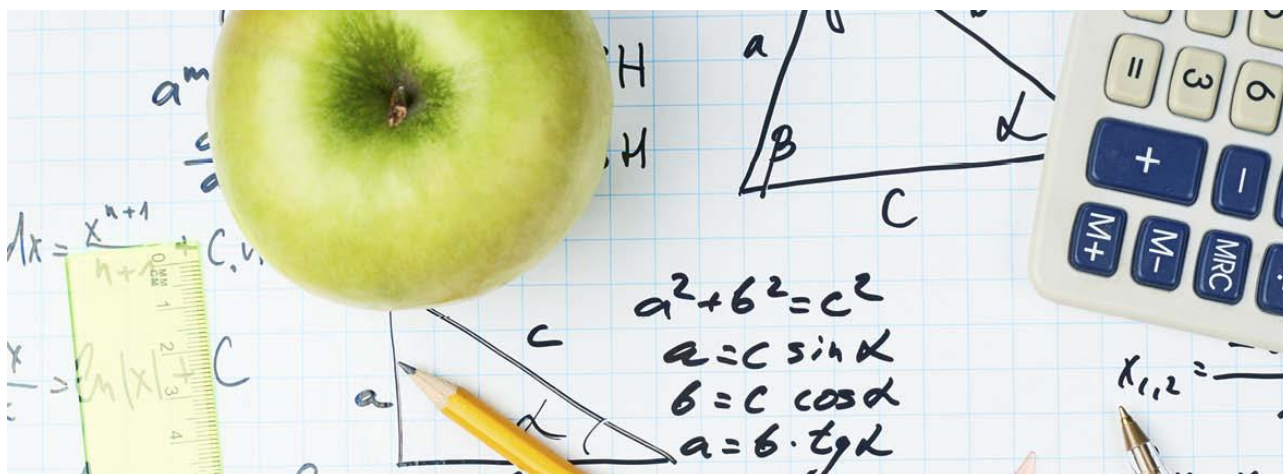


COHERE - Centre of Health Economics Research



An investigation into procedure (in)variance in the valuation of mortality risk reductions

By

Trine Kjær COHERE, Department of Business and Economics, University of Southern Denmark
Jytte Seested Nielsen, Newcastle University Business School, UK

COHERE discussion paper No. 4/2016

FURTHER INFORMATION
Department of Business and Economics
Faculty of Business and Social Sciences
University of Southern Denmark
Campusvej 55,
DK-5230 Odense M Denmark
www.cohere.dk

An investigation into procedure (in)variance in the valuation of mortality risk reductions

Trine Kjær¹ and Jytte Seested Nielsen²

Abstract

This study seeks to investigate whether elicited preferences are affected by the presentation of mortality risks in a stated preference survey. A three-way split sample discrete choice experiment was conducted in which respondents were asked to express their willingness-to-pay for public risk reducing initiatives under different but outcome equivalent representation formats. Our results demonstrate that respondents exhibit much stronger preferences for public life saving interventions when these are framed in terms of avoided fatalities compared to corresponding mortality risk reductions. Furthermore, we find that less numerate respondents are more susceptible to the inclusion of the number of fatalities in the representation format. The same pattern is observed for respondents who express a higher degree of concern for a traffic accident. In conclusion our findings may justify presenting both type of risk information in valuation of mortality risk reductions in public settings.

Keywords: discrete choice experiment, framing, mortality risk, procedure invariance, public policy, stated preferences, willingness-to-pay

JEL Classification J17, D6

¹COHERE, Department of Business and Economics, University of Southern Denmark, Campusvej 55, DK-5230 Odense M, Denmark, ²Newcastle University Business School, 5 Barrack Road, NE1 4SE Newcastle upon Tyne, United Kingdom, Contact per e-mail: Trine Kjær tkj@sam.sdu.dk; Jytte Seested Nielsen jytte.nielsen@ncl.ac.uk

1. Introduction

The value of mortality risk reduction is an important component of cost-benefit analysis of many environmental public policies such as air pollution reducing initiatives. The welfare economic approach to valuing reductions in mortality risk requires an estimate of the individual trade-off between changes in wealth and a small change in the probability of dying. Several methods can be used to estimate this trade-off including hedonic pricing and stated preference (SP) methods. The main advantage with SP is that it allows for the measurement of non-use values including altruistic preferences which can be relevant to incorporate when the risk reduction is a public good (where a public good is defined as a good which is freely accessible to all citizens and non-rival and non-excludable). In SP studies subjects are asked either directly or indirectly to state their willingness to pay (WTP) in monetary terms to reduce a specific risk. The majority of such previous SP studies have used contingent valuation (CV) techniques (e.g. Jones-Lee et al. 1985; Johannesson et al. 1996; Persson et al. 2001; Gerking et al. 2014), however recently discrete choice experiments (DCE) have likewise been applied (e.g. Andersson et al. 2016; Alberini and Scasny 2011; Cameron and Deshazo 2013; Johansson-Stenman & Martinsson 2008). Concerns have been expressed about the validity of WTP estimates from SP studies relating among other things to the difficulties involved in explaining and communicating small changes in risk to survey respondents (Corso et al. 2001; Hammitt and Graham 1999). A lack of understanding of or focus on the offered risk reductions may affect the valuation and cause biases such as scope insensitivities (Baron 1997; Frederick and Fischhoff 1998). When the risk reducing initiative is a public good, benefit can either be framed in terms of expected number of fatalities avoided in a given region or as a change in the risk of dying for the individuals in that given region (expressed either in terms of frequencies or probabilities). All type of representation formats have been widely applied in the literature and to date no consensus exists on what should be recommended practice. Furthermore no study has yet attempted to systematically compare the effect of these different representation formats on the implied valuations of outcomes. The present paper seeks to address this gap in the existing empirical research.

