How would you like your gain in Life Expectancy to be provided? An experimental approach

J. Seested Nielsen[1], Susan Chilton[2],
Michael Jones-Lee[2], Hugh Metcalf[2]

[1] Department of Health Economics, University of Southern Denmark
[2] Business School, University of Newcastle

No 2009/09

Newcastle Discussion Papers in
Economics: ISSN 1361 - 1837

How would you like your gain in life expectancy to be provided?

An experimental approach.

J. Seested-Nielsen

University of Southern Denmark

&

Chilton, S*., Jones-Lee, M. and Metcalf, H.

Newcastle University Business School - Economics

Abstract

This paper reports the results of an empirical study investigating people's preferences over three different types of perturbation to their survival function, each perturbation generating the same gain in life expectancy. Preferences over the three different perturbations were found to be distributed more or less evenly across the subject pool. Use of a novel experimental methodology generated economically consistent and intuitively plausible responses to (necessarily) hypothetical questions concerning improvements in life expectancy by first allowing respondents to gain experience while making similar choices in an incentivized setting involving financial risk. The results demonstrate the potential for economic experiments to contribute to the development of more robust methods for policy evaluation in domains where physical risk is an important factor.

JEL Classification J17; D6

* Correspondence: e-mail s.m.chilton@ncl.ac.uk

1. Introduction

Not surprisingly, many people understand an increase in life expectancy to mean an extension of the time for which they will survive, or more specifically, as an "add-on" at the end of life. This interpretation is quite natural, given that an increase in the endowment of most other quantifiable "goods" - such as income or leisure time - can quite legitimately be regarded as an addition to a pre-determined level. But of course given that life expectancy is, strictly speaking, the mean of a probability distribution of remaining survival time, then a gain in life expectancy represents an increment (or "add-on") to actual survival time *only if* the latter is taken to be an average over a large population. In addition to this and perhaps more significantly, as noted in Hammitt (2007), a given gain in life expectancy can in principle be generated by any one of an infinite number of different perturbations in an individual's survival function (which specifies the probability that the individual will still be alive at any given age). Thus, for example, a one-month gain in life expectancy could be generated either by a reduction in the risk of death during the coming year, *or* by an ongoing reduction over the next twenty years, *or* by a reduction that only takes effect in later years of life, and so on.

But does this actually matter? Viewed from the perspective of health and safety policy, the answer would appear to be unequivocally in the affirmative. Thus, for example, in the context of environmental health, over recent years there has been a heavy focus on defining and estimating the monetary "Value of a Life Year" (or VOLY) that is, the value of a one-year gain in life expectancy "in normal health". The US Office of Management now recommends that cost-benefit analyses undertaken in their agencies (including the US Environmental Protection Agency) apply both the VOLY and the VSL approach (Sunstein 2004). However, to the best of our knowledge, with a few exceptions[1] the work aimed at deriving such values has failed to take account of the fact that, as already noted, a year's gain in life expectancy can be generated in any one of a number of different ways and that those affected may have a marked preference for – and hence place a significantly higher value on – one way of generating the gain rather than another.

Furthermore, in addition to recent work directed at placing a monetary value on a life-year in the environmental health context, the past four decades have also seen considerable research effort – in some cases supported by substantial government funding – directed at placing monetary values on

[1] For example, Mason et al (2009).

different kinds of safety improvement in various contexts including transport, the workplace, nuclear power generation and so on. While this research has inevitably required a clear specification of the time pattern of the safety improvement concerned (i.e. whether it applies only to, say, the forthcoming year or, by contrast, over a more extended future period) the range of alternative future time paths actually considered has been rather limited.

All of this strongly suggests that investigation of the public's attitudes to the nature of the future time- path of health improvement or risk reduction is a matter of considerable importance. To put it bluntly, failure to address this issue leaves open the possibility that some types of healthcare or safety improvement could be seriously undervalued- and hence underprovided- relative to others which in fact warrant considerably lower priority (a more detailed discussion of the underlying rationale for this observation is deferred until later in this paper). The empirical study reported in this paper was therefore aimed at investigating people's preferences over three different (hypothetical) programmes which all deliver the same gain in life expectancy (LE) but where the risk reductions are distributed differently. There would appear to be two principal reasons why an individual might have a marked preference for one way of generating a given gain in life expectancy rather than another. The first, and obvious, possibility is that future reductions in the probability of death may be subject to some form of discounting relative to more immediate reductions. The second hypothesis concerning the reason why the nature of the probability distribution for the time of death may matter is that even though two different programmes could deliver precisely the same change in life expectancy, other moments of the probability distribution - such as the variance, (i.e. the dispersion around the mean) - could be expected to influence the valuation as well.

Earlier stated-preference empirical studies designed to examine people's valuation of a gain in life expectancy have generally failed to specify the way in which the gain is generated (see for example, Moore and Viscusi (1988), Johannesson and Johansson (1996) or Morris and Hammitt (2001)). This is to a large degree understandable, given the inherent difficulties that would be encountered in communicating to members of the public the essential properties of a survival function. A major aim, therefore, was to develop a protocol which addressed this almost universal shortcoming directly. Following extensive piloting, the main study had embedded within it a novel experimental procedure designed to improve respondent's understanding of the basic concepts involved by first allowing them to gain experience of making choices that were conceptually similar to those involving survival functions, but which in fact related to financial gambles offering the possibility of real monetary payouts to respondents.

The remainder of the paper is organised as follows. The next section describes the three different types of perturbation in the individual survival function that underpin the study and outlines the experimental design that was actually employed. Section 3 then provides a more detailed description of the experimental procedures used a) to familiarise respondents with the basic concepts to be used in the study and b) to elicit preferences over the three different ways of generating a given gain in life expectancy; results are then presented in Section 4; validity is examined in Section 5 and in Section 6 a discussion of results is presented. The final Section concludes.

2. Theory underpinning the experimental design

The general properties of the mortality and survival functions that form the basis of our analysis are well-established (see, for example, Jenkins, S. P. (2005)) and will be familiar to most readers. However, for the sake of clarity and completeness it would seem appropriate to provide a brief summary of the definition and key properties of these functions which are as follows:

f(t) – Probability density function for time t (≥ 0), of individual's death.

S(t) – "Survival function" which gives the probability that individual survives until at least time t.

θ(t) – "Hazard function" which gives the probability density of dying at time t conditional on surviving until time t.

From these definitions it then follows that:

$$S(t) = 1 - \int_0^t f(\tau)d\tau \quad (1)$$

and

$$\theta(t) = \frac{f(t)}{S(t)}. \quad (2)$$

Differentiating through equation (1) with respect to t then gives

$$f(t) = -S'(t). \quad (3)$$

In turn, by definition the individual's life expectancy, LE, is given by

$$LE = \int_0^\infty tf(t)dt. \quad (4)$$

From equations (3) and (4) using integration by parts it follows immediately that

$$LE = \int_0^\infty S(t)dt, \quad (5)$$

that is, LE is given by the area under the survival function. Clearly, then, a specific gain in LE can be generated as a result of an infinite number of different perturbations in the hazard function, resulting in different changes to the underlying survival function.

While the above general definitions are set in a continuous context, for the sake of simplicity - as explained below – the experiment actually employed in the study split possible remaining survival time into a limited number of discrete time periods. With appropriate modification, the definitions set out above also apply in the discrete time case. In particular, in the discrete time case the hazard function takes the form of a vector of "hazard rates", (p_1, p_2,…p_i,…), where p_i denotes the probability of dying at the beginning of the i^{th} period, conditional on having survived until then.

