Article:

Kjær, T., Nielsen, J. and Hole, A.R. orcid.org/0000-0002-9413-8101 (2018) An investigation into procedural (in)variance in the valuation of mortality risk reductions. Journal of Environmental Economics and Management, 89. pp. 278-284. ISSN 0095-0696

https://doi.org/10.1016/j.jeem.2018.04.004

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

Trine Kjær[1], Jytte Seested Nielsen[2], Arne Risa Hole[3]

[1]Trine Kjær (corresponding author) tkj@sdu.dk
COHERE, Department of Public Health
University of Southern Denmark
J.B. Winsløwsvej 9B, 1.
DK-5000
Odense C
Phone: 0045-65504085

[2] Jytte Seested Nielsen jytte.nielsen@ncl.ac.uk
Newcastle University Business School

[3]Arne Risa Hole a.r.hole@sheffield.ac.uk
Department of Economics
University of Sheffield

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. The authors have no conflicts of interests to declare. We thank Editor Roger von Haefen, Professor Dorte Gyrd-Hansen and two anonymous reviewers for their helpful comments and suggestions.

1. Introduction

Valuation of mortality risk reductions constitutes an important input to cost-benefit analysis of many environmental policies such as air pollution reducing initiatives. The welfare economic approach to valuing reductions in mortality risk requires an estimate of the individual's rate of substitution between wealth and risk (Jones-Lee et al., 1985). Several methods can be used to estimate this trade off including stated preference (SP) methods such as contingent valuation and choice experiments. The estimated absolute value of a marginal risk reduction, defined as the Value of Statistical life (VSL) is obtained by dividing the estimated willingness to pay (WTP) by the corresponding risk reduction. While some SP studies have investigated (in)sensitivity to the magnitude of risk reduction (see e.g. Andersson et al., 2016), studies on the effect of different but outcome equivalent presentation formats are scarce. Gyrd-Hansen et al. (2003) and others have found individuals to be sensitive to whether risk information is presented as absolute or relative risk reductions, and Zhai and Suzuki (2008) have found that the larger denominator of the fraction (e.g. 1/100 versus 10/1000), the less the WTP for a given risk reduction.

The benefit of public risk reducing initiatives can either be presented as a change in the risk of dying (expressed as frequencies or probabilities) or as the equivalent expected total number of fatalities avoided/lives saved over a given population[1]. The latter has also been termed the 'community analogy' (Calman and Royston, 1997). To illustrate, a reduction in risk from 2 in 10,000 to 1 in 10,000 in a community with 500,000 individuals can be presented as either a standard frequency (1 in 10,000) or a relative risk reduction (50%) or presented as a 'community analogy' frequency based on the number of individuals in the community (50 fatalities avoided). According to the assumption of *procedural invariance* (Tversky and Thaler,

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

1990), presentation format should not matter as long as the expected outcome is the same. The risk reduction format (frequency and/or relative risk reduction) has been used in previous SP studies, see eg. Jones-Lee et al. (1985) and Alberini and Scasny (2011), whereas the community analogy has been used for estimating WTP in eg. Andersson et al. (2016) and Rheinberger (2011). No study has so far attempted to systematically compare the effect of these different presentation formats on the implied valuations of outcomes, and the choice of presentation format appears to be rather idiosyncratic.

There is a large literature showing that risks and risk changes are not always perceived correctly by individuals, and that individuals have difficulties understanding how probabilities influence risk assessments. It has been proposed that affect (i.e. risk as feelings) may serve as a cue for many important judgments involving risk, and that different representations of risks are associated with affect to varying degrees (Finucane et al., 2000). Studies by Slovic and colleagues have showed that presentation of risks in the form of frequencies (e.g. 1 out of 100) created more frightening images than probabilities (Slovic et al., 2002). In addition, Rottenstreich and Hsee (2001) found a more pronounced overweighting of small probabilities relating to affect-rich outcomes compared to affect-poor outcomes. According to Slovic et al. (2002) affective responses occur rapidly and automatically and reliance on such feelings can be characterized as an ‘affect heuristic’.

