VON NEUMANN, J. and O. MORGENSTERN (1994), *Theory of Games and Economic Behaviour*, Princeton, NJ.

WAKKER, P. and D. DENEFFE (1992), "Eliciting von Neumann-Morgenstern Utilities when Probabilities are Distorted or Unknown", *Management Science,* 42, 1131–1150.

SUMMARY

Expected Utility Theory (EUT) underlies the Standard Gamble (SG) method for eliciting people's preferences towards safety policy and risky treatments. Increasingly surveys using this method decompose the SG questions into two or more intermediate questions. Under the EUT assumption of procedural invariance these indirect responses are theoretically equivalent to those elicited directly through one question. We investigate the issue empirically in two studies, both of which find that procedure invariance is violated. Despite this, and the differences in the direction of the discrepancies that we observe in these studies, we show how these results are in fact consistent with an alternative theory to EUT – Prospect Theory – and discuss the implications for future policy and research.

ZUSAMMENFASSUNG

Die Nutzenerwartungstheorie bildet die Grundlage der Glücksspielmethode zur Feststellung der individuellen Präferenzen in Fragen der Sicherheitspolitik und riskanter Behandlungen. Immer mehr Erhebungen nach dieser Methode zerlegen die Glücksspielfragen in zwei oder mehr Zwischenfragen. Nach der nutzenerwartungstheoretischen Annahme der Verfahrensinvarianz sind diese indirekten Antworten theoretisch den direkten Antworten auf nur eine Frage gleichwertig. Wir untersuchen die Frage empirisch in zwei Studien, die beide zu dem Ergebnis gelangen, dass die Verfahrensinvarianz verletzt wird. Dennoch und trotz der unterschiedlichen Richtung der festgestellten Diskrepanzen zeigen wir, dass diese Ergebnisse mit einer Alternative zur Nutzenerwartungstheorie – der Aussichtstheorie – übereinstimmen und erörtern die Folgen für die zukünftige Politik und Forschung.

RESUME

La théorie de l'espérance de l'utilité est sous-tend la méthode du "standard gamble" (SG), laquelle a pour objectif d'amener les individus à révéler leurs préférences par rapport à des politiques de sécurité ou des traitements risqués. De plus en plus, les enquêtes qui recourent à cette méthode décomposent les questions du SG en deux questions intermédiaires, voire plus. Sous la condition d'invariance de la procédure de la théorie de l'espérance de l'utilité, ces réponses indirectes sont théoriquement équivalentes à celles obtenues directement à l'aide d'une seule question. Nous abordons la question empiriquement dans deux études qui montrent toutes deux que la condition d'invariance de la

procédure est violée. Malgré cela, et malgré les différents types d'écarts que nous observons dans nos études, nous montrons comment ces résultats sont en fait cohérents avec une théorie alternative à l'utilité espérée – la théorie prospective –. Nous discutons ensuite les implications pour les politiques et recherches futures.