2.1. Description of the three life expectancy programmes

The experimental design involves three different programmes which change the hazard function in specific ways but which each deliver a six-month gain in LE, i.e. we do not investigate preferences for gains in LE, as such, but instead focus on preferences over different ways of generating a particular gain in LE The reason why a six month gain was chosen is because, compared with the risk reductions (and implied life expectancy gains) in a standard VSL-surveys, even a six month gain in life expectancy is a substantial good(for someone of average age, the risk reduction in a typical VSL survey would generate a gain in life expectancy of a few hours, or at most, days and certainly be less than a month). Hence, with the VSL-literature as the starting point it is hard to argue that a one year gain would be marginal. In addition, it is not possible by reducing only the hazard rate in the first decade to deliver a gain in life expectancy of one year for a 40-year old, since it would result in an implied survival probability greater than one. Finally, since for the average 40-year old in the UK, even a complete elimination of the risk of death over the coming decade would result in a gain in life expectancy of only about nine months, it was decided to set the gain at six months, to allow more scope for trading off gains in life expectancy than a lesser gain would provide.

The reasoning behind selection of the three programmes is in part pragmatic. While, as already noted, there are an infinite number of different possible perturbations in the probability density function for an individual's remaining length of life that generate a given gain in life expectancy, clearly from a practical point of view it would not be feasible to investigate other than a small subset, though in principle the methodology we are going to introduce is generalizable. However, the three programmes are representative of programmes that can actually be implemented in practice and represent perturbations which, in a sense, span a natural spectrum. In particular, programme OR offers a One-off Reduction in hazard rate for the coming period (as has frequently been assumed to be the case in theoretical and stated-preference work related to the valuation of safety improvement – see, for example, Jones-Lee (1976)), while the alternatives, programmes AR and PR, deliver an ongoing and sustained reduction in hazard rates over the remainder of an individual's life. In programme AR the Absolute Reduction in risk is constant, while in programme PR it is Proportional to the initial level of the hazard rate (as evidence suggests is the case, for example, in relation to reductions in particulate air pollution (Pope et al. 1995)). The three programmes are described formally in Figure 1.

Figure 1: Description of the three LE programmes applied in the experiment

OR; a one-off risk reduction δ ($\delta \in$ (-p,0)) in the hazard rate after which the hazard rates return to their previous level
AR; a permanent constant *absolute* risk reduction ω ($\omega \in$ (-p,0)) in hazard rates
PR a permanent constant *proportional* risk reduction k ($k \in$ (-1,0) in hazard rates

For practical experimental purposes and ease of understanding for respondents, the design is based around a five-period discrete-time model representing the statistical hazard rates in the five forthcoming decades for an average 40-year-old UK individual[2]. The advantages of restricting the design to only one age group are that the same hazard rates can be used for all respondents and the results are controlled for any effects of age. The principal aim of this study was to develop a methodology to overcome the shortcomings in previous LE studies discussed earlier but, given its success, it could in principle be adapted to other age groups to increase its policy relevance.

The basic definition of remaining life expectancy, LE, for a 40-year-old is given in equation (6)[3], in which p_i represents the existing hazard rate for a given decade:

$$LE = 10(1-p_{40}) + 10(1-p_{40})(1-p_{50}) + 10(1-p_{40})(1-p_{50})(1-p_{60}) + .. \quad (6)$$

with p_{90}=1, implying that for the purposes of the experiment, all people will die not later than age 90.

Equations (6) to (9) express the LE calculated in discrete time for the three different programmes specified in Figure 1.

$$LE_{OR} = 10(1-(p_{40}+\delta)) + 10(1-(p_{40}+\delta))(1-p_{50}) + ..,\delta \in \left(-p_{40},0\right) \quad (7)$$

$$LE_{AR} = 10(1-(p_{40}+\omega)) + 10(1-(p_{40}+\omega))(1-(p_{50}+\omega)) + ..,\omega \in \left(-p_i,0\right) \quad (8)$$

[2] *Had we used, for example, an annual hazard rate per year, this would have resulted in games of 50 bags, each containing 10,000 cards. We took the view that this would be an unfeasibly complex distribution for respondents to deal with.*

[3] Alternatively, but equivalently, LE can be defined as,

$$LE_{40} = 0p_{40} + 10(1-p_{40})p_{50} + 20(1-p_{40})(1-p_{50})p_{60} + ..$$

$$LE_{PR} = 10(1 - p_{40}(1+k)) + 10(1 - p_{40}(1+k))(1 - p_{50}(1+k)) + .., k \in (-1, 0) \quad (9)$$

As already noted, the hazard rate reductions δ, ω and *k* were set so as to ensure that all three programmes OR, AR and PR generate a six month gain in LE. The distribution of the risk reductions that was actually presented to respondents is shown in Table 1. All probabilities are expressed to a base of 1000.

Table 1. The risk figures describing the three programmes

Age[1]	**40-50**	**50-60**	**60-70**	**70-80**	**80-90**
Initial hazard rate (p_i)	20/1000	48/1000	121/1000	347/1000	652/1000
OR *changes* in hazard rate (δ)	14/1000	0	0	0	0
AR *changes* in hazard rate (ω)	5/1000	5/1000	5/1000	5/1000	5/1000
PR *changes* in hazard rate (kp_i)	1/1000	2/1000	5/1000	15/1000	28/1000

1. The figures represent an average 40-year-old individual (simple average between male and female, truncated at the age of 90) from the UK and are based on data from the Government Actuary's Department (mean 2003-2005). Refer to Appendix A for further description.

From Table 1 it is apparent that the total reduction in the hazard rates differs according to programme (14/1000 in OR, 25/1000 in AR and 51/1000 in PR) and equations (7) to (9) illustrate why this is so. For example, examining equation (7) it is clear that a change in hazard rate in the first decade influences the remaining survival probabilities as well. For a better intuition we will separate the total gain in LE into two beneficial effects, referring to the direct change in hazard rate as the "safety effect" and the indirect effect on remaining survival probabilities as the "survivor effect". A reduction in the hazard rate in any given period not only reduces the risk of death in that period (the 'safety effect') but also increases the probability of still being alive in later periods (the 'survivor effect). Consequently, a reduction in hazard rate in the first decade brings more 'survivor effect' than a reduction in hazard rate of the same magnitude in, say, the fourth decade.

But what will the respondents take into account when they make their decision? The increase in life expectancy afforded by each of the three programmes is the same. Under the assumption that individuals are risk-neutral and have an individual discount rate equal to zero we would expect people to be indifferent between the three distributions i.e. OR=AR=PR. Yet as shown in Cropper et al. (1994) and Viscusi & Aldy et al. (2003) people may well apply positive discount rates in assessing the timing of physical risk reductions for themselves or in assessing the relative desirability of saving people now or in the future. This in itself could result in a marked preference for OR over AR and AR over PR.

Timing could also matter with respect to the age at which an individual enjoys his/her risk reduction. For example, there has been evidence that the relationship between the valuation of a change in mortality risk and age can be depicted as an inverted U-form, hence people could have stronger preferences for "mid-life" risk reductions (see for example Aldy and Viscusi 2008 and Jones-Lee et al. 1985). This could results in a preference for AR

We could also imagine that people would show preferences in accordance with a utility function which weights the different periods against each other in a manner that differs from that characterized by risk-neutrality and no discounting. For instance, we could anticipate that for some individuals, the baseline risk in the first period is so small that it would be literally treated as zero – in which case any risk reduction in that period would have no value to the person concerned. By contrast, some other individuals could be expected to be uninterested in gains in the final period because of the high baseline risks in this period, the likelihood of not reaching that point in life and a high rate of time discounting.