Furthermore, research in psychology has demonstrated that numeracy skills have important consequences for judgement and decision making, and that inadequate numeracy may be an important barrier to an individual’s understanding of risks. There is evidence that numerate individuals are likely to pay more attention to numbers associated with a risk as they comprehend them better and use them in decisions. On the other hand, the less numerate are likely to be informed more by other sources of information such as emotions, implying that they are more susceptible to how messages are framed and how numbers are formatted (Peters et al., 2006; Reyna et al., 2009).

We extend the current literature on valuation of mortality risk reductions by systematically investigating the potential influence that different presentation formats have on the elicited values. For this purpose, a three-way split sample discrete choice experiment (DCE) was conducted that include two types of risk information presented either separately or jointly. Respondents were asked to express their WTP for risk reducing initiatives keeping the size of the outcomes constant across splits (all in the context of traffic). To further our understanding of the underlying causes of variation across formats, we investigate whether numerical abilities and affective feelings can explain some of the observed discrepancies in marginal WTP. As a proxy for the former we use subjective numerical skills whereas for the latter, we use survey responses relating to level of concern for traffic accidents.

We find that presentation format significantly affects preferences and that marginal willingness to pay for a risk reduction increases significantly when framed in terms of avoided fatalities compared to corresponding frequencies. Furthermore, we find evidence that the sensitivity to format is impacted by the numerical ability of the respondents as well as their affective reaction.

2. Materials and methods

2.1. Data

The sample was obtained from the Nielsen Company's online panel database in May 2013. The survey sample was representative of the adult Danish population with respect to gender and age. 3600 individuals were invited (by email) to participate in the survey. The response rate in the survey was 17% resulting in a sample of 600 equally split across three treatment groups. For those who started the survey, the completion rate was 77%. Prior to the actual data collection, the survey was tested in an online pilot study (n=200).

The first part of the final questionnaire contained socio-demographic questions as well as questions related to respondents own traffic behaviour including the following question; ‘I am very concerned of being in a car accident’, measured on a 1-5 point Likert scale ranging from highly disagree (1) to highly agree (5). Information about annual baseline traffic mortality risk was then provided followed by risk communication explaining the corresponding number of lives lost out of 100,000 randomly selected Danish citizens. The full risk communication text can be found in the accompanying online Appendix.

The risk reducing initiative was described as a mandatory public 10-year traffic safety intervention with annual payments and annual risk reductions. The DCE comprised of two attributes; the annual mortality risk reduction and a price attribute (framed as extra taxation). The attributes and levels are shown in Table 1 below. A D-efficient Bayesian design was developed 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). Respondents were randomly assigned to one of three treatment groups that only varied in terms of the representation of the risk reduction. In treatment group FATAL, respondents were given information about the number of fatalities avoided in the given population. In treatment group RISK, respondents were given information about the equivalent absolute change in mortality risk expressed in terms of frequencies of 100,000 individuals, and in treatment group BOTH, respondents were provided with both types of information. By including BOTH we can examine the effect of adding/removing information and not just replacing information, which therefore allows us to assess the relative salience of the two types of information. See Figure 1 for an example of the DCE format for treatment group BOTH.

FIGURE 1 AROUND HERE

TABLE 1 AROUND HERE

Subsequent to the DCE, respondents were asked a range of debriefing questions. These included self-assessed numerical skills measured on a 1-10 Likert scale ranging from poor numerical skills (1) to good numerical skills (10).

2.2. Econometric specification

The DCE data were analysed using a mixed logit model (Train, 2009). The model was estimated on the pooled sample with interactions for each treatment group. The utility function *U* of individual *n* for intervention *i* in choice set *t* is specified as

$$\begin{aligned} U_{nit} = {} & \beta_1 EFFECT_{nit} + \gamma_1 PRICE_{nit} + \\ & \left[\beta_2 EFFECT_{nit} + \gamma_2 PRICE_{nit}\right] RISK_n + \\ & \left[\beta_3 EFFECT_{nit} + \gamma_3 PRICE_{nit}\right] FATAL_n + \varepsilon_{nit} \end{aligned} \quad (1)$$

Where EFFECT and PRICE are the two attributes, $\beta=(\beta_1,\beta_2,\beta_3)$ *and* $\gamma= (\gamma_1,\gamma_2,\gamma_3)$ are coefficients to be estimated, and ε_{nit} is an error term which is assumed to be independent and identically distributed (IID) type I extreme value. The utility function for the opt-out alternative is specified as a linear function of a status-quo constant, interactions between the status-quo constant and RISK and FATAL, and an IID error term[2]. The coefficient on the status quo constant is specified to be normally distributed, which allows for correlation across the choices made by the same respondent. The remaining coefficients in the model are specified as fixed.

