To be presented at Managing the Social Impacts of Change from a Risk Perspective, Beijing, China 15-17 April 2009

Work in progress – please do not quote

How do you prefer your gain in life expectancy to be delivered?

An experimental approach linking hypothetical and incentivised choices

Jytte Seested Nielsen*[1], Susan Chilton[2], Michael Jones-Lee[2], Hugh Metcalf[2]

[1]Department of Health Economics
J.B. Winsløwsvej 9, 1
DK-5000 Odense C
fax: 0045 6550 3880
jsn@sam.sdu.dk

[2]University of Newcastle
Ridley Building
Claremont Road
Newcastle NE1 7RU
United Kingdom
s.m.chilton@newcastle.ac.uk
michael.jones-lee@newcastle.ac.uk
h.r.t.metcalf@newcastle.ac.uk

* To whom correspondence should be addressed

Abstract

This paper demonstrates the potential for economic experiments to inform regulatory decision-making and/or policy formation in the domain of health, safety and the environment. In this paper we empirically investigate preferences for different ways of generating a particular gain in life expectancy. Use of a novel experimental methodology improves the reliability of the (necessarily) hypothetical responses by first allowing respondents to gain experience in making similar choices in an incentivized setting. We investigate preferences for three different ways of generating an individual gain in life expectancy of six months. Preferences for three main orderings of the life expectancy programmes are found to be distributed more or less evenly across the subject pool.

1. Introduction

As part of health and safety policy, national governments implement a range of programmes that impact people at different points in their lives by reducing their risk of dying, and hence increase their life expectancy (LE). For example, some road safety programmes generate a one-period change in risk; here, when the intervention is removed, peoples' risk of dying returns to what it would have been without the intervention (these programmes relate to the framework for 'Value of Preventing a Statistical Fatality' (VPF)[1], Jones-Lee (1974)). Other programmes deliver an ongoing and sustained reduction in risk over the remainder of their lives, such as might occur under a programme of air pollution reduction and from which a 'Value Of a Life Year lost' (VOLY) type framework might emerge.

Little, if anything, is known about society's preferences over such programmes, i.e. whether people prefer programmes of the former or latter type or some mix of the two. This paper aims to investigate preferences for three different (hypothetical) programmes which all deliver the same gain in life expectancy but where the risks are distributed differently. One hypothesis why distribution may matter is that even though two different programmes could deliver the exact same change in life expectancy, other moments of the probability density function, in particular the variance, could be expected to influence the valuation as well.

If the distribution and the variance matter, then perhaps a new communication approach will be required. During piloting for this experiment and previous work on the valuation of a gain in LE (see e.g. Chilton et el (2004)) it has been suggested that if the respondents are presented with a gain in LE but not informed about the underlying probability distributions, a strong possibility exists that they perceive the gain in LE as a 'certain add-on' at the end of their life. If this were to be the case, respondents' could not differentiate between two LE programmes which generate the same LE gain but in different ways. Additionally, if they were asked to choose between two programmes which offer different LE gains no risk trade-offs would be made and the preference for any programme delivering a larger gain would always strictly dominate, regardless of the probability distribution. This issue is addressed in the present experiment. Also, previous methods directly valuing a gain in

[1] Or as it is also known, the 'Value of Statistical Life' (VSL)

life expectancy have not always described the gain in life expectancy accurately and hence the respondents have not been informed about the manner in which the gain in life expectancy is to arise.

By means of a survival curve, this paper aims to address these questions and develops a methodology to allow us to 1) provide a conceptual framework to facilitate a direct comparison of LE benefits from VPF- and VOLY-types of programmes, and 2) develop an empirical mechanism by which to investigate individual preferences for the different types of programme. This mechanism makes use of a novel experimental methodology to improve the reliability of the (necessarily) hypothetical responses, by first allowing respondents to gain experience of making similar choices in an incentivised setting.