In addition, we could expect other moments of the probability density function for time of death to influence the respondents' preferences. Some individuals could express concerns about whether they would be able receive the gains in life expectancy, implying that the more certain they were to receive the gains the better. In this paper we will use the variance, i.e. the dispersion around the mean, as a measure of the riskiness of receiving the benefit. All other things held equal, we would expect risk-averse individuals to have a preference for the programme with the smallest variance.

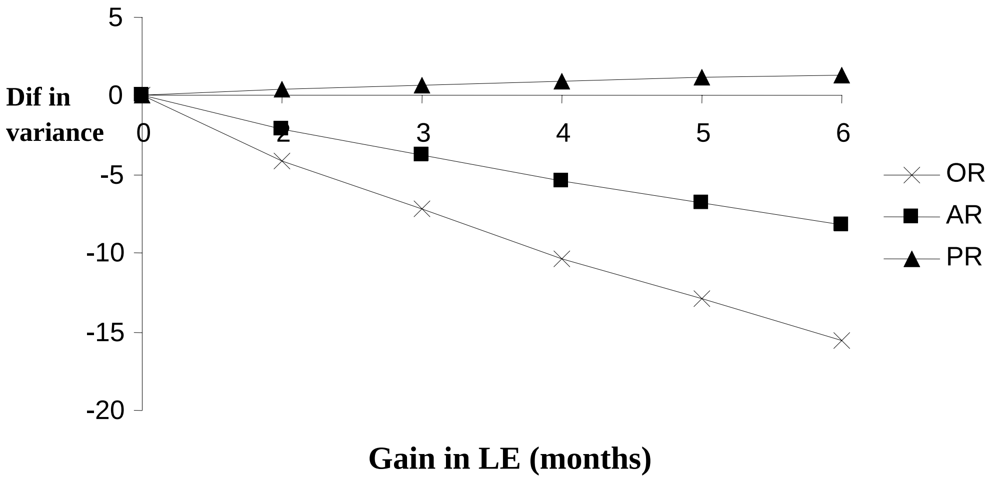
A calculation of the variance of the three programmes shows that programme OR offers the largest reduction in variance followed by AR, whereas programme PR actually increases the variance. More specifically, Figure 3 depicts the change in variance as a function of the gain in LE[4]. For example, a six-month gain in LE from programme OR results in a decrease in variance of about 15,whereas the same gain in LE from programme AR is associated with a reduction in variance of only about half that magnitude. By contrast, programme PR would, in fact, increase variance by roughly 2 units for a gain in LE of six months.

---

[4] Variance is calculated applying the standard formulae (Berry & Lindgren, 1996), which in this case is:

$$V(LE_{40}) = (0 - \overline{LE})^2 p_{40} + (10 - \overline{LE})^2 (1 - p_{40}) p_{50} + \dots -$$

Figure 3. Change in variance. The three programmes



3. Methods

Trading off different LE gains requires that respondents are familiar with the concept of probability and are willing to express their strength of preference for different distributions of probability changes. The core of our experimental design is 1) to elicit choices in relation to the three programmes via a series of paired comparisons, using a variant of the risk-risk trade-off method (Viscusi et al., 1991) and 2) to find the strength of preferences by deterring respondents' indifference points between appropriately modified versions of the three programmes in pairs.

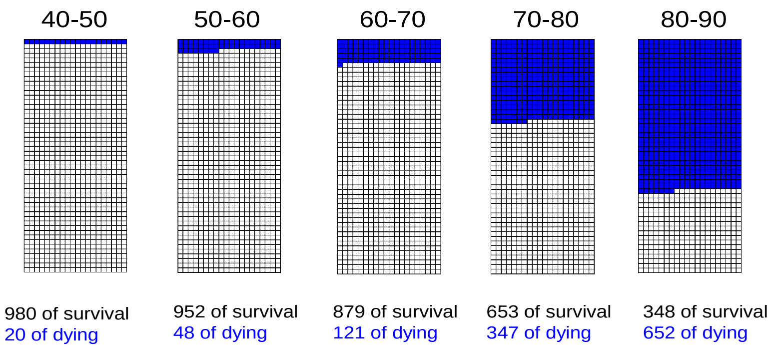
The experiment consisted of two different sections; the first section was a learning experiment, incentivised by the opportunity of winning a monetary prize (see Appendix C), whereas the second section involved hypothetical choices aimed at identifying preferences over different ways of generating a given LE gain. In addition to holding the LE gain constant across the three programmes, in order to ensure that respondents focused only on the time pattern of hazard rate reductions, no reference was made to the physical or medical process by which the reductions were brought about, so that the experiment was entirely "context-free". The protocol was developed, tested (n=46) and refined as an iterative process in consecutive, small groups prior to implementation in order to help improve, as much as possible, respondents' understanding of the various questions and tasks in the main study. In refining the protocol the two main challenges that arose were devising an acceptable procedure to communicate and elicit "indifference" from members of the public and identifying the minimum possible information that could be presented on the question and response sheets without compromising completely the scientific validity of the proposed programme changes. Thus, at a minimum, respondents required a reasonable degree of understanding and awareness of the initial risk they faced in the absence of a policy change, and the change in that initial risk distribution that would occur if their chosen Programme was implemented.

As mentioned above, the first part of the experiment created an incentivised learning environment, within which choices over different probability distributions of financial payoff were made. It allowed us to communicate to respondents the notion of conditional probabilities, letting them actually experience - in reality - the consequences of choosing to play one game containing a particular set of probability distributions as opposed to another. This allowed us to demonstrate that, amongst other things, games with different probability distributions can have the same expected pay-off and that expected payoffs are not the same as actual payoffs. This mirrors the process whereby different changes in the survival curve can lead to an identical gain in life expectancy. Therefore, when subjects arrived at the stage where they were to make choices between different life expectancy programmes they were already familiar with the notion of a remaining time path of different (conditional) risk distributions according to the type of programme implemented. In addition, the concept of an "indifference point" between two probability distributions was introduced in the money experiment and hence when the respondents arrived at the life expectancy section, the idea of finding an indifference point between two alternative options was a familiar task. Overhead slides, explanations and response sheet in the LE experiment mirrored as far as possible the material used in the money experiment.

In summary, the incentivised money experiment was introduced to resolve some of the information processing and cognitive problems associated with preference elicitation over LE gains, and hence to ameliorate one of the potential causes of any hypothetical bias that might arise. In somewhat more formal terms, the aim of the monetary experiment was to take advantage of the "rationality spillover" phenomenon introduced in Cherry et al. (2003). Rationality spillover refers to the idea that inducing rationality by a market-like discipline (monetary setting) can spill over to a non-market setting (life expectancy).

In the LE experiment, the respondents were asked to choose between the three different programmes; OR, AR, PR in pairs (named X, Y, Z in the experimental application). The respondents had already been introduced to versions of the underlying probability distributions in the money game in which they were labelled A, B, C. They were told that all three programmes would cost more or less the same and deliver a gain in LE equal to six months. To introduce the respondents to the three different programmes, we followed the risk communication approach described in Krupnick et al. (2002). This is illustrated in Figure 4, which shows the respondents' hazard rates over the next five decades. A blue square represented death, and a white survival, in each decade.