We extend the current literature on valuation of mortality risk reductions by investigating the potential influence that different representations of mortality risk reductions has on the elicited valuation estimates. For this purpose a three-way split sample DCE was conducted. The DCE was chosen as an alternative to CV due to the more explicit focus on the trade-off

of attributes, with-in sample scope sensitivity and the gain in statistical power due to multiple observation per individual (Adamowicz et al. 1998; Bateman et al. 2002; Goldberg and Rosen 2007). Respondents were asked to express their WTP for risk reducing initiatives under different representation formats keeping the size of the outcomes constant across splits (all in the context of traffic). In survey arm I (FATAL), respondents were given information about the number of fatalities avoided in the given population. In arm II (RISK), respondents were given information about the equivalent absolute change in mortality risk expressed in form of frequencies of 100,000 individuals, and in arm III (BOTH) respondents were provided with both types of information. By including BOTH we can examine the effect of adding information (and not just replacing information). BOTH thus represents an intermediate stage, that enable us to investigate what type of information respondents base their choices on and whether this differs across segments Gigerenzer and collaborators, have found support that individuals understand information about risk much better when given information in the form of frequencies than when given information in the form of probabilities (Gigerenzer 1996; Gigerenzer & Todd 1999; Hoffrage & Gigerenzer 1998). For this reason we deliberately decided to focus on representation of risks as frequencies. Our design allow us to systematically compare whether the trade-offs between risk and income are procedure invariant. More specifically, we address this issue by comparing marginal WTP estimates across the three different presentation formats. Representation format may play a role if differences in descriptions trigger affective reactions. We therefore investigate whether the sensitivity to changes in format will depend on the level of respondents' self-assessed affective feelings. Finally, as a proxy for numeracy, we examine whether self-assessed numerical skills influence the impact of framing. We analyse whether respondents with poorer numerical skills react differently towards the presentation format of the risk reduction than individuals with higher numerical skills.

The remainder of the paper is organised as follows. The subsequent section provides an overview of previous relevant research. Section three describes survey design, data collection, model specification, and outlines the hypothesis to be tested. Results are presented in section four and discussed in section five. The final section concludes.

2. Background

The standard economic model of preferences for wealth and mortality risk (Jones-Lee 1974, Hammitt et al., 2004) assumes that an individual's welfare can be represented as

$$EU(s, w) = su_a(w) + (1 - s)u_d(w) \quad [\text{EQ 1}]$$

where s is the individual's chance of surviving the current period and $u_a(w)$ and $u_d(w)$ represent her utility as a function of wealth conditional on survival and death in the current period, respectively. The individual's Value of a Statistical Life (VSL) is derived by differentiating EQ 1 holding utility constant

$$VSL = \frac{dw}{ds} = \frac{u_a(w) + u_d(w)}{su'_a(w) + (1-s)u'_d(w)} \quad [\text{EQ 2}]$$

Equation 2 describes the marginal rate of substitution (MRS) between risk and wealth. In surveys, respondents are asked about their WTP for a small finite risk reduction Δp and VSL is subsequently estimated as the ratio between the mean WTP and Δp .

Reductions in mortality risks Δp can either be represented in terms of changes in mortality risks or an expected number of fatalities avoided/lives saved¹ by multiplying the change in probability by the number of individuals in the given region. The latter has also been termed the 'community analogy' i.e. when the risk is presented as the number of mortal accidents within a given area (Calman and Royston, 1997). To illustrate, a risk reduction from e.g. 2 in 10,000 to 1 in 10,000 in a community with 500,000 individuals can be presented as an absolute risk reduction framed as either a standard frequency (a risk reduction of 1 in 10,000), a relative risk reduction (in this case a 50% reduction in the mortality risk) or a 'community analogy' frequency based on the number of individuals in the community (50 fatalities avoided)². A basic assumption in economic theory is that preferences are stable and that expression of preferences will not be affected by the procedure used to reveal them (Bateman et al., 2002); an assumption also known as *procedure invariance* (Tversky & Thaler 1990). Accordingly, we would not expect any difference with respect to representing the risk reduction in terms of expected number of avoided fatalities or an equivalent change

¹ Strictly speaking, a life cannot be saved but can be extended. On the other hand, a fatality can be avoided.

² The risk reduction can also be presented as a probability (in this case a risk reduction in the mortality risk of 0.01%)

in the mortality risk. In both cases the overall expected outcomes are the same. Table 1 provides an overview of previous SP studies listed according to representation format. It appears that choice of representation format is rather unsystematic and that both formats have been more or less equally adopted in public valuations of mortality risk reductions.

Table 1. List of previous SP studies (public valuation of risk reductions) according to framing format

Description	Format	Measurement	Study
RISK	Frequency format and/or relative risk reduction	WTP	Jones-Lee et al., (1985); Alberini & Scasny (2011); Johannesson et al., (1996); Maier et al. (1989); Svensson & Johansson (2010); Zhai & Suzuki (2008)
FATAL	Lives saved or avoided fatalities	WTP	Andersson et al. (2016); Andersson & Lindberg (2009); Desaignes & Rabl (1995); Hultkrantz et al., (2006); Rheinberger (2011); Strand (2005); Zhu (2004)
		Relative valuation	Chilton et al. (2002); Cropper et al. (1992); Johansson-Stenman & Martinsson (2008); Jones-Lee & Loomes (1995)

It is well documented that the concept of risk is complex and difficult to communicate. There is a large literature showing that risks and risk changes are not always perceived correctly by individuals and that individuals have difficulties understanding probabilities influencing risk assessment (Visschers et al. 2009; Slovic 2000). Some SP studies have investigated the influence of risk representation and format on elicited WTP values. Most of these studies have investigated sensitivity to the magnitude of risk reduction with some evidence of insensitivity to scope using CV (Corso et al. 2001; Goldberg and Rosen 2007; Hammitt and Graham 1999; Persson et al. 2001) and DCE (Andersson et al. 2016). On the other hand, studies on the effect of different but outcome equivalent representations formats are scarce. Some studies have found individuals to be sensitive to whether risk information is presented as absolute or relative risk reductions (Baron 1997; Gyrd-Hansen et al. 2003). The study by Zhai and Suzuki (2008) found that the choice of frequency scale (i.e. either in terms of ‘of every 100’, ‘of every 1,000’ or ‘of every 10,000’) had some effect on valuation. They found that the larger the common denominator, the less the WTP to reduce the risk of tsunamis. A

comparable result was found in Denes-Raj and Epstein (1994) where many subjects indicated a preference for a 7 in 100 winning gamble rather than 1 in 10.