We also estimated a restricted form of the above model without the treatment group interactions, in which the variance of the error term can vary by treatment group. This allowed

[2] We have also run models interacting socio-demographics with the status-quo. This did not change our results.

us to carry out a Likelihood Ratio test of the restricted model versus the unrestricted model in equation (1) testing for parameter homogeneity across the three treatment groups (Swait and Louviere 1993). These regression results can be found in the online Appendix.

The marginal WTP (MWTP) for a mortality risk reduction (i.e. the marginal rate of substitution between income and risk) can be calculated as the ratios of the estimated parameters. Hence $\text{MWTP}_{\text{BOTH}} = -\frac{\hat{\beta}_1}{\hat{\gamma}_1}$, $\text{MWTP}_{\text{RISK}} = -\frac{\hat{\beta}_1 + \hat{\beta}_2}{\hat{\gamma}_1 + \hat{\gamma}_2}$ and $\text{MWTP}_{\text{FATAL}} = -\frac{\hat{\beta}_1 + \hat{\beta}_3}{\hat{\gamma}_1 + \hat{\gamma}_3}$ give the estimated WTP for a mortality risk reduction for the three treatment groups respectively. Confidence intervals for the WTP estimates were obtained using the delta method. 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 influences risk decisions. To examine further any differences in sensitivity to presentation 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 being in a traffic accident as a proxy for affective feelings. In keeping with previous findings, we expect the more numerate individuals to be better able to understand the communication of risk and to be less susceptible to how the risk is presented. On the other hand, we expect respondents who express a higher level of concern for traffic accidents to perceive risks to be greater (i.e. acting more in affect) and therefore to be more susceptible to the presentation format. It is important to emphasise that we also expect concerned respondents legitimately to value risk-reducing initiatives higher than less concerned respondents for all presentation formats[3]. Respondents were categorized into two sub-groups according to their answer to the question on self-assessed numerical skills, with good skills

[3] This will not influence our test results as we only test for procedural invariance within sub-groups and not between sub-groups.

defined as those with skills above the median, i.e. expressing >7 (43%).[4] Similarly respondents were split into two sub-groups according to their concern of being involved in a traffic accident with the very concerned sub-group defined as those expressing 'agree' or 'highly agree' to the question (25%). Only a small and non-significant correlation (0.0053) was observed between the two generated variables thus resulting in two distinctively defined measures. Two additional models were estimated by adding a full set of interactions between the existing explanatory variables, as described above, and a dummy for either high self-assessed numerical skills or for being concerned about traffic accidents. All models were estimated in Stata using 500 Halton draws to approximate the log-likelihood function.[5]

2.3. Testing for procedural invariance

The assumption of procedural invariance (Tversky and Thaler, 1990) is tested by examining the effect of different but outcome-equivalent representations of risk reductions on the elicited marginal WTP values. Specifically, we test the following hypothesis;

H_0: $\text{MWTP}_{FATAL} = \text{MWTP}_{RISK}$

A rejection of the null-hypothesis implies that the two presentation formats lead to different valuations of the same outcome and hence imply differences in VSL estimates. To supplement this main comparison, we also test for any significant difference resulting from combining information about risk and fatalities in the presentation (thereby adding information instead of replacing). Hence, we test H_0: $\text{MWTP}_{FATAL} = \text{MWTP}_{BOTH}$ and H_0: $\text{MWTP}_{RISK} = \text{MWTP}_{BOTH}$,

[4] We find that our results are similar when using a score greater than 6 instead of 7 as the threshold, although they are sensitive to increasing the threshold to 8. Only 21% percent of the sample report a score higher than 8, while 57% and 43% report a higher score than 6 and 7, respectively. This means that by using 8 as the threshold we are carrying out a different comparison, i.e. individuals in the top-fifth of the self-reported numerical skills distribution versus the rest, instead of (roughly) the top half versus the bottom half of the distribution when using 6 or 7 as thresholds.