Figure 4. Initial risks. Overhead slide presented to the respondents



Essentially the same risk communication approach was used in the money experiment which was designed to mirror the LE experiment and was conducted prior to it in order familiarise respondents with the basic risk concepts. More specifically, it was explained to respondents that in the first round of the money experiment they would be allowed to make one draw from a bag containing 1000 counters, 20 of which were green and the remaining 980 white. If the respondent drew a white counter in this first round then a) he/she would be allotted one token which would be entered for a draw offering the possibility of winning a prize[5] at the end of the experiment and b) he/she would be allowed to proceed to the second round. If, on the other hand, the counter drawn was green the respondent would receive no token and would not be allowed to proceed to the next round. At the second round of the money experiment respondents were allowed to draw from a bag which again contained 1000 counters, but this time comprised 48 greens and 952 whites. If a white counter was drawn, then the respondent was allocated a second token for the money prize draw and was allowed to proceed to the third round, whereas if the counter drawn was green then no token was allocated and the respondent was not allowed to proceed any further. At the third round the counter split was 121 green and 879 white; at the fourth round, 347 green and 653 white and at the fifth and final round 652 green and 348 white. Clearly, therefore, the probability of receiving a "desirable white counter draw" at each round, conditional on having "survived" until that round, was set to be precisely the same as the conditional probability of surviving each of the five decades that respondents would face in the LE experiment. In addition, it was pointed out to respondents that drawing a white counter at any particular round could be regarded as providing two beneficial effects that parallel the "safety" and "survivor" effects of the LE experiment, namely a) gaining a token for the money draw and b) being allowed to proceed to the next round. The importance of these two effects was stressed since they are distributed differently in the three LE experiment programmes, and the separation of the two effects was intended to clarify to respondents why the total reduction in hazard rates is not identical in the three programmes, even if the LE gain is the same. An extract of the experimental protocol is included in Appendix C, exemplifying the explanations given to respondents.

After the respondents had indicated their preferences between a pair of programmes or indicated that the policy makers could choose (thus displaying indifference), they were asked to explain their answers by responding to an open-ended question to help assess the validity of the data. Subsequently, if they had not already done so, the respondents were asked to find their indifference point. The procedure used for finding the indifference point utilised a series of pairwise comparisons of cards. One card showed the risk reductions offered under their least preferred Programme, and remained constant across the choices. The other cards showed smaller risk reductions than were initially offered under their more preferred programme (there were 4

[5] The prize was in the range £20-50 and was adjusted on a per-session basis so that the expected pay-off was equal across different sized groups. The expected pay-off from participating in the experiment was£27 (£ 20 show-up fee and an expected pay-off from the draw of £ 7).

alternatives for AR and PR, but only 3in the case of OR[6]). In other words, the risk reduction in the respondent's preferred programme was decreased until the respondent was indifferent between the two choices.[7] So, for example, if OR (14/1000, 0/1000, 0/1000, 0/1000, 0/1000, 0/1000) was preferred to AR, then in the next round OR would be changed to (8/1000, 0/1000, 0/1000, 0/1000, 0/1000, 0/1000) which would then be compared to (an unchanged)AR. Respondents were then asked to identify the card that made it hardest for them to choose[8]. Respondents were also provided with templates on which programme cards could be sorted and they appear to have found this useful. In addition, the procedure for finding the indifference point was already familiar to the respondents since they had been taken through a similar process in the money experiment, in which indifference was illustrated by a "toss of a coin". All respondents, including intransitive respondents (who were only identified later in the analysis stage), succeeded in finding their indifference point under this procedure.

4. Findings of the study

Experimental sessions which each involved 4-10 participants were carried out during March-April 2008 in Newcastle-upon-Tyne. All participants were recruited to be in the "broad" 40-year-old age span. In the terminology of Harrison & List (2004), the experiment is therefore an "artefactual field experiment", since the population survey is a non-standard pool (i.e. not students). Appendix D describes the sample according to socio-demographic characteristics.

All respondents were instructed to answer on an individual level and no open-ended discussions were introduced during these sessions. A session lasted between 1-1½ hours; all experimental sessions were moderated by one of two members of the research team, aided by a trained assistant or assistants depending on numbers of participants. In total, 130 individuals participated; however 11 (8.5%) individuals displayed intransitive preferences[9] in the LE survey and were excluded from the sample, as was one respondent above the age of 50. This gives a final sample size of 118.[10]

---

[6] The reason for offering them only three different versions of the OR strategy is that the unchanged AR programme would strictly dominate the fourth OR option.

[7] By decreasing the risk reduction in their preferred option, a natural boundary is established (0 gain in LE) for the respondent's choice of indifference.

[8] Based on experiences from the piloting phases, all cards were coloured differently instead of given a number to indicate their ranking. The cards in the final survey were coloured in a randomized way; in this way the heuristic of choosing the same number (rank) in each indifference choice is avoided.

[9] Since a very high proportion of the respondents with intransitive preferences indicated indifference in the last questions, this could indicate that the intransitivity was caused by fatigue. It appears that the intransitivity here may be caused by seeking to avoid the relatively cognitively demanding indifference question as opposed to a particular

In Table 3 below, the simple descriptive statistics are listed, i.e. the respondents' choices in each of the three comparison questions in the LE section (by the number of respondents). The number in brackets represents a count of the respondents who made similar choices in the money experiment and the LE experiment. For example, two individuals were indifferent between the two programmes in the first comparison question in both the money and the LE experiment. In Appendix E the results of a direct comparison of the choices made in the money and the LE sections are shown. With a few exceptions the results indicate significant differences between the money and LE choices in almost all cases. Therefore, we infer that the choices are not merely a reflection of automatic repetition in that the *LE* responses were not unduly anchored or dependent on responses in the monetary experiment. This accords with the results from the earlier piloting phase of the experiment.

Table 3. Results of the three different comparison questions in the experiment

	Life expectancy
Comparison no 1 (OR vs. PR)	
OR	43(13)
PR	63(45)
Indifferent	12(2)
Comparison no 2 (AR vs. PR)	
AR	65(29)
PR	43(27)
Indifferent	10(1)
Comparison no 3 (OR vs. AR)	
OR	37(11)
AR	67(50)
Indifferent	14(2)

If the results in Table 3 were merely a consequence of random choices we would expect that within each comparison the respondents would divide more or less equally. A test for equality of proportions has therefore been carried out. For each comparison (No. 1, 2, 3) we can reject ($\alpha = 1\%$)

decision strategy not supported by rational choice. Thus, these individuals have not been given further attention in this paper.

[10] 13 individuals showed intransitive preferences in the finance game. However, since the finance experiment is mainly seen as a learning exercise these individuals have been retained in the sample.

that the proportions are identical and, for example, significantly more respondents have chosen PR than OR in the first comparison.

Table 4 combines the individuals' answers to the three comparison LE questions into a preference ordering of the three programmes.

Table 4. Preference orderings. The number of respondents with each preference ordering

Preference ordering	**Number of respondents**
PR>AR>OR	30
OR>AR>PR	29
AR>PR>OR	22
AR>OR>PR	11
AR=OR=PR	8
PR>OR>AR	5
Other	13
Total	**118**

”>”, ”<” indicates a strict preference whereas ”=” designates indifference

It is evident from Table 4 that in their pair wise comparisons the majority (70%) of respondents have indicated a preference ordering of the three LE programmes in one of three main ways: PR>AR>OR, OR>AR>PR and AR>PR>OR. The preferences for these three main orderings of the LE programmes are distributed more or less evenly across the subject pool. It is also clear from Table 4 that only a small proportion of the individuals were indifferent between all three programmes[11].

Regarding strengths of preferences, Table 5 relates only to respondents displaying a strict preference and contains their elicited indifference points in the LE survey; i.e. the number of months of gain in life expectancy the respondents were on average, willing to “sacrifice” from their preferred strategy in order to reach indifference between the two programmes in question.

[11] These results come with the caveat that order effects may have some influence, since practical limitations relating to non-computerised administration limited our flexibility. Results elsewhere suggest that the responses are valid and plausible, and sufficiently robust for the purposes they are used for in this paper.