The remainder of the paper is as follows: the subsequent section describes the analytical idea and outlines the experimental design. The experimental methods are outlined in section 3, results are then presented, and in section 5 a discussion of results is made. The final section concludes.

2. A gain in life expectancy; a definition

All changes in mortality risks can be described as a change in the individual's survival curve. Let the actual survival time t of an individual be regarded as the value of a variable T which can take any non-negative value (T>t). The different values that T can take have a probability distribution and we call T the random variable associated with the survival time.

The hazard function, h(t) describes the probability of dying in a given period (from t to t+δt), conditional on having survived until period t, and the hazard function is related to the survival curve in the following way:

$$h(t) = \frac{f(t)}{S(t)} \ ,$$

in which f(t) is the probability density function for t. The survival curve, S(t) expresses the probability that the individual survives for at least t periods; S(0) = 1, $\lim_{t \to \infty} S(t) = 0$ and the area under the survival curve identifies LE:

$\int_0^{\infty} S(t)dt = LE$ (Collett 1994;Kamerud 1989). A particular life expectancy can be generated by various survival curves. Hence a specific gain in LE can be generated as a result of different perturbations in the hazard function, resulting in different changes in the survival curve.

2.1. Description of the three life expectancy programmes

In the experimental design we will examine three different programmes which change the hazard function in various 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 ways of generating a particular gain in LE.

The reasoning behind selection of the three programmes outlined below is in part pragmatic. While there are an infinite number of different possible perturbations in the probability density function for an individual's remaining length of life, it would not be feasible to investigate them all; although in principle the methodology we are going to introduce is generalizable. In any case, the following three programmes are however representative of programmes that may actually be implemented in practice and represent perturbations which from a mathematically point are very interesting. Programme X offers a one-period reduction in hazard rates (analogous to the VPF framework), whiles the alternatives, programmes Y and Z, deliver an ongoing and sustained reduction in hazard rates over the remainder of an individual's life. In programme Y the reduction in risk is absolute while in programme Z it is relative (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

X) an absolute risk reduction δ ($\delta \in$ (-p,0)) in the *first period* hazard rate after which the hazard rates return to their previous level
Y) a permanent constant *absolute* risk reduction ω ($\omega \in$ (-p,0)) in hazard rates
Z) a permanent constant *relative* risk reduction kp (k $\in$ (-1,0) in hazard rates (p)

The experimental 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. The advantages of restricting the design to only one age group are that the same hazard rates can be used for all respondents and results from a relatively small sample size is controlled for any effects of age. The methodology could in principle be adapted to other age groups. However, the main objective here is the methodological development.

The set-up is illustrated in EQ 1[2], in which p_i represents the existing hazard rates for a given decade:

$$LE_{40} = 10(1-p_{40}) + 10(1-p_{40})(1-p_{50}) + 10(1-p_{40})(1-p_{50})(1-p_{60}) + .. \quad \text{EQ 1}$$

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

In the experiment, changes in hazard rates from three different LE programmes are compared, all of which generate a six-month gain in LE. The three programmes are described formally in Figure 1, while EQ 2-4 express the LE calculated in discrete time for the three different programmes, and Table 1 lists the numbers presented to the respondents in the experiment.

$$LE_x = 10(1-(p_{40}+\delta)) + 10(1-(p_{40}+\delta))(1-p_{50}) + .., \delta \in (-p_{40}, 0) \quad \text{EQ 2}$$

$$LE_Y = 10(1-(p_{40}+\omega)) + 10(1-(p_{40}+\omega))(1-(p_{50}+\omega)) + .., \omega \in (-p_i, 0) \quad \text{EQ 3}$$

$$LE_Z = 10(1-p_{40}(1+k)) + 10(1-p_{40}(1+k))(1-p_{50}(1+k)) + .., k \in (-1, 0) \quad \text{EQ 4}$$

[2] Or, equivalently $LE_{40} = 0p_{40} + 10(1-p_{40})p_{50} + 20(1-p_{40})(1-p_{50})p_{60} + ..$