[5] Increasing the number of draws to 1,000 did not have a qualitative impact on the results.

respectively. Including this second step enables us to investigate what type of information respondents base their choices on, and whether this differs across segments. All hypotheses were tested using Wald tests adjusted for clustering at the individual level.

3. Results

Descriptive statistics can be found in the online Appendix. No significant differences were found with respect to gender, age, household income, number of individuals in household, and proportion with higher education across the three treatment groups (with the exception of age between FATAL and BOTH). Estimated marginal WTP and 95% CIs are reported in Table 2 whereas Table 3 presents the test statistics for our primary and secondary hypotheses. Finally, regression results are reported in Table 4 both for the full sample as well as for respondents segmented according to numerical skills and concerns for traffic accident.

Overall, our results suggest that the marginal WTP estimates are affected by presentation format. According to the Swait-Louviere test we reject the null hypothesis of parameter homogeneity across the treatment groups. The Wald test rejects our main procedural invariance hypothesis that $WTP_{FATAL} = WTP_{RISK}$, implying that presenting risks in terms of avoided fatalities or reduced mortality risks (frequency format) significantly influences the trade off between income and risk. 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 rather than in terms of frequencies. All our findings are robust to conducting separate regression analyses for each treatment group[6].

INSERT TABLE 2 AROUND HERE

[6] Given that the average household size in our sample is 2.3 our WTP results correspond to VSL estimates of DKK 24-54 million (EUR 3.2-7.2). This range of VSL estimates are well within the interval observed more recently in the literature (Lindhjelm et al., 2011; Hultkrantz & Svensson, 2012).

INSERT TABLE 3 AROUND HERE

Focusing on the full sample, the following pattern in WTP estimates is observed: $\text{MWTP}_{\text{FATAL}} > \text{MWTP}_{\text{BOTH}} > \text{MWTP}_{\text{RISK}}$. Hence adding information about risk (fatalities) has a negative (positive) impact on marginal WTP estimates. In addition, comparing RISK to BOTH we see that marginal WTP is altered significantly (at the 10% level), indicating that the new information of fatalities causes an inflation of stated WTP[7].

We divided respondents into two sub-groups according to numerical skills and find some interesting patterns. First, we find that both segments are sensitive to the presentation format with larger observed differences in absolute WTP across treatment groups for poorer numerically skilled respondents (here $\text{MWTP}_{\text{FATAL}}$ is over three times as large as $\text{MWTP}_{\text{RISK}}$). Furthermore, for the less numerate, we find no difference between $\text{MWTP}_{\text{FATAL}}$ and $\text{MWTP}_{\text{BOTH}}$ but a significant difference (at the 5% level) between $\text{MWTP}_{\text{BOTH}}$ and $\text{MWTP}_{\text{RISK}}$, implying that the following pattern is; $\text{MWTP}_{\text{FATAL}} = \text{MWTP}_{\text{BOTH}} \neq \text{MWTP}_{\text{RISK}}$. Our results suggest that additional information about mortality risk, in the form of frequencies, does not alter perceived preferences for this sub-group, indicating that less numerate individuals most likely base their choices on the information about 'avoided fatalities'. In contrast, we do not observe any difference between $\text{MWTP}_{\text{RISK}}$ and $\text{MWTP}_{\text{BOTH}}$ for the sub-group of respondents who see themselves as numerate (p=0.92). Hence, these respondents do not appear to be sensitive to the additional inclusion of information about 'avoided fatalities'. This suggests that respondents with high self-perceived numerical skills base their valuation on the risk information when provided with both types of information. Furthermore, we are not able to reject the null-hypotheses of equal WTPs between FATAL and RISK as well as FATAL and BOTH for the

[7] As our Wald tests are adjusted for clustering at the individual level this and other reported results should be viewed as conservative in the direction of not detecting significant differences.