Table 5. Indifference points (LE)

LE (in months)	Number of respondents	Mean willingness to give up (in months)	t-test of equality (p-value)
Comparison no 1 (OR vs. PR)			
OR	43	1.7	2.3(0.025)**
PR	63	2.2	
Comparison no 2 (AR vs. PR)			
AR	65	1.7	2.9(0.004)***
PR	43	2.3	
Comparison no 3 (OR vs. AR)			
OR	37	1.5	2.5(0.01)**
AR	65	2.0	

*,**,*** Significant at 0.1, 0.05 and 0.01 levels, respectively

To illustrate, in the first comparison question, on average, individuals with a preference for OR were willing to "give up" 1.7 months of life expectancy under OR in order to reach indifference between PR and OR, i.e. on average, individuals were indifferent between a 6-month gain in LE from PR and a 4.3 month gain in LE from OR.

Significant differences are apparent in a comparison of the indifference points within the three comparison questions (based on a t-test). Hence, individuals with a marked preference for the PR programme were on average willing to give up a greater gain in LE (i.e. accept a higher risk increase) to arrive at their indifference point, whereas participants with preferences for OR were the least willing-to-trade LE gains[12]. This on average suggests stronger preferences for programme PR.

4.1. Regression results

Two separate regressions have been carried out. Firstly, based on the responses to the three choices, the subset (n=109) of respondents with a strict preference for either OR (n=32), AR (n=36) or PR (n=41) have been grouped into three response categories *OR*, *AR*, and *PR* (thereby leaving out the respondents who displayed indifference in all three choices (n=8) and one individual who was indifferent between the two most preferred programmes). Since the response categories [P(α=j|**s**), in

[12] This could be an artefact of the fact that individuals choosing OR only received three cards (compared with the four in the other cases). However, there is also a significant different between AR and PR.

which j= $\widetilde{OR}$, $\widetilde{AR}$, $\widetilde{PR}$ are unordered, a multinomial logit regression is carried out to identify the extent to which changes in socio-demographic variables affect the choice of these categories, where **s** is a vector of socio demographic variables which are reported in Table 6 below. All coefficients are relative to the base category $\widetilde{PR}$.

Table 6. Multinomial logistic regression on three response categories

	$\widetilde{OR}$	$\widetilde{AR}$
	Coefficient. (std. errors)	Coefficient. (std. errors)
Male	-0.14 (0.49)	0.22 (0.50)
Below age 40	0.86 (0.68)	0.75 (0.69)
Children	1.39 (0.74)*	1.14 (0.73)
Individual income	-1.8e-04 (2.1e-04)	-6.4e-04 (2.8e-04)**
Higher education	0.47 (0.57)	0.76 (0.57)
Constant	-1.4 (0.89)	-0.68 (0.88)
N	109	109
Base category	$\widetilde{PR}$	$\widetilde{PR}$

*,**,*** Significant at 0.1, 0.05 and 0.01 levels, respectively

In addition to the variables above, variables indicating whether or not the respondent has taken out health insurance and for respondent's self reported health have been included but turned out to be insignificant. The results indicate that some socio demographic variables have a significant influence on the choice of type. As might be expected, respondents with children, and therefore with significant and immediate responsibilities for dependants, have a higher propensity to choose the more risk averse choice, $\widetilde{OR}$. On the other hand, respondents with a higher individual income have a lower probability of choosing AR compared to the more risky PR category, perhaps reflecting the fact that they have greater flexibility and control over their immediate future.

Secondly, an ordinary least square regression (Table 7) was carried out in which the dependent variable is the number of months the respondent was willing-to-trade (deltaLE). The data set contains three observations (one for each of the three pair wise choices) per respondent and the estimated standard errors are robust and clustered by individual respondent. The model is described below.

$$\text{DeltaLE}_i = \beta_1 D_{(OR\_PR)_i} + \beta_2 D_{(AR\_PR)_i} + \beta_3 D_{(OR\_AR)_i} + \beta_4 D_{(OR\_PR)_i} D_{(posvar)_i} + \beta_5 D_{(AR\_PR)_i} D_{(posvar)_i} + \beta_6 D_{(OR\_AR)_i} D_{(posvar)_i} + \varepsilon_i$$

All D's are dummy variables and $D_{OR\_PR} D_{posvar}$ etc. are interaction variables. Hence for example, $D_{(OR\_PR)_i} D_{(posvar)_i}$ represents the extra willingness-to-trade if the respondent has a preference for a positive variance in the choice between OR and PR. Hence, in Table 7, the coefficient on $D_{(OR\_PR)_i} D_{(posvar)_i}$ (β_4=0.42) indicates that the respondent will give up 0.42 months more if the respondent preferred PR to OR in the choice between OR and PR. On the other hand, the coefficient on $D_{(OR\_PR)_i}$ (β_1=1.49) tells us that a respondent with a preference for OR on average is willing to give up 1.49 months in LE. Accordingly, these results are equivalent to the results in Table5.

All interaction variables are significant at a 5% level indicating that within each choice there is significantly higher propensity to trade for respondents with preference for the option with the highest variance. Noticeably, though, as depicted in Figure 3, the largest difference in variance is in the comparison between OR and PR, followed by the choice between AR and PR and the choice between AR and OR. However, carrying out a separate regression (not reported here) it cannot be rejected that $\beta_4 = \beta_5 = \beta_6$ and hence the willingness to trade is not significantly influenced by whether for example OR is compared with AR or PR. Socio demographic variables have been included both as separate regressors and as interactions with the dummy variables. In all cases, they turn out not to be significant.

Table 7. Ordinary least square regression. Standard errors adjusted for 118 clusters

DeltaLE	Coefficient. (Robust std. errors)
$D_{(OR\_PR_i)}$	1.49 (0.13)***
$D_{(AR\_PR)_i}$	1.68 (0.13)***
$D_{(OR\_AR)_i}$	1.67 (0.13)***
$D_{(OR\_PR)_i} D_{(posvar)_i}$	0.42 (0.19)**
$D_{(AR\_PR)_i} D_{(posvar)_i}$	0.67 (0.23)***
$D_{(OR\_AR)_i} D_{(posvar)_i}$	0.47 (0.21)**
N	354
Number of clusters	118
R^2	0.75

*,**,*** Significant at 0.1, 0.05 and 0.01 levels, respectively

Overall, our data set suggests that socio demographic characteristics influence the individual's preference for either OR, AR or PR. However, the strength of preferences, i.e. the willingness-to-trade (deltaLE), appears to be a function of the initial choice rather than the underlying socio demographics. This may well be a consequence of the limited number of fixed choice sets that were offered to respondents, meaning that once they had chosen their initial 'type' (OR, AR or PR), only then were they only given the opportunity to trade life expectancy gains

5. Validity

As already noted, it is apparent that on an individual level the participants displayed different preferences in the money and LE experiments. These differences are taken as an indication of preference validity; since we infer from this that on average the respondents have taken the change of decision context into consideration and have not automatically responded in the same way in the two experiments[13].

In Table 4 we see that preferences for the three main orderings of the three LE programmes are distributed more or less evenly across the subject pool, which naturally could raise concern that the choices were purely random. However, there are two reasons for believing that this was not the case. In the first place, if the responses had been based purely on random choice we would not have seen the significant differences listed in Table 5. Also, only a small proportion of the subject pool displays intransitive preferences, whereas if the choices had resulted from purely random choice we would have expected to see more individuals displaying intransitive preferences. Results in Table 7 suggest that willingness to trade was correlated with the variance of the distribution in the direction predicted by theory.