It has been proposed that affect (i.e. risk as feelings) may serve as a cue for many important judgments including judgements of risk, and that different representations of risk are tagged to varying degrees with affect (Slovic et al. 2002). As shown in a number of studies, these affective feelings influence decisions without conscious deliberative input (Peters 2006). Specifically studies by Slovic and colleagues showed that representations of risk in the form of frequencies created more frightening images (vividness) than probabilities (Lowenstein et al 2001; Slovic et al. 2002; Slovic et al. 2004). That preferences are affected by affective reactions has further been supported by Rottenstreich and Hsee (2001) who found affect-rich outcomes to result in more pronounced overweighting of small probabilities than affect-poor outcomes. It has been suggested by Finucane et al. 2000 that in the process of making judgements, people consult or refer to an 'affective pool' (containing all the positive and negative tags associated with the representations consciously or unconsciously). According to Slovic et al 2002, affective responses occur rapidly and automatically. They argue that reliance on such feelings can be characterized as the affect heuristic (Slovic et al 2002). The relationship between personal experiences, affect and risk perception has also been investigated within the environmental economics literature. In analyses of the public perception of risk changes it was found that risk perception and affect reciprocally influence each other (van der Linden 2014; Leiserowitz 2006).

Furthermore it has been suggested that people with higher cognitive ability differ from those with lower cognitive ability and these differences might influence decision making (Frederik 2005). In a SP survey on mortality risk, Andersson and Svensson (2008) investigated whether respondents with a lower cognitive ability were the main drivers of decision-making heuristics. They found that WTP answers from respondents with a higher cognitive ability were less flawed by scope bias implying that the decision process for respondents with higher cognitive ability was less driven by heuristics. Research in psychology has demonstrated that numeracy skills may have important consequences for decision making, and that inadequate numeracy may in particular be an important barrier to individuals' understanding of risks (Peters 2012). There is evidence that numerate individuals are likely to pay more attention to numbers associated with a risk, and to better comprehend them and use them in decisions whereas less numerate are more likely to be informed less by numbers and more by other

sources of information such as emotions (Peters et al. 2008). In addition less numerate individuals also appear to be more susceptible to how messages are framed and how numbers are formatted possibly because they are less able (or less likely) to translate numbers across different contexts (Peters et al. 2006; Peters et al 2011).

To the best knowledge of the authors, this study is the first stated preference study that has been conducted to systematically compare the effect of different but outcome equivalent presentation formats (as identified in Table 1) on the valuation of mortality risk reductions.

3. Materials and methods

3.1 Data

The questionnaire survey was conducted as an interactive web-design in May 2013 using an Internet panel. The survey was tested in an online pilot study (n=200) in the autumn 2012. One of the purposes of the pilot study was to establish the levels of the price attribute in the DCE. Based on the results of the pilot study the questionnaire was amended. The questionnaire consisted of four parts.

Part 1: Warm-up and introductory questions. Introductory socio-demographic questions and questions related to respondents own traffic behaviour (both car and other means of transportation) and health status. These questions were included as warm-up exercise and to establish how much the respondents use different modes of transportation etc.. As a proxy for “affect” we measure whether the respondents were concerned of being involved in a traffic accident. A 1-5 Likert scale was used ranging from ‘strongly agree’ to ‘strongly disagree’.

Part 2: Information on baseline risk. Information on the yearly baseline traffic mortality risk (constituting 240 Danish lives lost yearly). This was followed by further risk communication explaining the corresponding number of lives lost out of 100,000 random Danish citizens. The information was framed as following: *“Since there are 5.5 million people in Denmark, every year 4 individuals out of 100,000 will die in the traffic. That is, every Dane has on average a risk of 4 in a 100,000 for dying in a traffic accident”*. It has been suggested that in a stated preference survey a verbal probability analogy is a good supplement to numerical

probabilities/frequencies (e.g. Corso et al. 2001, Hammitt & Graham 1999). Hence, to put the numbers into perspective, the respondents were told that 100,000 are the number of people living in the city of Aalborg (the fourth largest city in Denmark). In addition the respondents were also given the information that 100,000 is twice the number of seats in “Parken” (the national football stadium in Copenhagen).

Part 3: The SP task. The risk reducing initiative was described as a mandatory public 10-year initiative with yearly payment reducing the risk of dying in a traffic accident. All the risk reducing initiatives were accompanied by a description of two to three concrete initiatives which in combination could generate the specific risk reduction. Information about the initiatives was included to increase the reliability of the survey. In addition it was stated that the initiatives would not reduce the risk of a traffic accident with less serious outcomes than fatal. Prior to the DCE respondents were presented with a short version of a cheap talk script which not only focused on increasing the validity of the WTP response by referring to the concept of opportunity cost, but also stressed the existence of other types of risks that one could alternatively pay for. In addition respondents were asked to consider carefully whether they actually would be willing to pay the specified amounts for the risk reductions in question.

The DCE comprised of two attributes: the annual mortality risk reduction and a price attribute (framed as an extra taxation). The attributes and levels are shown in Table 2 below. A D-efficient Bayesian design was conducted using Ngene software (ChoiceMetrics 2009) with priors from the pilot study. This led to a final design with a total of 10 choice sets consisting of two hypothetical alternatives (A and B) and one opt-out (i.e. no initiative). Each respondent received all 10 choice sets. Respondents were randomly assigned to one of three survey arms only varying according to the representation of the risk reduction. In FATAL outcomes were framed in terms of number of fatalities avoided (the ‘community analogue’ frequency). In RISK outcomes were described as the corresponding change in mortality risk using a frequency format per 100,000, and in BOTH respondents were provided with both types of information. An example of each of the three DCE versions are shown in Appendix Figure A1-A3.