numerate respondents suggesting that they overall are less affected by format[8]. Focusing on RISK, we see a relative large difference in marginal WTP according to numerical skills, indicating a pronounced variation in how the two sub-groups comprehend information on risk reductions. Specifically, we observe a significant lower WTP in RISK among the less numerate respondents suggesting a potential underestimation of marginal WTP due to difficulties in understanding changes in frequencies (p=0.12). In our study, we use a question on subjective numerical skills as a proxy for numeracy. Although previous literature has found a correlation between subjective and objective measurements 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 consistent with the accumulating body of literature (referenced previously) 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 presentation format. Our a priori expectation was that more concerned individuals are also more likely to be steered by affect in their decision-making process. To the extent that the FATAL format leads to a more affective reaction we would argue that more concerned respondents would be more sensitive to information on avoided fatalities, thus increasing their WTP when this information is provided; i.e. $MWTP_{FATAL} = MWTP_{BOTH} > MWTP_{RISK.}$ We would also expect to observe the opposite pattern for the less concerned, as they are expected to be less steered by feelings and not as likely to change their answers when provided with additional information about avoided fatalities i.e. $MWTP_{RISK} = MWTP_{BOTH} < MWTP_{FATAL}$. According to Table 3, 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,

[8] It should be noted that the size of the WTP values are very different8 suggesting that our non-significant test results could be driven by sample size and therefore should be interpreted with caution.

we find that WTP is higher for the more concerned respondents. This is not surprising as more concerned respondents are likely to value risk-reducing initiatives more strongly. The higher valuation is consistently observed across all three formats, with the largest discrepancy in WTP for BOTH, which reinforces the earlier finding that the two sub-groups rely on different risk information in their valuation. More specifically, concerned respondents focus their attention on the information on avoided fatalities whereas less concerned encompasses the information on frequencies and thus adjust their valuation accordingly.

TABLE 4 AROUND HERE

In this study, we use a DCE which previously has been externally validated in another public good context (Carlsson and Martinsson, 2001) and shown to produce scope sensitivity to risk (Alberini and Ščasný 2011). In our study, we also find the parameter for the marginal utility of risk to be significant 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 a biased estimate of true risk exposures) (Reyna et al., 2009).

4. Conclusions

The present study contributes to the broad literature on valuation of mortality risk reductions highlighting another source of the observed disparity in VSL estimates found across SP studies (Lindhjem et al., 2011). We find that the framing of mortality risk plays an important role in the valuation of mortality risk reductions. More specifically, our results demonstrate that describing the effect *only* in terms of 'avoided fatalities' could be argued to cause an overestimation of respondents' true valuation due to a more affective reaction suggesting that information on avoided fatalities creates more 'frightening images' than information on frequencies. Furthermore, we find that the observed discrepancies at least partly are influenced by numerical ability of the respondents. In particular, our results seem to indicate that less numerate individuals have difficulties comprehending risk information and thus focus their attention on the 'easy-to-evaluate' avoided fatalities information. The same pattern is observed for respondents who express a higher degree of concern for traffic accidents; they too are more susceptible to information about avoided fatalities. We do not intend to postulate that one framing approach is inherently superior to, or less susceptible of bias, than the other. Moreover, our study does not provide an answer on how to mitigate the influence of presentation format on risk valuations. We believe that it is a question for future research to examine whether learning mechanisms (such as a 'rationality spillover device' similar to the one used in Nielsen et al. (2010)) could help respondents perceive the risks levels in the same way across different presentation formats, or whether respondents ultimately differ in the type of assistance they need in making decisions. Our findings suggest that there is an additional challenge for researchers valuing changes in risks to understand how numerical skills and affect interacts with presentation format to influence the comprehension and use of numbers. This is a finding that should be of broad relevance to all areas where risk information to the public is pivotal, including decisions about the environment and health.

Acknowledgement

Blinded

References

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., 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.

Calman, K.C., Royston, G., 1997. Risk Language and Dialects. British Medicine Journal 315, 939–942.

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.

ChoiceMetrics, 2009. Ngene 1.0 User manual & reference guide. The cutting edge in experimental design.

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.

Gyrd-Hansen, D., Kristiansen, I.S., Nexøe J., Nielsen, J.B., 2003. How do individuals apply risk information when choosing amongst health care interventions? Risk Analysis 23(4), 697-704.

Hultkrantz, L., Svensson, M., 2012. The value of a statistical life in Sweden: A review of the empirical literature. Health Policy 108, 302–310

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.