Generally speaking, responses to open-ended questions following a choice experiment, whilst not a strong validity check in their own right, can nevertheless be viewed as being broadly indicative of respondents' understanding of the choices that they faced and the degree of careful thought they might have applied in the exercise, to a reasonable extent. Our conclusion is that in this study the answers to the open-ended questions add further support to the conclusion that responses in the choice experiment were not randomly generated. Some examples of the responses to the open-ended questions are as follows:

RESPONDENT 30; "I'd prefer to increase my chances of reaching 70" *(preference for OR over PR, female 43)*

RESPONDENT 17; Take chance of surviving 40ies, Y increases chances as I age *(preference for AR over OR, male aged 45)*

---

[13] Also, when asked to find their indifference point, only 9 (8%) individuals chose a similarly coloured card each time. Hence deciding on a colour and then choosing this colour automatically every time a choice had to be made did not appear to be a commonly applied heuristic in this instance either.

RESPONDENT 32; “because chance of dying before you're 50 not as high so may as well lower your chances in your early retirement so that you can recap the benefits of your working life” *(preference for PR over OR, female, aged 45)*

RESPONDENT 72; “Having children I feel they should still need me in my 40's therefore if I can increase life expectancy at this stage then I would feel like I had achieved something*” (preference for OR over PR, female, aged 35).*

Overall, a large proportion of the responses indicated that they had paid significant attention to risk concepts and comprehended the implications in terms of actually receiving any LE gain and acknowledged the trade-offs inherent in their choices. More than 70% of the participants based their arguments around terms like “changes in risks/chances/odds” and many referred to differing opportunities and responsibilities at differing points of life. The general impression resulting from answers to the open-ended questions in that respondents were, on the whole, able to articulate evidence of clear and coherent reasoning in line with their quantitative choices.

6. Discussion of findings

The main result from this experiment is that most people display a marked preference for one way of generating a particular gain in LE rather than another. It is also evident from Table 5 that the majority of respondents have ranked the LE programmes in one of three ways: PR>AR>OR, OR>AR>PR or AR>PR>OR. The preferences for these three main orderings of the LE programmes are distributed more or less evenly across this subject pool. Hence, if these results are replicable over a larger sample and other age groups, the implications from a policy perspective are of considerable importance. In particular, it would appear that a) the tendency to date of both theorists and policymakers to ignore the nature of the perturbation in the survival function that gives rise to a gain in life-expectancy constitutes a serious oversight and b) that in spite of this, there does not seem to be a clear majority preference for one way of generating a given gain in life expectancy as opposed to another.

A further important policy implication arises for benefit-cost analysis. Based on these results we could envisage that even though an individual might place little or no value on a risk reduction of the programme OR type, they could nonetheless be willing to pay a significant amount for a risk reduction of the programme of the AR or PR type. One possible implication of this is that using an existing Value of Statistical Life (VSL) estimate based on current-period risk reduction as the starting point for the calculation of a VOLY could potentially fail to take into account the fact that some people might be willing to pay for a risk reduction arising from an AR or PR type programme, but not for a risk reduction arising from an OR type of programme, leading to an underestimation of the benefits from programmes with the characteristics of AR and PR. Similarly, VSL estimates derived from data involving ongoing variations in risk, as in the case of compensating wage differential studies – see, for example, Viscusi (1978) – might not provide an appropriate basis for deriving VOLY estimates for programmes of the OR type. Indeed, differences in the time profile of variations in risk may go some way towards explaining the variations in the implicit VSL derived from different contexts, including labor markets, automobile safety devices, bicycle helmets, domestic fire detectors and air pollution reported in Viscusi and Aldy (2003). In short, the result emphasises the importance of providing a clear specification of the distribution and timing of the risk reduction when valuing life expectancy gains or estimating a VSL.

Respondents chose between programmes that had different distributions of risk. According to theory, their preferences are clearly informed by a mix of individual risk aversion (or proneness) and time preferences and also the baseline level of risk. Unfortunately, our restricted set of programmes does not allow us to identify the magnitude of all of these effects with great precision, but we do find that the least risk-averse respondents are more willing-to-trade off gains in LE which in accordance with our expectation. However, in addition to the influence of risk aversion, our results could also be a straightforward reflection of the effect of discounting or that respondents applied some other form of subjective weighting in their underlying utility function. For example, evidence from the piloting and from the qualitative answers given by respondents after they had chosen between the programmes suggests that for some respondents the baseline risk in the first period was so small that they discounted it to zero, whereas others were uninterested in gains in the final period. To illustrate, two respondent quotes are included below:

RESPONDENT 15 "To keep reducing my chance of death as I get older, I am more likely to survive from 40-49 anyway so I would prefer additional chances later in life" *(preference for AR over OR, female respondent, age 34).*

RESPONDENT 37: "Better to survive in 40-49 than 80-89 so give more chance with ORX" *(preference for OR over PR, male respondent, age 45)*

These respondent quotes suggest that a subjective weighting function might have been used. For example, one such function could follow the simple but albeit extreme assumption, of setting a combination of the first and last period gains from risk reductions to equal zero. As can be seen below this would generate the following preference orderings:

First period gain set to Zero : PR>AR>OR,

Last period gain set to Zero : OR>AR>PR,

First and Last period set to Zero : AR>OR>PR.

Assuming that some of the respondents took some account of the riskiness (or variance) of the choices then the preference ordering; AR>PR>OR becomes part of the choice set as well.

Comparing these orderings to those generated by the respondents shows that these weighting functions can help explain almost all of the preference orderings made by respondents. Clearly this is an area for further research.

Of course, for the purposes of our particular experiment, our results are based strictly on the trade-offs between probability distributions and do not take potential preferences for different contexts (e.g. air pollution or traffic) into account. While Chilton et al. (2006) find little evidence of any contextual effects across different types of instantaneous, premature deaths (i.e. all within a VPF framework and excluding causes that involve protracted periods of suffering prior to death, such as cancer), it may be that context and probability distributions interact in some way to influence preferences in the other types of programme.

The present experiment also addresses three key practical problems associated with the actual application of the risk-risk trade-off method relating to changes in mortality risks. The first is peoples' well-documented difficulty in dealing with changes in small probabilities (see, for example Corso et al., 2001), which one might reasonably expect to be even more pronounced when the focus is on probability distribution. The second, less well-documented (but often observed in pilot studies) problem is the apparent inability that people have to discriminate finely over different periods of time. This often manifests itself as an unwillingness to trade any gain in LE whatsoever or, in the case of a willingness-to-pay (WTP) study, offering zero WTP for often quite substantial LE gains. Our study appears to have ameliorated this, at least to a reasonable degree (see the discussion surrounding Table 5). The third is that the task must, of necessity, be hypothetical in nature. We show that applying an incentivised experiment to help respondents understand the key features of the "delivery mechanism" of the LE gain generates consistent and credible quantitative choices in the subsequent hypothetical survey and more specifically only a small proportion of the sample displays intransitive preferences.

Within the constraints of this study, we were unable to consider other age groups or to analyse other probability distributions. These are clearly areas for further research in the future, as is disentangling the consequences of differences in time and risk preferences from initial risk effects. Parallel developments and findings in the finance literature aiming to isolate and analyse preferences for skewness from those of variance and expectation (Alderfer & Bierman 1970; Barberis & Huang 2008) should be brought into the study of physical risk changes.