Table 1: The risk figures describing the three programmes

Age	**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
X *changes* in hazard rate (δ)	14/1000	0	0	0	0
Y *changes* in hazard rate (ω)	5/1000	5/1000	5/1000	5/1000	5/1000
Z *changes* in hazard rate (kp_i)	1/1000	2/1000	5/1000	15/1000	28/1000

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 2003-3005). 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 X, 25/1000 in Y and 51/1000 in Z) and EQ 2-4 illustrate why this is so. For example, examining EQ 2 it is clear that a change in hazard rate in the first decade influences the remaining conditional 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 conditional upon on this, the 'survivor effect'. Implementation of a LE programme delivers a direct change in hazard rate ('safety effect') whereas the 'survivor effect' arises as a consequence of the safety effect, conditional upon survival. Consequently, a reduction in hazard rate in the first decade brings more 'survivor effect' than an equivalent reduction in hazard rate ('safety effect') in the fifth decade.

The structural characteristics of the way a gain in LE is generated can be described in terms of a decision tree approach also mirroring the concept of option value in financial theory. Figure 2 below describes the initial life expectancy in this way, in a five period discrete-time model for a 40 years old average individual.

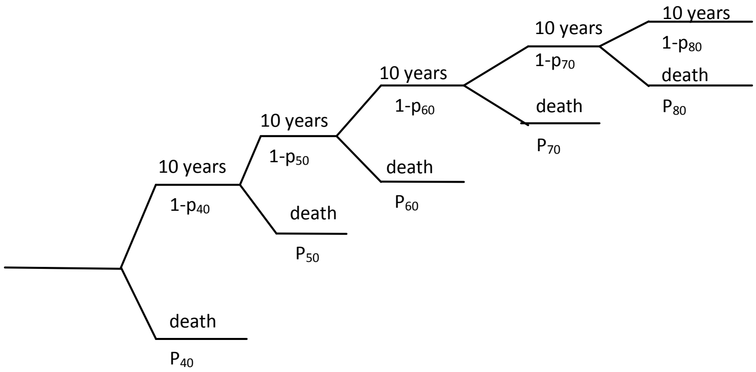


Figure 2.The initial LE described by use of a decision tree

The figure above hence describes the initial situation which from a game theoretic perspective can be described as 'nature' or the initial 'LE game' faced by the individuals.

In the empirical application we offer the respondents the three different programmes, which are illustrated in the decision tree setting in Appendix B, each of which offers a change in the probability of dying which is distributed differently. So the game now is, first respondents chose their preferred programme and then nature takes its course.

We will investigate three different programmes which each deliver the exact same gain in life expectancy, mathematically, i.e. the expectation is the same. However, we could expect other moments of the probability density function to influence the respondents' preferences. In this paper we will pay attention to the variance, i.e. the dispersion around the mean. All other things held equal we would expect risk averse individuals to have a preference for the programme with the smallest variance i.e. the programme in which you have the highest chance of getting exactly the mean.

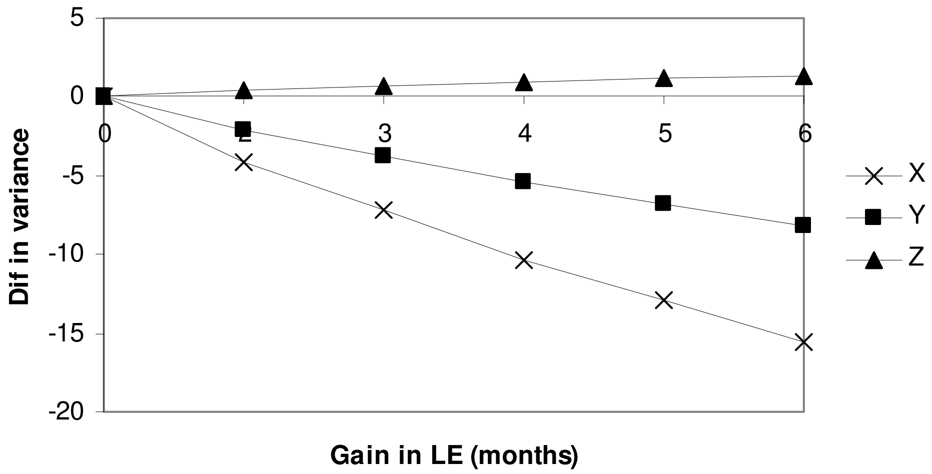
A calculation of the variance of the three programmes shows that programme X offers the largest reduction in variance followed by Y programme Z increases the variance. This is illustrated in Figure 3 below.