Lindhjem, H., Navrud, S., Braathen N.A., Biausque, 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.

Nielsen, J. S., Chilton, S., Jones-Lee, M., Metcalf, H., 2010. How would you like your gain in life expectancy to be provided? An experimental approach. Journal of Risk and Uncertainty 41(3), 195–218.

Peters, E., Västfjäll, D., Slovic, P., Mertz, C. K., Mazzocco, K., Dickert, S., 2006. Numeracy and decision making. Psychological Science 17(5), 407-413.

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.

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., 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.

Swait, J., Louviere, J., 1993. The Role of the Scale Parameter in the Estimation and Comparison of Multinomial Logit Models. Journal of Marketing Research 30(3), 305-314.

Train, K., 2009. Discrete choice methods with simulation. Cambridge. University Press, Cambridge.

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

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.

Tables and figures

Table 1. Attributes and levels in the DCE

Attributes	Treatment group	Description	Levels
EFFECT	FATAL	Number of avoided fatalities every year[1]	60 120 180
	RISK	Annual mortality risk reduction[1]	1/100,000 2/100,000 3/100,000
	BOTH	Both types of information (FATAL + RISK)	
PRICE		Extra annual household tax payment (in DKK)	100 500 1200 2000 5000

Notes: [1] annual mortality risk of 1/100,000 is equivalent to saving 60 lives (and so forth)

Table 2. Marginal WTP [95%CI] per 1/100,000 risk reduction. Reported in DKK (2013).

Treatment group	All	Numerical skills		Very concerned for car accident	
		Poor skills	Good skills	Agree	Disagree/Neutral
FATAL	1236 (N=200)	1269 (N=109)	1186 (N=91)	1679 (N=53)	1105 (N=147)
	[925;1547]	[895; 1642]	[649; 1723]	[932; 2426]	[768; 1442]
RISK	560 (N=200)	415 (N=116)	761 (N=84)	637 (N=49)	531 (N=151)
	[348;772]	[146; 684]	[411; 1111]	[188; 1086]	[292; 770]
BOTH	888 (N=200)	1049 (N=118)	735 (N=82)	1586 (N=49)	700 (N=151)
	[602;1175]	[575;1521]	[392; 1078]	[632; 2540]	[424; 975]

Notes: CIs are adjusted for clustering at the individual level.

Table 3. Test statistics. P-values are reported.

	All	Numerical skills		Very concerned for car accident	
		Poor skills	Good skills	Agree	Disagree/ neutral
$WTP_{FATAL} = WTP_{RISK}$	<0.01***	<0.01***	0.19	0.02**	<0.01***
$WTP_{FATAL} = WTP_{BOTH}$	0.11	0.47	0.16	0.88	0.07*
$WTP_{RISK} = WTP_{BOTH}$	0.07*	0.02**	0.92	0.08*	0.36
$WTP_{FATAL} = WTP_{FATAL}$		0.80		0.17	
$WTP_{RISK} = WTP_{RISK}$		0.12		0.68	
$WTP_{BOTH} = WTP_{BOTH}$		0.29		0.08*	

Notes: Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01. Test statistics are adjusted for clustering at the individual level.