Table 2. Attributes and levels in the DCE

Attributes	Split	Description	Levels
Outcome	Arm I: RISK	Annual mortality risk reduction*	1/100,000
			2/100,000
			3/100,000
	Arm II: FATAL	Number of avoided fatalities every year*	60
			120
			180
Arm III: BOTH	Both types of information (Arm I + II)		
Cost		Extra annual household tax payment (in DKK)	100
			500
			1200
			2000
			5000

*Note that an annual mortality risk of 1/100,000 is equivalent to saving 60 lives (and so forth)

Part 4: Follow-up questions. Subsequent to the DCE respondents were presented with some general follow up questions regarding their valuation of the initiatives (e.g. certainty question, protest responses). Finally the questionnaire included a range of debriefing questions aimed to shed some light on the decision process. These include a question on self-assessed numerical skills measured on a 1-10 scale with ‘1’ being poor numerical skills and ‘10’ being good.

3.2. Econometric specification

The DCE data were analysed using an error component model (belonging to the family of mixed logit models) following Train (2003). Separate models were estimated for each survey arm. The utility function U for individual n of alternative i and choice set t is specified as

$$U_{nit} = V_{nit} + \varepsilon_{nit} = \alpha_{status_quo} + \beta_1 EFFECTS_n + \beta_2 PRICE_n + \varepsilon_{nit} + \mu_n E_{nit} \quad [EQ 3]$$

where α is the alternative-specific constant for the status quo (specified as choosing no intervention), β the parameters for each of the two attributes, and ε the error term assumed independent and identically distributed (IID) with type I extreme value distribution. Finally, μ is a random term with zero mean and error component E denoting the alternative specific

random individual effects. By applying this model specification we account for substitution (correlation) patterns between the policy interventions introducing heteroscedasticity in its variance and allow for repeated choices by each respondent.

Mean WTP for a mortality risk reduction (i.e. marginal rates of substitution between income and risk) was calculated as the ratio in parameters $(-\frac{\beta_1}{\beta_2})$, and WTP standard errors were obtained using the Delta Method (Hole 2007).

3.3. Testing for procedure invariance

The neoclassical assumption of procedure invariance (Tversky & Thaler 1990) is tested in the present paper by examining the effect of different but outcome equivalent representations of risk reductions on the elicited WTP values. According to standard economic theory we should not observe any differences in valuation with respect to framing in terms of avoided fatalities or change in mortality risk. Henceforth we specify the main hypothesis H_1 : Procedural invariance which we test according to the following hypothesis

$$H_0: WTP_{FATAL} = WTP_{RISK}$$

A rejection of the null-hypothesis implies that the two presentation formats lead to different valuations of the same outcome and hence will cause differences in the estimates of a value of a statistical life. To supplement this main comparison, we test for any significant difference resulting from combining information about risk and fatalities in the presentation (thereby adding information instead of replacing). Hence, as supporting hypotheses (H_{2a} and H_{2b}) we test $H_0: WTP_{FATAL} = WTP_{BOTH}$ and $H_0: WTP_{RISK} = WTP_{BOTH}$, respectively. Including this second step enable us to investigate which type of information respondents' use in their decision process and whether this differ across segments.

Previous literature suggests that risk representation might influence elicited WTP values since different representations might trigger different affective reactions. In addition evidence exist that numeracy influence risk decisions. To examine further any differences in sensitivity to representation format across individuals, we perform a series of sub-group analyses in which respondents are categorized according to two explanatory factors; 1) self-assessed numerical skills, and 2) level of concern for a traffic accident as a proxy for affective feelings. In keeping with previous findings we expect numerate individuals to be better able

to understand the communication of risk and to be less susceptible to any difference in representation than lower numerate individuals. On the other hand, we expect respondents who express a higher level of concern for traffic accidents to perceive greater risk (i.e. acting more in affect) and therefore to be more susceptible to representation format. Respondents were categorized into two groups according to their answer to a question on self-assessed numerical skills with good skills defined as those expressing >7 (43%); and similarly categorized according to concerns of being involved in a traffic accident with the concerned group defined as those expressing “agree” or “highly agree” to the question on concern related to being involved in traffic accidents (25%). Only minor correlation (0.0053) was observed between the two generated variables thus resulting in two distinctively defined measures.

The procedure invariance hypothesis and supporting hypotheses were tested using pairwise Wald tests (Wooldrige 2002). Data was analysed using Stata software.

4. Results

The survey was carried out in May 2013. The sample was obtained from The Nielsen Company’s online panel database. The survey sample is stratified with respect to gender and age as a representative sample of the Danish population above 18 years. 3600 individuals were invited (through their email) to participate in the survey. 780 individuals accepted the invitation of which 600 (200 in each of the three survey arms) completed the questionnaire leading to a completion rate at 77%. No significance was found in the differences in gender, age, household income, number of persons in household, and higher education across the three survey arms (with the exception of age between FATAL and BOTH). Table A1 in the Appendix lists selected descriptive statistics.

Estimated mean WTP and 95% CIs are reported in Table 3 whereas Table 4 lists test statistics for the hypotheses (H_1 and H_{2a+2b}). Results are reported both for the full sample as well as for subgroups segmented according to numerical skills and concerns for traffic accident. Overall, our results show that WTP estimates are affected by how the risk reducing initiative is presented. According to the test statistics we can reject our main procedure invariance hypothesis that $WTP_{FATAL} = WTP_{RISK}$ implying that the representation of risks in terms of

avoided fatalities or reduced mortality risks (frequency format) does matter. This finding is confirmed across sub-groups and thus seems very robust. Looking at the size of the estimates we see that WTP values are considerably higher (in most cases more than double) when the risk reduction is framed in terms of fatalities avoided compared to framing in terms of reduced mortality risk.

Focusing on the full sample the following pattern in WTP estimates is observed: $WTP_{FATAL} > WTP_{BOTH} > WTP_{RISK}$. Hence adding information about risk (fatalities) seems on average to have a negative (positive) significant impact on the mean WTP estimates. This suggests an underlying heterogeneity as to what type of information respondents base their decision on. When given both type of information mean WTP increases (decreases) indicating that some respondents switch their decision according to the new information of risk (fatalities) causing an inflation (deflation) of stated WTP.