Figure 3 depicts the change in variance of LE as a function of the gain in LE[3]. For example, a six-month gain in LE from programme X results in a decrease in variance of about 15. Since this represents the largest decrease in variance, the variance of programme Z and Y strictly dominates the one-period change in the hazard rate (X) for a six-month gain in LE. This implies that X will be the choice of the most risk-averse. A different way of interpreting this is that programme X exhibits the largest change in variance according to LE gain (the steepest slope). Thus when an individual prefers X, a reduction in life expectancy gain from six to five months would have a larger absolute impact on the change in variance than an equivalent reduction in LE for an individual who preferred Y or Z to X.

[3] Variance is calculated applying the standard formulas (Berry & Lindgren 1996), adapted to

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

Figure 3: Change in variance, the applied data



3. Methods

3a. LE section

Trading off different LE gains requires that respondents are familiar with probabilities and are willing to express their strengths of preferences 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 eliciting respondents' indifference points between the three programmes in pairs.

The experiment consists of two different sections; the first section is a learning experiment, incentivised by the opportunity of winning a monetary prize, whereas the second section offers the opportunity of obtaining LE gains in a hypothetical set-up. The LE experiment takes place in a context-free setting which enables two of the three variables to be held constant (context and life expectancy gain) while isolating and varying the variable of interest (changes in) the probability distribution. The protocol was developed, tested (n=46) and refined prior to implementation in order to ensure in the main study that respondents would understand the various questions and tasks.

In the LE experiment, the respondents were asked to choose between the three different programmes; X, Y, Z in pairs. All three programmes would cost more or less the same and deliver a gain in LE equal to six months.[4] Attention was drawn to the fact that there are two different beneficial effects from surviving a decade. In the LE experiment, these are called the 'safety effect' and 'survivor effect'; explained, respectively, to the respondents in terms of 'you did not die' and 'you now have the possibility of future decades'. The importance of these two effects was stressed since they are distributed differently in the three programmes, and separation of the two effects intended to clarify to respondents why the total reduction in hazard rates is not identical in the three programmes despite that the gain in life expectancy remains constant.

After the respondents had indicated their preferences for 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, the respondents were asked to find their indifference point for the choice just made, following a procedure introduced to them in the money experiment. The procedure used was that the risk reduction in the respondent's preferred programme was decreased until the respondent was indifferent between the two choices.[5] To establish the point of indifference, all respondents were given a template containing the risk reductions, the final risks and the gain in LE in their less preferred programme. In addition they received an envelope containing 4 (3)[6] cards with different versions of their preferred programme which all offered a smaller risk reduction and hence a lower gain in LE than their preferred programme. The respondents were then asked to identify the card that made it hardest for them to choose[7].

3b Money experiment

As mentioned, the LE experiment was preceded by a money experiment to allow for the creation of an incentivised learning environment, in which respondents could be informed about the elements

[4] Risk reductions giving a gain of 6 months were chosen since the results in Chilton et al. (2004) indicate that most respondents perceive a gain in life expectancy equal to six months as 'worth buying' (whereas for a substantial proportion of the respondents one month, for example, was too little).

[5] 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.