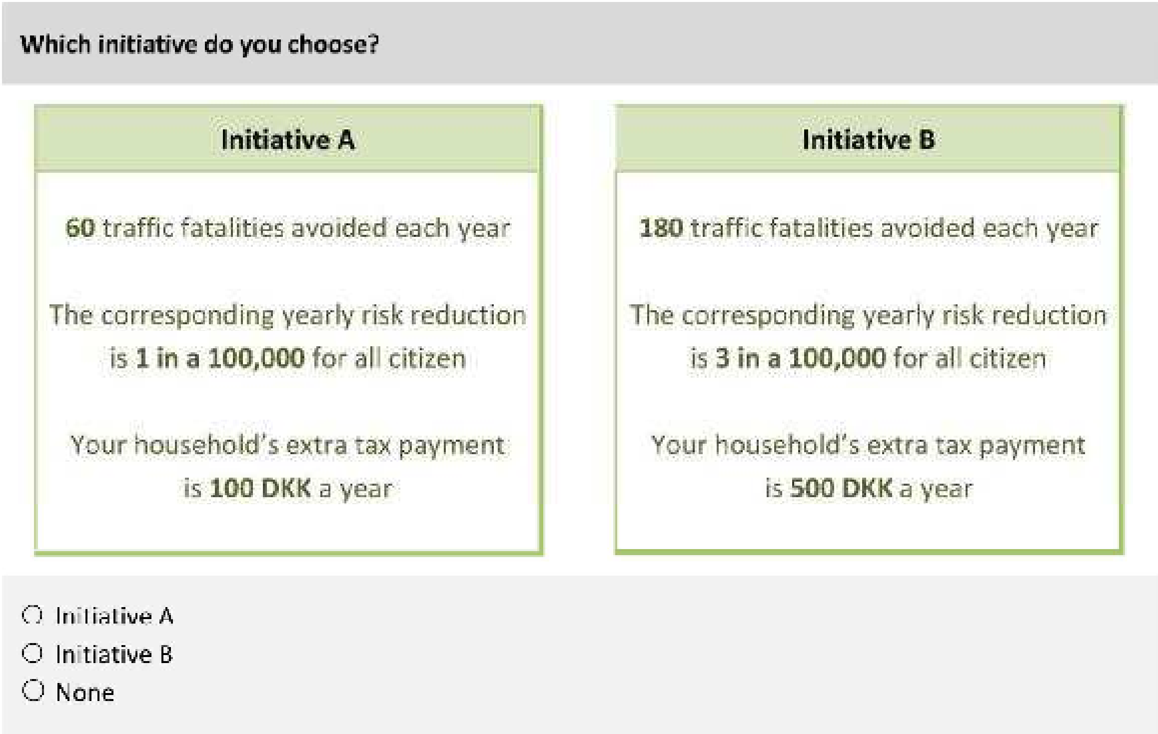
Table 4. Regression results

Variable	Model 1: Baseline model		Model 2: Good numerical skills interactions				Model 3: Concerned about traffic accidents interactions			
	Coefficient	SE	Coefficient	SE	Interaction	SE	Coefficient	SE	Interaction	SE
EFFECT	50.964***	7.290	49.219***	9.753	5.103	14.688	49.159***	8.617	7.913	16.645
PRICE	-0.057***	0.006	-0.047***	0.007	-0.027**	0.014	-0.070***	0.009	0.034***	0.012
sq[1] (mean)	-1.344***	0.302	-1.352***	0.408	-0.006	0.601	-1.084***	0.340	-1.223	0.766
sq[1] (SD)	3.969***	0.225	3.978***	0.225			3.955***	0.224		
RISK[2] × *EFFECT*	-7.040	10.447	-16.721	14.115	22.779	20.909	-8.147	12.056	2.978	24.487
RISK × *PRICE*	-0.021**	0.011	-0.031**	0.013	0.026	0.022	-0.007	0.013	-0.038	0.025
RISK × *sq*	0.611	0.443	0.362	0.577	0.623	0.898	0.450	0.509	0.837	1.052
FATAL[3] × *EFFECT*	20.176*	10.601	26.383*	14.243	-15.790	21.392	18.165*	12.532	8.353	23.645
FATAL × *PRICE*	0.000	0.008	-0.013	0.010	0.032*	0.018	0.009	0.011	-0.023	0.017
FATAL × *sq*	-0.076	0.425	-0.085	0.560	0.039	0.866	-0.281	0.489	1.008	1.020
Respondents	600		600				600			
Observations	6000		6000				6000			
Swait-Louviere test P-value	0.000		0.000				0.000			
Log-likelihood	-4695.0		-4681.2				-4663.5			

Notes: Columns 2 and 3 report the results from the baseline model without either numerical skills or traffic accident interactions, while columns 4-7 and 8-11 report the results from the models with numerical skills and traffic accident interactions, respectively. Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01. Reported SEs are adjusted for clustering at the individual level.