To examine sensitivity to framing format with respect to individual differences in numerical skills, respondents were divided into two groups; one group expressing poor/intermediate numerical skills and the other group expressing good numerical skills. Here we see some interesting patterns. First, we find that both groups of respondents are sensitive to the representation of risk reductions with larger observed differences in absolute WTP across arms for poorer numerically skilled respondents (here WTP_{FATAL} and WTP_{BOTH} are almost three times larger than WTP_{RISK}). Furthermore, we find no significant difference between WTP_{FATAL} and WTP_{BOTH} for lower numerate implying that the following pattern is detected; $WTP_{FATAL} = WTP_{BOTH} \neq WTP_{RISK}$. Our results suggest that additional information about the mortality risk in form of frequencies do not alter perceived preferences for this group, indicating that lower numerate individuals most likely base their choices on the information on ‘avoided fatalities’. In contrast we do not observe any difference between WTP_{RISK} and WTP_{BOTH} for respondents who see themselves as numerate (i.e. $WTP_{RISK} = WTP_{BOTH} \neq WTP_{FATAL}$). These respondents do not appear to be sensitive to the additional inclusion of information on ‘avoided fatalities’. This suggests that respondents with high self-perceived numerical skills most likely base their valuation on the risk information when provided with both types of information

Table 3. Marginal WTP [95%CI] per 1/100,000 risk reduction for groups defined by numerical skills. WTP reported in DKK

	All		Numerical skills		Concern for car accident					
			Poor skills	Good skills	Much concerned	Less concerned				
Arm I: FATAL	1233 (N=200)	[1084;1381]	1264 (N=109)	[1077;1451]	1183 (N=91)	[9416;1425]	1679 (N=53)	[1325;2032]	1100 (N=147)	[937;1262]
Arm II: RISK	569 (N=200)	[444;694]	421 (N=116)	[251;591]	773 (N=84)	[5874;9592]	636 (N=49)	[402;871]	544 (N=151)	[396;692]
Arm III: BOTH	887 (N=200)	[735;1039]	1047 (N=118)	[804;1289]	735 (N=82)	[551;919]	1587 (N=49)	[1148;2028]	696 (N=151)	[542;849]

Table 4. Wald test statistics: Procedure invariance

	All	Numerical skills		Concern for car accident	
		Poor skills	Good skills	Much concerned	Less concerned
$WTP_{FATAL} = WTP_{RISK}$	<0.01***	<0.01***	<0.01***	<0.01***	<0.01***
$WTP_{FATAL} = WTP_{BOTH}$	<0.01***	0.16	<0.01***	0.75	<0.01***
$WTP_{RISK} = WTP_{BOTH}$	<0.01***	<0.01***	0.78	<0.01***	0.16
$WTP_{FATAL} = WTP_{FATAL}$		0.60		<0.01***	
$WTP_{RISK} = WTP_{RISK}$		<0.01***		0.51	
$WTP_{BOTH} = WTP_{BOTH}$		0.05**		<0.01***	

Focusing on the RISK arm we see a high discrepancy in mean WTP according to numerical skills, indicating a pronounced variation in how the two groups comprehend information on risk reductions. This findings is in accordance with the accumulating body of literature (referred to previously in this paper) demonstrating that peoples' ability to assess risks is correlated with numeracy.

Furthermore, we examined whether the level of concern for being in a car accident influenced sensitivity to representation format. Our a priori expectation was that more concerned individuals also are more likely to be steered by affect in their decision making process. To the extent that the FATAL format leads to more affective reaction we would argue that more concerned respondents would be more inclined to react on information on avoided fatalities thus increasing their WTP when this information is provided; i.e. $WTP_{FATAL} = WTP_{BOTH} > WTP_{RISK}$. In support, we would also expect to observe the opposite pattern for the less concerned as they are expected to be less steered by feelings and thus correct their answers when provided with the risk information, thus being less sensitive to the added information on avoided fatalities i.e. $WTP_{RISK} = WTP_{BOTH} < WTP_{FATAL}$. According to Table 4, these patterns are confirmed suggesting that affect is likely to explain part of the overall divergence in the elicited preferences for risk reductions across formats. Furthermore we find that WTP is higher for the more concerned respondents. This is not surprising as more concerned respondents legitimately are likely to value risk reducing initiatives higher. The higher valuation is consistently observed for all three formats however with largest discrepancy in WTP for FATAL and a non-significant difference in WTP for RISK which seems to verify that concerned respondents are very sensitive to information on avoided fatalities.

5. Discussion

We find that framing of mortality risk reductions plays an important role in the valuation of mortality risk reductions using a stated preference approach. Our findings are important as they highlight yet another source to the observed disparity between Value of Statistical Life estimates across studies seen in for example the meta-analysis by Lindhjem et al., 2011 which didn't control for the difference in framing investigated here. Specifically, we find that mean WTP based on information on number of fatalities avoided is significantly larger than WTP

based solely on information about mortality risk reduction. One likely explanation is that the FATAL format creates more frightening images due to explicit mentioning of the number of casualties in the population. In such case, the FATAL format would induce greater perceptions of risk due to the more affective-laden description leading to higher valuations. As stated by Peters et al (2006), affect sometimes may help and other times hurt decision processes; which occurs will among other things depend on how affect influences the information processing that takes place in the construction of preferences. In contrast, it could be argued that the RISK format requires better ability to understand numerical frequencies thus causing confusion which could lead to underestimation of the true risk valuation. According to Dieckmann et al (2009) less numerate individuals may (due to difficulties in risk evaluations) underweight the stated likelihood. If this is to be confirmed we would expect to see a relative lower valuation among less numerate respondents in the RISK format. Correspondingly, we do find the lowest WTP estimate of all for this subgroup (DKK 421). In addition, we find that individuals with low subjective numeracy tend to be more sensitive to presentation format (wider WTP interval) than high numeracy individuals. Likewise we find that the subgroup of respondents who express a high level of concern for a traffic accident is more susceptible to the inclusion of the number of fatalities in the representation format.

Our findings are in keeping with previous referred literature demonstrating that framing of risky choices affect individual decision making. In the past, stated preference research on framing effects in the valuation of mortality risks has often analysed differences in elicited preferences relating to whether an outcome was expressed in terms of an increase in the number of life saved or an equivalent decrease in number of fatalities (Li et al., 2010; Okder 2012). The original explanation for this type of framing effect showing differences across the loss/gain domain was based in prospect theory (Kahneman & Tversky, 1979; Tversky & Kahneman, 1992) with the empirical finding of a reference point and that people tend to be risk averse in gains but risk seeking in losses. Any change in preference found across presentation format in our analyses cannot be explained by a variation in terms of gain/loss domain and hence cannot be contributed to prospect theory.