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

[7] 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

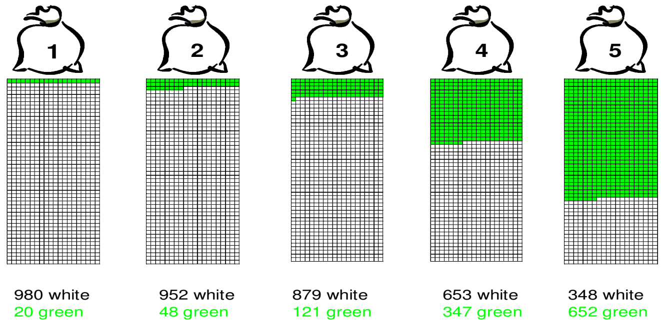
needed in order to be able to trade-off different probability distributions with the same expected gain. More specifically, the idea was to educate the respondents about conditional probabilities, letting them experience the consequences of different probability distributions and hence demonstrate to them that different probability distributions can have the same expected pay-off. 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 lifetime of different (conditional) risk distributions according to the type of programme implemented. Overhead slides, explanations and response sheet in the LE experiment mirrored as far as possible the material used in this money experiment. In addition, no new tasks were introduced to the respondents in the hypothetical setting.

The particular money experiment applied was built upon 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[8].

To introduce the respondents to the white-green composition of the bags we followed the risk communication approach described in Krupnick et al. (2002). This is illustrated in Figure 4, which depicts the first set of bags (the baseline) that reflect the initial hazard rate for a 40-year-old individual.

[8] 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.

Figure 4: Baseline set of bags. Overhead slide presented to the respondents



The same illustration outline was used later in the LE experiment showing the respondents' remaining life expectancy over the next five decades as in Figure 4, in which a blue square represented death and a white square survival in each decade

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 survival effects introduced to the respondents in the subsequent LE experiment.

The tokens collected in the games were entered into a draw for a prize[9] 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. The outline of the money game is briefly presented in Table 2 notice

[9]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).

though the expected number of tokens in the baseline game is 3.46, which is 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 2: Outline of the money experiment

Game 1	Baseline game	• Practice game (played out for token) • E(tokens)= 3.46
	Option games	• Some green cards replaced with white cards • E(tokens)=3.51 in each option
Game 2	*A vs C*	• 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.
Game 3	*B vs C*	
Game 4	*A vs B*	
Game 5	Indifference game	• 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.

X, Y, Z, in terms of the distribution of risk reduction, corresponded precisely to option A, B, C in the money games.

It had been noted in piloting the experiment that a significant number of respondents seemed to have some difficulty in shifting abruptly from money choices to life expectancy choices. Whilst perhaps gratifying in one sense, they sometimes almost complete concentration on the probability distributions in the money games meant that a number of respondents only realized they had changed context in the second or third life expectancy choice. Hence, in order to give the respondent a chance to give some thought to the shift of context from money to LE and to emphasise that there are no right or wrong choices – some brief arguments both in favour and against changing the risk ‘now’ versus ‘later’ were presented to the respondents before they made any choices concerning LE. The arguments have been put together from issues brought forward in the piloting phase. Whilst there is a chance that some respondents may have been so influenced that they changed their preferences on hearing these arguments, two things probably mitigate this effect in that the arguments were delivered in fairly neutral and bland terms and, as noted, respondents were exposed to both types of opinion.

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 (not students). Appendix C describes the sample according to socio-demographic characteristics.

All respondents recorded their answers on an individual level and no open-ended discussions were introduced during the group sessions. A session lasted between 1-1½ hours; all interviews were moderated by one of two members of the research team, and five trained assistants assisted during the group sessions. In total 130 individuals were interviewed; however 11 (8.5%) individuals displayed intransitive preferences[10] in the LE experiment and were excluded from the sample, as was one respondent above the age of 50. This gives a final sample size of 118.[11]

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 D 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 choices in almost all choices. Therefore, we can infer that the choices have not been made by automatically recording the same preferences in the hypothetical LE experiment as in the incentivised money experiment. This is in accordance with the results from the piloting phase of the experiment.