7. Conclusion

While the findings of our study indicate that people appear to have clear preferences over different ways of generating a given gain in life expectancy, preferences for three main types of LE programme are found to be distributed more or less evenly across the subject pool of 40-year-old individuals. All programmes generate a six-month gain in LE and if replicable over a larger sample and other age groups, the implication of these results from a policy perspective is that a mix of different LE generating policies is warranted, i.e. both programmes generating a one-period change in risk and programmes delivering an ongoing and sustained reduction in risk over the remainder of their lives. We further highlight the potential problems that may arise as a result of basing estimates of the Value of a Life Year (VOLY) on a predetermined Value of Statistical Life (VSL), given that the latter will inevitably have been derived from data relating to a quite specific form of risk reduction and may therefore not constitute the appropriate basis for computing the value of a gain in life expectancy generated by some other form of perturbation in the survival function.

From a methodological perspective, we would argue that the experimental procedure employed in the study represents a significant advance in helping respondents to acquire a better understanding of how gains in life expectancy are actually delivered (and what is meant by this concept), at least intuitively. Whilst we used paper-and-pen administration in this study, the procedures would be easily transferable to computer. Indeed, doing so would open up the possibility of exploring the impact on LE preferences of the many other different types of perturbations of the hazard function.

Finally, on a more general note, our findings demonstrate the potential for the development of methods that precede hypothetical choice tasks with incentivized tasks to encourage more reliable responses.

Acknowledgment

This work was supported by the Health Insurance Foundation in Denmark (Helsefonden) and the Danish Centre of Excellence named AIRPOLIFE (Air Pollution in a Life Time Health Perspective).

References

Alderfer, C. P. & Bierman, H. 1970, "Choices with risk: beyond the mean and variance", *The Journal of Business*, vol. 43, no. 4, pp. 341-353.

Aldy J.E, & Viscusi, W. K. 2008 Adjusting the Value of a Statistical Life for Age and Cohort Effects. *The Review of Economics and Statistics*, vol. 90, no. 3. pp. 573-581.

Barberis, N. & Huang, M. 2008, "Stocks as Lotteries: The Implication of Probability Weighting for Security Prices", *American Economic Review*, vol. 98((5), pp. 2066-2100

Berry, D. A. & Lindgren, D. A. 1996, *Statistics: Theory and Methods*, Second edn, Duxbury Press, Belmont California, USA.

Cherry, T.L., Crocker T.D., Shogren J.F. 2003, "Rationality spillovers", *Journal of Environmental Economics and Management*, vol. 45, pp. 63-84

Chilton, S., Jones-Lee, M. W., Kiraly, F., Metcalf, H., & Pang, W. 2006, "Dread risks", *Journal of Risk and Uncertainty*, vol. 33, pp. 165-182.

Chilton, S., Covey, J., Jones-Lee, M. W., Loomes, G., & Metcalf, H. 2004, *Valuation of health benefits associated with reductions in air pollution*. Final report. DEFRA

Corso, P. S., Hammit, J. K., & Graham, J. D. 2001, "Valuing Mortality-Risk reduction: Using Visual Aids to Improve the Validity of Contingent Valuation.", *Journal of Risk and Uncertainty*, vol. 23, no. 2, pp. 165-184.

Cropper, M.L., Aydede, S.K.,& Portney, P.R. 1994, "Preferences for Life Saving Programs: How the Public Discount Time and Age", Journal of Risk and Uncertainty, vol. 8, pp. 243-265.

Hammitt, J. K. 2007, "Valuing Changes in Mortality Risk: Lives Saved Versus Life Years Saved", *Review of Environmental Economics and Policy*, vol. 1, no. 2, pp. 228-240.

Harrison, G. W. & List, J. A. 2004, "Field experiments", *Journal of Economic Literature*, vol. 42, no. 4, pp. 1013-1059.

Jenkins, S. P. 2005, *Survival Analysis*, iser.essex.ac.uk

Johannesson, M. & Johansson, P. O. 1996, "To Be, or Not to Be, that is the Question: An Empirical Study of the WTP for an increased Life Expectancy at an advanced Age.", *Journal of Risk and Uncertainty*, vol. 13, pp. 163-174.

Jones-Lee, M. W. 1976, "The Value of Changes in the Probability of Death or Injury", *The Journal of Political Economy*, vol. 82, no. 4, pp. 835-849.

Jones-Lee MW, Hammerton M, Phillips P R. 1985. "The Value of Safety: Results of a National Sample Survey", The Economic Journal, vol. 95 no. 377 pp. 49-72.

Mason, Helen, Michael Jones-Lee, and Cam Donaldson 2009, "Modelling the Monetary Value of a QALY: A New Approach based on UK Data". *Health Economics,* Vol. 18*,* pp.933-950

McDowell, I & Newell, C 1996. *Measuring health: A guide to rating scales and questionnaires. (2nd ed.*) New York: Oxford U Pr.

Moore, M.J. & Viscusi, W.K.1988. The Quantity-Adjusted Value of Life, Economic Inquiry, vol. 26, pp. 369-388.

Krupnick, A., Alberini, A., Cropper, M., Simon, N., O'Brien, B., Goeree, R., & Heintzelman, M. 2002, "Age, Health and the Willingness to Pay for Mortality Risk Reductions: A Contingent Valuation Survey of Ontario Residents." *Journal of Risk and Uncertainty*, vol. 24, no. 2, pp. 161-186.

Morris, J. & Hammitt, J. K. 2001, "Using Life Expectancy to Communicate Benefits of Health Care Programs in Contingent Valuation Studies", *Medical Decision Making*, vol. 21, no. 6, pp. 468-478.

Pope, C. A., Thun, M. J., Namboodiri, M. M., Dockery, D. W., Evans, J. S., Speizer, F. E., & Health, C. W. 1995, "Particulate air Pollution as a Predictor of mortality in a Prospective Study of U.S. Adults", *Am J Respir Crit Care Med*, vol. 151, pp. 669-674.

Sunstein, C. R. 2004, "Lives, life-years, and willingness to pay", *Columbia Law Review*, vol. 104, no. 1, pp. 205-252.

Viscusi, W. K. 1978, “Labor Market Valuations of Life and Limb: Empirical evidence and policy implications. *Public Policy,* vol. 26, pp. 359-386.

Viscusi, W. K., Magat, W. A., & Huber, J. 1991, "Pricing environmental health risk: survey assessments of risk-risk and risk-dollar trade-offs for chronic bronchitis.", *Journal of Environmental Economics and Management*, vol. 21, pp. 35-51.

Viscusi, W.K. Aldy, J.E. 2003 “The Value of a Statistical Life: A Critical Review of Market Estimates Throughout the World”, Journal of Risk and Uncertainty, vol. 27 no. 6, pp. 5-76

Appendix A

Assumption required for the calculations of life expectancy gains

Data stems from Government Actuary's Department (mean 2003-2005) and reflects a simple average between female and male. The simulations are carried out from the age of 40 until the age of 90.

The hazard rates are calculated the following way: $\frac{l_{50} - l_{40}}{l_{40}} = p_{40}$

l_x is the number of survivors to exact age x of 100.000 live births who are assumed to be subject throughout their lives to the mortality rates experiences in the three year period to which the data relates.

Accordingly, in the money game, $1000p_{40}$ is the number of green cards in the first bag and the number of white cards is $1000(1\text{-}p_{40})$. In the money game this is the appropriate starting point for an estimation of E(tokens) since you either 'survive' a whole bag or 'die". However, the LE –section should technically be continuous since there is both a probability you could die in the beginning of the decade or at the end of the decade. By assuming that you need to survive to the end of the decade this means that we have overestimated the probability of dying slightly and hence underestimated the current life expectancy.