Despite that the two different framings variants imply an identical number of expected avoided fatalities, it is of some importance to emphasise that the framing in the RISK arm relates to the valuation of a change in probabilities (one expected death with some non-zero

variance), whereas the framing in the FATAL arm relates to the valuation of saving a certain unidentified person (no variance). Hence the two frames will not necessarily be regarded equally desirable by society due to the different probability distributions (Keeney 1980; Jones-Lee & Sugden (1988) but still they are, as illustrated in Table 1, used interchangeably in empirical surveys which has been the primary motivation for the choice of study design in the present paper.

There are some limitations to this study that needs to be stressed. We use a question on subjective numerical skills as a proxy for numeracy. Although previous literature has found a correlation between subjective and objective measurement of numeracy (Fagerlin et al 2007) inclusion of a validated and more precise measure would have been preferred. Despite this we do observe some interesting and significant differences across samples that are in keeping with previous literature suggesting robustness of our results with respect to numeracy. The usefulness of applying stated preference studies to valuation of health risks has received a lot of attention in the literature. In particular, SP studies have been criticised for not being sensitive to the magnitude of the risk change (Persson et al 2001) which could be attributed to poor survey design (Hammit and Graham 1999). In this study, we use a DCE which previously have been externally validated in another public good context (Carlsson and Martinsson 2001) and shown to produce scope sensitivity to risk (Alberini and Ščasný 2011; Carlsson et al 2010). In our study we also find the parameter for the marginal utility of risk to be significant (results available from authors on request) for all sub-samples implying that respondents in general have exhibited sensitivity to scope. Our results thus seem to support previous findings that people in general are capable of making accurate risk comparisons but differ in their ability to assess risk magnitude and understand risk formats leading to biased estimate of true risk exposures (Reyna et al 2009).

We would like to strike that these results do not imply that the general public is irrational because some of them do not understand numbers as well as scientists. Furthermore we do not intend to postulate that one framing approach is inherently superior to, or less susceptible of bias, than the other. Our findings do however suggest that there is an additional challenge to researchers including those conducting valuation of risk changes as to understand how number ability interacts with how numbers are presented in order to influence the comprehension and use of numbers. Furthermore, our study highlights how presentation of risk generally influences judgment and decision making and to varying degrees depending on

the level of concern and subjective numeracy. A finding that is of broad relevance to all areas where risk information to the public is pivotal and expected to be used to reduce own risk including decisions about environment and health.

6. Conclusion

The present study contributes to the broad literature on valuation of mortality risk reductions examining procedure invariance in the presentation of mortality risk. Our results demonstrate that respondents change their choice behaviour exhibiting much stronger preferences for public life saving interventions when these are framed in terms of ‘avoided fatalities’. Hence, describing the effect *only* in terms of ‘avoided fatalities’ might cause an overestimation of respondents’ true valuation of the health effects due to a more affective reaction. For valuation based solely on information about mortality risk reduction we identify significant differences according to numerical skills. This result suggests a potential underestimation of WTP with this format due to the difficulties in understanding changes in frequencies in low numerate individuals. On the other hand in the ‘avoided fatality’ arm we don’t see discrepancies in valuations between respondents with poor and good numerical skills indicating that this way of presenting information is perceived equivalently irrespective of numerical skills. In conclusion the mixed findings – that different type of respondents seem to base their choices on different type of information - may justify for presenting both type of risk information in future valuation of mortality risk reductions in public settings.

Acknowledgement

This work was carried out as part of the project IMPROSA (Improving Road Safety) supported by the Danish Research Council. The opinions expressed in the article are solely the responsibility of the authors and do not necessarily reflect the views of the research project’s sponsors. Furthermore we like to thank discussants and participants at NHESG 2014 and HESG 2016 for very useful comments.

Appendix 1: Descriptive statistics

Table A1. Descriptive statistics (selected). Mean values (SD) reported

	All	Split I (FATAL)	Split II (RISK)	Split III (BOTH)
Gender (males in %)	0.50	.47	.54	.48
Number of persons in household	2.3 (1.2)	2.4 (1.2)	2.2 (1.1)	2.2 (1.2)
Age	47.93 (16.8)	45.85 ¹ (17.0)	48.39 (16.1)	49.56 (17.4) ¹
Yearly household income in DKK	334,118 (197,536)	324,850 (194,183)	351,462 (203,385)	325,872 (194,183)
Higher education (in %)	.455	.415	.455	.495
Very concerned of being in a car accident (<i>highly disagree (1) to highly agree (5)</i>)	3.26	3.20	3.30	3.28
Numerical skill (<i>1-10 with 10 as highest</i>)	6.72 (2.2)	6.6 (2.1)	6.8 (2.2)	6.7 (2.4)

¹ Significant difference between split I and split III (p=0.03)

Appendix 2: Example of each of the three DCE formats

Figure A1: ARM I: RISK

Which initiative do you choose?

Initiative A	Initiative B
The gain in yearly risk reduction is 1 in a 100,000 for all citizen	The gain in yearly risk reduction is 3 in a 100,000 for all citizen
Your household's extra tax payment is 100 DKK a year	Your household's extra tax payment is 500 DKK a year

Initiative A
 Initiative B
 None

Figure A2: ARM II: FATAL

Which initiative do you choose?

Initiative A	Initiative B
60 traffic fatalities avoided each year	180 traffic fatalities avoided each year
Your household's extra tax payment is 100 DKK a year	Your household's extra tax payment is 500 DKK a year

Initiative A
 Initiative B
 None

Figure A3: ARM III: BOTH

Which initiative do you choose?