[10] 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 and not by a particular decision strategy not supported by rational choice. Thus, these individuals have not been given further attention in this paper.

[11] 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.

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

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

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

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

Table 4: Preference orderings. Number of respondents with each preference ordering

Preference ordering	**LIFE**
Z>Y>X	30
X>Y>Z	29
Y>Z>X	22
Y>X>Z	11
Y=X=Z	8
Z>X>Y	5
Other	13
Total	**118**

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

It is evident from Table 4 that the majority (70%) of respondents in their pair-wise comparisons have indicated a preference ordering of the three LE programmes in one of three main ways: Z>Y>X, X>Y>Z and Y>Z>X. 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 5 that only a small proportion of the individuals were indifferent between all three programmes.

Regarding strengths of preferences, Table 5 below contains information about the indifference point in the LE experiment; i.e. the number of months the respondents on average were willing to 'sacrifice' from their preferred strategy for reaching indifference between the two programmes in question.

Table 5: Indifference point in the LE experiment. Only with respondents displaying a strict preference in Table 3

LE (in months)	**Number of respondents**	**Mean willingness to give up (in months)**	**T-test(p-value)**
Comparison no 1 (X vs. Z)			
X	43	1.7	2.3(0.025)**
Z	63	2.2	
Comparison no 2 (Y vs. Z)			
Y	65	1.7	2.9(0.004)***
Z	43	2.3	
Comparison no 3 (X vs. Y)			
X	37	1.5	2.5(0.01)**
Y	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 X were willing to 'give up' 1.7 months of X to reach indifference between Z and X, i.e. on average, individuals were indifferent between a 6-month gain in LE from Z and a 4.3 (6-1.7) month gain in LE from X.

Significant differences are found 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 Z programme were on average willing to trade-off more risk reduction/gain in LE to arrive at their indifference point, whereas participants with preferences for X were the least willing-to-trade[12]. This on average suggests stronger preferences for programme Z.

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

4.1. Regression results

Regressions were carried to find which variables influence the 'willingness-to-trade off life expectancy' i.e. the indifference point. We used an ordered probit model in which the dependent variable was the number of months the respondent was willing-to-trade (deltaLE), hence the data set contains three observations per respondent which we assume to be independent observations.

We report the results of three different models. In Model I the respondents have been grouped into three different groups, X, Y and Z. The dummy variable X is the group of respondents with a strict preference for X (X>Y>Z or X>Z>Y or X>Y=Z) etc. It can be seen that both variable X and Y are significantly negative which indicates that respondents with a strict preference for either X or Y are less willing-to-trade than the individuals with a strict preference for Z. This, as X and Y have lower variances and assuming our individuals are risk averse is in accordance with expectations.

Model II includes the pairwise difference in variance (deltavar) as a regressor, instead of the respondent's type, and is hence positive if the respondent had a preference for the option with the highest variance. The sign of deltavar is significant and positive which indicates that respondents with a preference for a positive change in variance have a higher propensity to trade. Again this is in accordance with expectations.

Finally, Model III is a combination of models I and II. In this model only one 'type Y' is a significant variable. This could be an indication of that the dummies for 'type' and the variance variable could be picking up on the same effect.

In all models socio-demographic variables have been included but proved to be insignificant even when using a general- to- specific modelling approach.

Table 6. Regression results

DeltaLE	**Model I**	**Model II**	**Model III**
	Coef (std. error)	*Coef (std. error)*	*Coef (std. error)*
Y	-0.372 (0.145)**		-0.293(0.156)*
X	-0.469(0.151)***		-0.251(0.218)
deltavar		0.016(0.005)***	0.010(0.008)
Pseudo R2	0.0134	0.0115	0.0157
N	327	327	327

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

5. 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. This is an important finding. It is evident from Table 5 that the majority of respondents have ranked the LE programmes in one of three ways: Z>Y>X, X>Y>Z or Y>Z>X. 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 implication from a society perspective is that a mix of different LE generating policies is warranted, i.e. both programmes generating a one-period change in risk (the VPF framework) and programmes delivering an ongoing and sustained reduction in risk over the remainder of their lives.