Formally, it would have been more accurate to apply the following approximation instead of equation 6:

$LE_{40}= 0.5*10((1\text{-}p_{40}) + (1\text{-}p_{40})(1\text{-}p_{50}) + (1\text{-}p_{40})(1\text{-}p_{50})(1\text{-}p_{60})) + 0.5*10(1+ (1\text{-}p_{40}) + (1\text{-}p_{40})(1\text{-}p_{50}) + (1\text{-}p_{40})(1\text{-}p_{50})(1\text{-}p_{60}))$

This would approximate the continuous curve more accurately. However, since the main purpose of this experiment is to show that a method works, it has been decided to mirror the money game in the LE section. Also, since we wish to examine whether preferences shift across contexts, we must keep all other things constant.

Appendix B

Money experiment

The experiment applied was built around four different incentivised games – three option games and one indifference game – the purpose of which are described below. An additional baseline game was included which served as a practice game. Each game potentially contained five rounds and each round of a game consisted of a bag containing a total of 1,000 cards (a mix of green and white cards) where the specific mix reflected 'game specific' hazard rates for the next five decades. Each participant drew a card from a bag. The card was subsequently placed back into the bag before the next participant drew a card[14].

In the option games of the money experiment, participants were asked to make pair-wise comparisons of three games. All games (i.e. sets of bags) reflect a changed distribution of hazard rates according to the changes described in Table 1; however, in the money game they were labelled A, B and C instead of X, Y and Z. Following a choice of which 'game' (i.e. set of bags) to play, the participants either 'survived' a bag (drew a white card), thereby collecting a token and able to continue in the game; arrived at the end of the game; or were eliminated prematurely (drew a green card). Furthermore, it was explained to respondents that drawing a white card has two beneficial effects: 1) by 'surviving' that bag they have obtained a token to enter the draw, and 2) they are in a position to progress to the next bag(s) and hence have the chance of winning additional tokens. Changing the distribution of the cards changes both of these effects and the two effects mirror the safety and survivor effects introduced to the respondents in the subsequent LE experiment.

The tokens collected in the games were entered into a draw for a prize[15] which was adjusted on a per-session basis to make the expected pay-off across different group sizes equal. Each token was equivalent to one entry into the draw and the winner of the prize was determined by a random draw at the end of each session. A brief outline of the format for the money experiment is presented in Table B1 notice though the expected number of tokens in the baseline game is 3.46, which is

[14] Before the respondents were allowed to draw a card from a bag the bag was shaken and it was assured that the respondent only drew one card and that they could not see the colour of the card before it was drawn.
[15] The prize was in the range of GBP 20-50 depending on group size and the expected pay-off from participating in the experiment was GBP 27 (GBP 20 show-up fee + an expected pay-off from the draw equal to GBP 7).

equivalent to 3.46 decades in the LE section, the current LE for a 40-year-old British individual (truncated at the age of 90).

Table B1: Format for the money experiment

<table>
<tr><td>Game 1</td><td>Baseline game</td><td>• Practice game (played out for token)
• E(tokens)= 3.46</td></tr>
<tr><td></td><td>Option games</td><td>• Some green cards replaced with white cards
• E(tokens)=3.51 in each option</td></tr>
<tr><td>Game 2</td><td>A vs C</td><td rowspan="3">• Three pair-wise comparisons are made or a coin is tossed when the respondent is indifferent between the two sets of bags.
• One randomly decided choice is played out for tokens.
• The respondents play their preferred set of bags in the randomly chosen game.</td></tr>
<tr><td>Game 3</td><td>B vs C</td></tr>
<tr><td>Game 4</td><td>A vs B</td></tr>
<tr><td>Game 5</td><td>Indifference game</td><td>• Indifference point is established for the randomly chosen game the respondents just played.
• A coin is tossed to decide which set of bags to play out for tokens.</td></tr>
</table>

OR, AR, PR, in terms of the distribution of risk reduction, corresponded precisely to option A, B, C in the money games.

Appendix C

Extract from the experimental protocol

- Let us say that you were offered the opportunity to choose between three on-going programmes, X, Y and Z. All would cost more or less the same. All programmes would increase the average life expectancy by **6 months.** Although in each case this would be achieved differently. By this I mean that the risks of dying in each decade would be different in each programme**.**

HAND OUT HANDOUT 2.

Handout 2 shows you your current risk. Programme X,Y and Z would change these in different ways.

- **X VS Z**

i) We will now ask you to choose between programme X and programme Z. Let us now have a look at the two programs in details.

OVERHEAD

On the top of this overhead you see the risk of dying in each decade as they are before we make any changes. Underneath is programme X indicating the change in the risk of dying in each decade offered by programme X. In programme X we have reduced the risk of dying with 14/1000 in the first decade. This gives us a new risk of dying in the first decade of 6/1000. The risk of dying in the second decade is still 48/1000, 121/1000 in the third, 347/1000 in the forth and 652/1000 in the fifth.

Underneath is programme Z indicating the change in risk of dying in each decade offered by programme Z. In programme Z we have changed the risk of dying differently in each decade. We have reduced the risk in each decade, starting with 1/1000 in the first decade which gives a new risk of dying equal to 19/1000. We reduce the risk of dying with 2/1000 in the second period to 46/1000. In the third decade we reduce the risk of dying with 5/1000 to 116/1000. In the forth decade we reduce the risk by 15/1000 to 332/1000 and in the fifth decade with 28/1000 to 624/1000.

ii) The change in life expectancy is 6 months in each programme. Note that the change in life expectancy is the same in both programmes even though we don’t have the same total risk reduction i.e.14/1000 in X vs 51/1000 in Z.

This is because you get a lot of survivor effect from reducing the risk in decade 1, since you have many periods ahead of you. On the other hand; reducing the risk in the fifth decade doesn’t give you any survivor effect. In order to get the same change in life expectancy from X and Z you need a bigger safety effect in later decades in Z. This mirrors the situation in the finance game were we changed different number of green cards in the different options but ended up with the same total effect as in the expected number of tokens.

iii) Please take your time and think about what programme X and programme Z would mean to you. Please indicate whether you prefer X or Z in Q5 in the response sheet. If you would be equally happy with either i.e. you wouldn't mind letting the policy makers choose which programme to implement; please tick the 'equally happy' box.

After you have made your choice I will ask you to answer Q6 in your response sheet.

Appendix D
Sample demographics

Description	**Variable name**	**N**	**Mean**
Proportion of males	Male	130	0.43
Age (in years)	Age	130	41
Below 40	Below age 40(=1 if below the age of 40, else 0)	130	0.55
Education (1=primary, 2= secondary, 3= higher)	Higher education(=1 if higher education, else 0)	130	2.5
Health (5 categories. The first question in the Short Form 36 (SF-36) health survey[16])	Good (=1 if health status is the respondent's self-reported health is good or excellent)	130	3.8
Proportion with children	children	130	0.67
Proportion with a health insurance	insurance	130	0.28
Mean individual income per month (GBP) (8 response categories)	Individual income	129	1800
Mean household income per month (GBP) (8 response categories)	Household income	124	2850

[16] See e.g. **McDowell I & Newell C (1996).**

Appendix E

Results of the three different comparison questions in the experiment

	LE	Money	Signed rank test(p-value)
Comparison no 1 (X vs. Z)			
X	43(13)	26	-2.6(0.01)**
Z	63(45)	78	2.1(0.04)**
Indifferent	12(2)	14	0.42(0.67)
Comparison no 2 (Y vs. Z)			
Y	65(29)	47	-2.4(0.01)**
Z	43(27)	64	2.9(0.004)***
Indifferent	10(1)	7	-0.78(0.44)
Comparison no 3 (X vs. Y)			
X	37(11)	27	-1.5(0.12)
Y	67(50)	82	2.1(0.03)**
Indifferent	14(2)	9	-1.1(0.25)

*,**,*** Significant at 0.1, 0.05 and 0.01 levels, respectively