Initiative A	Initiative B
<p>60 traffic fatalities avoided each year</p> <p>The corresponding yearly risk reduction is 1 in a 100,000 for all citizen</p> <p>Your household's extra tax payment is 100 DKK a year</p>	<p>180 traffic fatalities avoided each year</p> <p>The corresponding yearly risk reduction is 3 in a 100,000 for all citizen</p> <p>Your household's extra tax payment is 500 DKK a year</p>

Initiative A
 Initiative B
 None

References

- Adamowicz, W., Boxall, P., Williams, M., & Louviere, J. (1998). Stated preference approaches for measuring passive use values: choice experiments and contingent valuation. *American Journal of Agricultural Economics* 80(1), 64-75.
- Alberini, A., & Ščasný, M. (2011). Context and the VSL: Evidence from a stated preference study in Italy and the Czech Republic. *Environmental and Resource Economics*, 49(4), 511-538.
- Andersson, H., & Lindberg, G. (2009). Benevolence and the value of road safety. *Accident Analysis & Prevention*, 41(2), 286-293.
- Andersson, H., Svensson, M. (2008). Cognitive ability and scale bias in the contingent valuation method, *Environmental & Resource Economics* 39(4); 481-495
- Andersson, H., Hole, A. R., & Svensson, M. (2016). Valuation of small and multiple health risks: A critical analysis of SP data applied to food and water safety. *Journal of Environmental Economics and Management*, 75, 41-53.
- Baron, J. (1997). Confusion of Relative and Absolute Risk in Valuation. *Journal of Risk and Uncertainty* 13; 301-309
- Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., Swanson, J. (2002). *Economic valuation with stated preference techniques: a manual*. Economic valuation with stated preference techniques: a manual.
- Calman, Sir Kenneth C. and Geoff Royston. (1997). Risk Language and Dialects. *British Medicine Journal* 315, 939-942
- Cameron, T. A., & DeShazo, J. R. (2013). Demand for health risk reductions. *Journal of Environmental Economics and Management*, 65(1), 87-109.
- Carlsson, F., & Martinsson, P. (2001). Do hypothetical and actual marginal willingness to pay differ in choice experiments?: Application to the valuation of the environment. *Journal of Environmental Economics and Management*, 41(2), 179-192.
- Carlsson, F., Daruvala, D., & Jaldell, H. (2010). Value of statistical life and cause of accident: A choice experiment. *Risk Analysis*, 30(6), 975-986.
- Chilton, Susan, Judith Covey, Lorraine Hopkins, Michael Jones-Lee, Graham Loomes, and Anne Spencer. (2002). Public Perceptions of Risk and Preference-Based Values of Safety. *Journal of Risk and Uncertainty* 25(3), 211-232.
- ChoiceMetrics (2009). *Ngene 1.0 User manual & reference guide*. The cutting edge in experimental design.

- Corso, P. S., Hammitt, 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* 23(2), 165-184.
- Cropper, M. L., Aydede, S. K., & Portney, P. R. (1992). Rates of time preference for saving lives. *The American Economic Review* 82(2); 469-472.
- Denes-Raj, V., Epstein, S. (1994). Conflict Between Intuitive and Rational Processing: When People Behave Against Their Better Judgment. *Journal of Personality and Social Psychology* 66(5); 819-829
- Desaigues, B., & Rabl, A. (1995). Reference values for human life: an econometric analysis of a contingent valuation in France. In *Contingent valuation, transport safety and the value of life* (pp. 85-112). Springer Netherlands.
- Dieckmann, N. F., Slovic, P., & Peters, E. M. (2009). The use of narrative evidence and explicit likelihood by decision makers varying in numeracy. *Risk Analysis*, 29(10); 1473-1488.
- Fagerlin, A., Zikmund-Fisher, B. J., Ubel, P. A., Jankovic, A., Derry, H. A., & Smith, D. M. (2007). Measuring numeracy without a math test: development of the Subjective Numeracy Scale. *Medical Decision Making* 27(5); 672-68
- Finucane, M.L., Alhakami, A., Slovic, P., Johnson, S.M. (2000). The affect heuristic in judgments of risks and benefits *Journal of Behavioral Decision Making* 13; 1-17
- Frederick, S., Fischhoff, B. (1998). Scope (in)sensitivity in elicited valuations. *Risk Decision and Policy* 3(2); 109-123
- Frederick, S (2005). Cognitive Reflection and Decision Making. *Journal of Economic Perspectives*, 19(4); 25 -42
- Gigerenzer, G. (1996). The psychology of good judgment: frequency format and simple algorithms. *Medical Decision Making* 3; 273-280
- Gigerenzer, G., Todd, P.M. (1999). *Simple Heuristics that Make Us Smart*. Oxford University Press. Oxford. New York
- Goldberg, I., & Roosen, J. (2007). Scope insensitivity in health risk reduction studies: A comparison of choice experiments and the contingent valuation method for valuing safer food. *Journal of Risk and Uncertainty*, 34(2); 123-144.
- Gyrd-Hansen D, Kristiansen IS, Nexøe J, Nielsen JB. (2003). How do individuals apply risk information when choosing amongst health care interventions? *Risk Analysis* 23(4); 697-704
- Hammerton, M., Jones-Lee, M. W., & Abbott, V. (1982). Technical Note—Equity and Public Risk: Some Empirical Results. *Operations Research* 30(1); 203-207.
- Hammitt, J. K., & Graham, J. D. (1999). Willingness to pay for health protection: inadequate sensitivity to probability? *Journal of Risk and Uncertainty* 18(1); 33-62.