Based on these results we could envisage that even though an individual could have a willingness-to-pay (WTP) = £0 for a risk reduction of the programme X type they could have a WTP > £0 for a risk reduction of the programme Z type. Taken together this means that using an existing VPF estimate as the starting point for the calculation of a VOLY could potentially not take into account that people would be willing to pay for a risk reduction arising from the Z programme but not for a risk reduction arising from an X type of programme. Importantly also, it implies that we need to better describe the distribution and timing of the risk reduction when valuing life expectancy gains or a VPF.

Respondents chose between programmes that had different distributions of risk. According to theory, their preferences are clearly informed by a mix of individual risk and time preferences and also the baseline level of risk, Unfortunately, our restricted set of programmes does not allow us to tease out all of these effects but in our regressions we find that the differences in the variance of the LE gains does have explanatory power, implying that our earlier assumption of risk aversion seems to hold, this is also replicated on the analysis by type of respondent. 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), it may be that context and probability distributions interact in some way to influence preferences in the other types of programme.

The present experiment addresses two key problems in an application of the risk-risk trade-off method relating to changes in mortality risks. The first is peoples' well-documented inabilities in dealing with changes in small probabilities (see, for example Corso et al. 2001)), much less *distributions* of these probabilities. The second is that the task must, by necessity, be hypothetical in nature. We show that the experimental mechanism delivers consistent and credible quantitative choices in the hypothetical setting and more specifically only a small proportion of the sample displays intransitive preferences.

Future research is needed to establish whether our results apply for other age groups and to analyse other probability distributions. Moreover, research is needed in order to disentangle the consequences of differences in time and risk preferences from initial risk effects. The theoretical and empirical work carried out 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.

8. Conclusion

Preferences for three main orderings of three LE programmes 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 society perspective is that a mix of different LE generating policies is warranted, i.e. both programmes generating a one-period change in risk (the VPF framework) and programmes delivering an ongoing and sustained reduction in risk over the remainder of their lives.

We show that the experimental mechanism developed delivers consistent and credible quantitative choices in the hypothetical setting. We would argue that this mechanism provides respondents with a more accurate understanding of how gains in life expectancy are delivered (and what is meant by this concept) than endeavours in this area to date. Finally, on a more general note, it demonstrates 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). The authors are very grateful to Xinyuan Mu, Fan Zhang, Jing Shen, Wenhua Metcalf, Jochen Fehling and Ania Bartczak for assisting with the groups and for very helpful suggestions in the earlier design phase.

References

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

Barberis, N. & Huang, M. 2008, "Stocks as Lotteries: The Implication of Probability Weighting for Security Prices", *American Economic Review*, vol. *forthcoming*.

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*.

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.

Collett, D. 1994, *Modelling survival data in medical research* Chapman & Hall/CRC, London.

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.

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

Jones-Lee, M. W. 1974, "The Value of Changes in the Probability of Death or Injury", *The Journal of Political Economy*, vol. 82, no. 4, pp. 835-849.

Kamerud, D. B. 1989, "Mortality Risk and Life Expectancy", *The Journal of the Operational Research Society*, vol. 40, no. 2, pp. 199-200.

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.

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.

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.

Appendix A

Assumption required for the calculations of life expectancy gains

Data stems from Government Actuary's Department (mean 2003-3005) 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.