[1] status quo constant; [2] treatment group RISK; [3] treatment group FATAL

Figure 1. An example of a DCE for treatment group BOTH presenting both types of risk information



Notes: the text about annual risk reduction is dropped in treatment group FATAL whereas the text about fatalities is dropped in treatment group RISK

ONLINE APPENDIX

Appendix 1: Descriptive statistics (selected). Mean values (SD)/median reported

	All	Treatment group		
		FATAL	RISK	BOTH
Males	50.0%	47.0%	54.0%	48.0%
Number of individuals 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	45.5%	41.5%	45.5%	49.5%
Very concerned of being in a car accident	3.26(1.18)/3	3.20(1.19)/3	3.30 (1.19)/3	3.28 (1.14)/3
Highly agree (1)	8.61%			
Agree (2)	16.89%			
Neutral (3)	31.42%			
Disagree (4)	26.18%			
Highly disagree (5)	16.89%			
Numerical skills[2]	6.72 (2.2) /7	6.6(2.1) /7	6.8 (2.2)/7	6.7 (2.4) /7

[1]Significant difference between FATAL and BOTH (p=0.03)

[2]1-10 Likert scale from poor (1) to good (10)

Appendix 2: Survey text (selected). Translated by the authors

Text on risk communication

In recent years, around 240 Danes have died in the traffic every year. There are approximately 5.5 million people living in Denmark. This means that every year 4 individuals out of 100,000 people in Denmark will die in a traffic accident.

As a comparison, you can think about the population in Aalborg which is around 100,000. It is therefore the same as saying that every year, 4 people in Aalborg would die as a result of a traffic accident. 100.000 is also twice the population in Roskilde or Vejle. Or twice the number of seats in Parken, Copenhagen.

Introductory text to DCE

Imagine that the government is considering implementing one of two potential interventions. Both will reduce the risk of dying in a traffic accident for you, your family and others over the next decade. The intervention could be one of the following:

- more street lightening in mornings and evenings
- initiatives to decrease the number of bicycle accidents caused by a lorry turning right when bicyclists are driving straight ahead
- better marking of pedestrian walkways and road lanes
- better signage

In the following 10 questions, you will be presented with a choice between two different initiatives which will deliver different reductions in the number of fatalities at different prices. The interventions will for an extra tax payment per household reduce the risk of dying in the traffic. The risk of a traffic accident with non-fatal outcomes will not be reduced by the interventions. We will ask you to choose which of the initiatives you would prefer the government to implement. You can also choose to indicate that the government shouldn't implement any initiatives.

Remember that the risk of dying in a traffic accident as you are presented for in this survey is only one form of risk you face in life. Therefore, we will also ask you to think about how important you think it is to reduce exactly this risk and how much you would be willing to pay out of your household’s annual budget over the next decade.

Appendix 3: Regression results – restricted models used in Swait-Louviere tests

	Model 1: Baseline model		Model 2: Good numerical skills interactions				Model 3: Concerned about traffic accidents interactions			
Variable	Coefficient	SE	Coefficient	SE	Interaction	SE	Coefficient	SE	Interaction	SE
Utility										
EFFECT	48.819***	3.912	43.398***	4.816	13.625*	7.270	50.070***	4.565	1.271	8.559
PRICE	-0.058***	0.003	-0.052***	0.004	-0.015**	0.006	-0.067***	0.004	0.023***	0.006
sq[1] (mean)	-1.128***	0.188	-1.224***	0.242	0.204	0.355	-1.039***	0.218	-0.595	0.423
sq[1] (SD)	3.696***	0.256	3.722***	0.258			3.827***	0.267		
Scale of error term										
RISK[2]	0.304***	0.076	0.365***	0.098	-0.151	0.130	0.192**	0.082	0.221	0.168
FATAL[3]	-0.072	0.077	0.004	0.098	-0.197	0.131	-0.163*	0.084	0.220	0.162
Number of respondents	1800		1800				1800			
Number of observations	18000		18000				18000			
Log-likelihood	-4722.162		-4718.420				-4700.542			

Notes: Significance levels are shown as *p<0.1, **p<0.05, ***p<0.01. The scale of the error term μ_n is inversely related to the error variance: $\mathrm{var}(\varepsilon_{nit}) = \pi^2 / 6\mu_n^2$. μ_n is modelled as an exponential function of the variables listed under 'Scale of error term'.

[1] status quo constant; [2] treatment group RISK; [3] treatment group FATAL