- Hammitt, J.K., Liu, J.-T. (2004). Effects of Disease Type and Latency on the Value of Mortality Risk, *Journal of Risk and Uncertainty* 28(1); 73-95.
- Hoffrage, U., Gigerenzer, G. (1998). Using natural frequencies to improve diagnostic inferences. *Academic Medicine* 73(5); 538-540.
- Hole A.R. (2007) A comparison of approaches to estimating confidence intervals for willingness to pay measures. *Health Economics*, 16(8), 827–840.
- Hultkrantz, L., Lindberg, G., & Andersson, C. (2006). The value of improved road safety. *Journal of Risk and Uncertainty* 32(2); 151-170.
- Johannesson, M., Johansson, P. O., & O'Connor, R. M. (1996). The value of private safety versus the value of public safety. *Journal of Risk and Uncertainty* 13(3); 263-275.
- Johansson-Stenman, O., & Martinsson, P. (2008). Are some lives more valuable? An ethical preferences approach. *Journal of Health Economics* 27(3); 739-752.
- Jones-Lee, M. (1974). The value of changes in the probability of death or injury. *The Journal of Political Economy* 82(4); 835-849.
- Jones- Lee, M. W., & Sugden, R. (1988). Valuing Particular as Opposed to Statistical Life. *Seoul Journal of Economics* 1(2), 163-169.
- Jones-Lee, M. W., Hammerton, M., & Philips, P. R. (1985). The value of safety: results of a national sample survey. *The Economic Journal* 95(377); 49-72.
- Jones-Lee, M. W., & Loomes, G. (1995). Scale and context effects in the valuation of transport safety. *Journal of Risk and Uncertainty* 11(3), 183-203.
- Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica: Journal of the Econometric Society*; 263-291.
- Keeney, R. L. (1980). Equity and public risk. *Operations Research*, 28(3-Part-I); 527-534.
- Leiserowitz, A. 2006. Climate change risk perception and policy preferences: the role of affect, imagery, and values. *Climate Change* 77; 45-72.
- Li, M., Veitri J., Galvani, A.P. & Champan, G.B. (2010). How do People Value Life? *Psychological Science*. 21(2); 163-167.
- Linden, S. Van der. 2014. On the relationship between personal experience, affect and risk perception: The case of climate change. *Eur. J. Soc. Psychol.* 44; 430-440.
- Lindhjem, H., Navrud, S., Braathen N.A., Biaisque, V. (2011) Risk Analysis, Valuing Mortality Risk Reductions from Environmental, Transport, and Health Policies: A Global Meta-Analysis of Stated Preference Studies. *Risk Analysis* 31(9); 1381-1406.

- Lowenstein, G.F., Weber, E.U., Hsee, C.K., Welch, N. (2001). Risk as Feelings. *Psychological Bulletin* 127(2); 267-286.
- Maier, G., Gerking, S., Weiss, P. (1989). The Economics of Traffic Accidents on Austrian Roads: Risk Lovers or Policy Deficit? *Empirica – Austrian Economic Papers* 16(2), 177-192.
- Okder, H. (2012). The illusion of the framing effect in risky decision making. *Journal of Behavioral Decision Making* 25(1), 63-73.
- Persson, U., Norinder, A., Hjalte, K., & Gralén, K. (2001). The value of a statistical life in transport: findings from a new contingent valuation study in Sweden. *Journal of Risk and Uncertainty*, 23(2), 121-134.
- Peters, E. (2006). The functions of affect in the construction of preferences. In *The Construction of Preference*. S. Lichtenstein & P. Slovic, Eds.: 454–463. Cambridge University Press. New York
- Peters, E. (2012). Beyond comprehension the role of numeracy in judgments and decisions. *Current Directions in Psychological Science*, 21(1), 31-35.
- Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., & Dickert, S. (2006). Numeracy and decision making. *Psychological Science* 17(5), 407-413.
- Peters, E., Hart, P. S., & Fraenkel, L. (2011). Informing Patients The Influence of Numeracy, Framing, and Format of Side Effect Information on Risk Perceptions. *Medical Decision Making* 31(3), 432-436.
- Peters, E. (2008). Numeracy and the perception and communication of risk. *Annals of the New York Academy of Sciences* 1128(1); 1-7.
- Peters, E. (2012). Beyond comprehension the role of numeracy in judgments and decisions. *Current Directions in Psychological Science* 21(1); 31-35.
- Reyna, V. F., Nelson, W. L., Han, P. K., & Dieckmann, N. F. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological bulletin* 135(6); 943.
- Reyna, VF., Nelson WL., Han PK, Dieckmann NF. (2009). How numeracy influences risk comprehension and medical decision making. *Psychological Bulletin* 135(6); 943-73.
- Rheinberger, C. M. (2011). A mixed logit approach to study preferences for safety on alpine roads. *Environmental and Resource Economics* 49(1); 121-146.
- Rottenstreich, Y., Hsee, C.K. (2001). Money, kisses and electric shocks: On the Affective Psychology of Risk. *Psychological Science* 12(3); 185-190
- Slovic, P.E. (2000). *The perception of risk*. Earthscan Publications.

Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G., (2002). The affect heuristic. In: Gilovich, T., Griffin, D., Kahneman, D. (Eds.), *Heuristics and Biases: The Psychology of Intuitive judgement*. Cambridge University Press, New York.

Slovic, P., Finucane, M.L., Peters, E., MacGregor, D.G. (2004). Risk as Analysis and Risk as Feelings: Some Thoughts about Affect, Reason, Risk, and Rationality, *Risk Analysis* 24(2); 311–322.

Strand, J. (2001). Public-and private-good values of statistical lives: Results from a combined choice-experiment and contingent-valuation survey (No. 2001, 31). Memorandum, Department of Economics, University of Oslo.

Svensson, M., & Vredin Johansson, M. (2010). Willingness to pay for private and public road safety in stated preference studies: Why the difference?. *Accident Analysis & Prevention*, 42(4), 1205-1212.

Train, K. (2003). *Discrete choice methods with simulation*. Cambridge. University Press, Cambridge.

Tversky, A., Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science* 211; 453-458.

Tversky, A., & Kahneman, D. (1992). Advances in prospect theory: Cumulative representation of uncertainty. *Journal of Risk and Uncertainty* 5(4); 297-323.

Tversky, A., Thaler R.H. (1990). Anomalies Preference Reversals *Journal of Economic Perspectives* 4 (2); 201 -21.

Visschers, V. H. M., Meerrens, R. M., Passchier, W. W. F. & Vries, N. N. K. (2009). Probability Information I Risk Communication: A Review of the Research Literature. *Risk Analysis* 29(2); 267-287.

Wooldridge J.W (2002). *Econometric analysis of cross section and panel data*. MIT Press. Cambridge, Massachusetts.

Zhai, G., Suzuki T. (2008) Effects of Risk Representation and Scope on Willingness to Pay for Reduced Risks: Evidence from Tokyo Bay, Japan. *Risk Analysis* 28(2); 513-522

Zhu, W. (2009). Valuation of life: a study using discrete choice analysis (No. 2004: 3). Oslo University, Health Economics Research Programme.