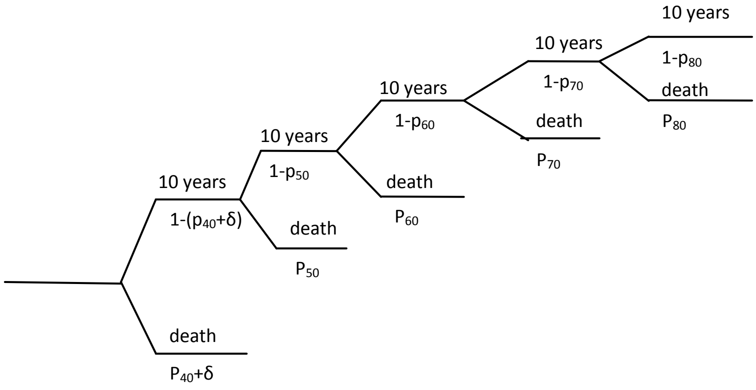
Accordingly, in the money game, 1000p_{40} is the number of green cards in the first bag and the number of white cards is 1000(1-p_{40}). In the money game this is the appropriate starting point for an estimation of E(tokens) since you either 'survive' a whole bag or 'die". However, the LE –section should 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 EQ 2:

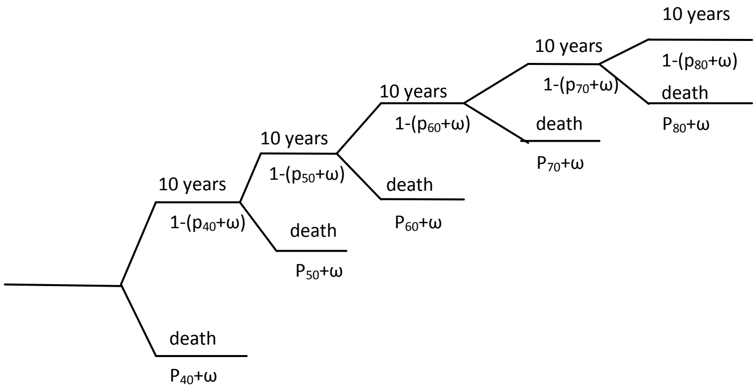
LE_{40}= 0.5*10((1-p_{40}) + (1-p_{40})(1-p_{50}) + (1-p_{40})(1-p_{50})(1-p_{60})) + 0.5*10(1+ (1-p_{40}) + (1-p_{40})(1-p_{50}) + (1-p_{40})(1-p_{50})(1-p_{60}))

This would approximate the continuous curve more correctly. 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 want to see whether preferences shift across context we want to keep all other tings constantly.

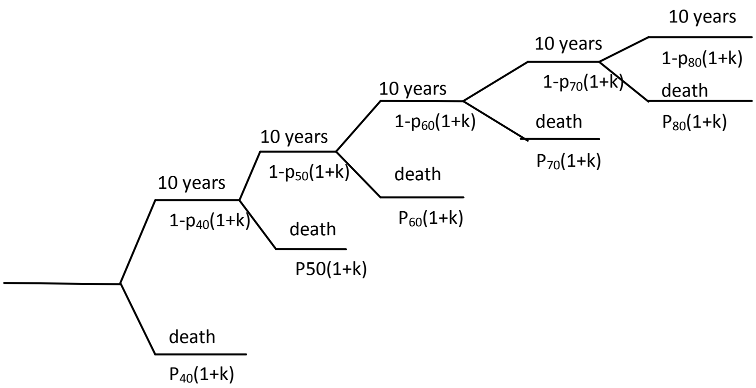
Appendix B



Programme X



Programme Y



Programme Z

Appendix C

Sample demographic

	N	**Mean**
Proportion of males	130	0.43
Age (in years)	130	41
Education (1=primary, 2= secondary, 3= higher)	130	2.5
Health (5 categories from SF-36)	130	3.8
Proportion with children	130	0.67
Proportion with a health insurance	130	0.28
Mean individual income per month (GBP)	129	1800
Mean household income per month (GBP)	124	2850

Appendix D

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