Accounting for uncertainty in the presentation and valuation of travel time in SP experiments

Ariane Kehlbacher*[1], Kerstin Stark[1], and Laura Gebhardt[1]

[1]Researcher, German Aerospace Center DLR, Berlin, Germany

Short summary

Uncertainty is an important consideration when making decisions which means it is important when valuing travel time reliability in SP experiments. This study makes two contributions. First, it demonstrates how to quantify and communicate the uncertainty of quantities of interest when providing value estimates of travel time and reliability to decision-makers. To this effect Bayesian inference is used to propagate uncertainty of parameter estimates to quantities of interest. We find that in particular the estimate of the value of reliability is uncertain, something which would have been missed, if only point estimates are used in its computation. Second, this study uses a novel approach to represent travel time uncertainty in the SP experiment by using quantile dotplots, which we find respondents to perceive as helpful and understandable.
Keywords: value of reliability, travel time uncertainty, quantile dotplots, Bayesian inference

1 Introduction

Uncertainty is an important consideration when valuing travel time reliability in SP experiments. First, it is an important consideration when making transport investment decisions based on value estimates such as the value of reliability (VOR) or value of travel time (VTT). However, studies typically provide a single best estimate only ignoring all sources of uncertainty. Decision-makers using these values may therefore be over-confident, which can waste resources when projects that are provide little value are given the green light. This study therefore uses Bayesian inference and demonstrates how parameter uncertainty can be easily propagated to any quantities of interest that are computed from the parameter estimates.
Second, uncertainty about the exact travel time needs to be successfully communicated to participants, so that they are able to make trade-offs between travel time unreliability and other attributes. To this effect, studies use two different approaches to communicating travel time uncertainty to participants in SP experiments. The first approach involves the use of numerical (e.g. Bates et al., 2001; Li et al., 2010) or textual information Black & Towriss (1993); Small et al. (1999); Alonso-Gonzalez et al. (2020). The second approach uses visual representations of travel time unreliability using histograms (e.g. Copley et al., 2002; Tilahun & Levinson, 2010). A different approach is based on findings in the literature on visual communication and statistical reasoning. It finds that visual presentations using quantile dotplots (Wilkinson, 1999) representing a distribution where dots are sampled proportional to the quantiles of the distribution, are better suited to communicating uncertain quantities. Kay et al. (2016) evaluate different formats for visualizing uncertainty in bus transit prediction. They identify quantile dotplots as best for communicating bus transit time uncertainty on mobile applications. They jointly encode the underlying data as well as the uncertainty, which is preferred to using a extrinsic representations of uncertainty, which risk being viewed as peripheral, and consequently discounted. Castro et al. (2022) also find quantile dotplots to outperform other uncertainty presentation formats in a resource allocation task, while Frans et al. (2023) are unable to find a direct relation between visualization format and decision quality. This study conducts a mode choice experiment involving shuttles where travel time is visualised using quantile dotplots so that travel time magnitude and uncertainty are jointly communicated.

2 Methods

Data

Data collection took place in February and March of 2021. The study area is the German city of Hamburg and its adjacent districts. The online survey was administered to a panel of 1,000 respondents, of which 987 completed the survey. Respondents whose main transport modes were walking or cycling were excluded from the sample as well as those with work commute distances of more than 100 km. The final sample size is $N = 745$ respondents who each completed $S = 9$ choice cards with three alternatives each, resulting in a total of 6,705 observations. Descriptive sample statistics are reported in Table 1.

Table 1: Socio-demographic characteristics of the sample including possibility of arriving earlier and later at work (N=745). [Ref] indicates the level that is used as reference in the model estimation.

Variable	Category	%	Variable	Category	%
distance	$\leq$ 10 km	51	arrive early	1 to 10 min	58
	11-20 km	24		11 to 30 min	27
	21-50 km	20		more than 30 min [Ref]	15
	51-250 km	0	arrive late	1 to 10 min	60
age	$\leq$ 30	18		11 to 30 min	23
	31-50	38		more than 30 min [Ref]	17
	51-99	44	monthly hh income	< EUR 1500	17
gender	female/div.	49		EUR 1500-2999	38
	male [Ref]	51		EUR 3000-4999 [Ref]	33
education	high	58		EUR 5000-6999 [Ref]	9
	low [Ref]	42		EUR 7000+ [Ref]	3
work	full time	52			
	part-time [Ref]	48			

SP experiment and survey

Alternatives in the SP experiment have attributes mean and standard deviation of the travel time distribution and travel costs. The levels of the travel time mean were pivoted around respondents' self-reported travel times using a semi-pivot design (see Table 2). Actual travel times are based on the last work commute or leisure trip by car or public transport prior to the Covid-19 pandemic. The design of the choice cards was optimised using the software NGENE to obtain a D-efficient design (Rose & Bliemer, 2009) and choice cards with a dominant alternative were removed from the final choice set. Each respondent completed nine choice cards an example of which is shown in Figure 1. It shows how quantile dotplots are used to depict the travel time distributions. A total of ten dotplot was chosen as it was deemed large enough to convey the overall shape of the travel time distribution, whilst still being visually countable on a small screen. Respondents were also asked

Table 2: Attributes and their levels. M(travel time) is the mean of the travel time distribution. SD(travel time) is the standard deviation of the travel time distribution.

actual travel time	M(travel time) (in minutes)	travel cost (in EUR)	SD(travel time) (in minutes)
1 to 19 min	10, 15, 20	3, 4, 5, 6, 7, 8	1.05, 1.12, 1.16
20 to 29 min	20, 25, 30	3, 4, 5, 6, 7, 8	1.03, 1.05, 1.12
30 to 39 min	30, 35, 40	4, 5, 6, 7, 8, 9	1.01, 1.05, 1.11
40 to 49 min	40, 45, 50	4, 5, 6, 7, 8, 9	1.01, 1.05, 1.08
over 50 min	50, 55, 60	5, 6, 7, 8, 9, 10	1.01,1.02, 1.05

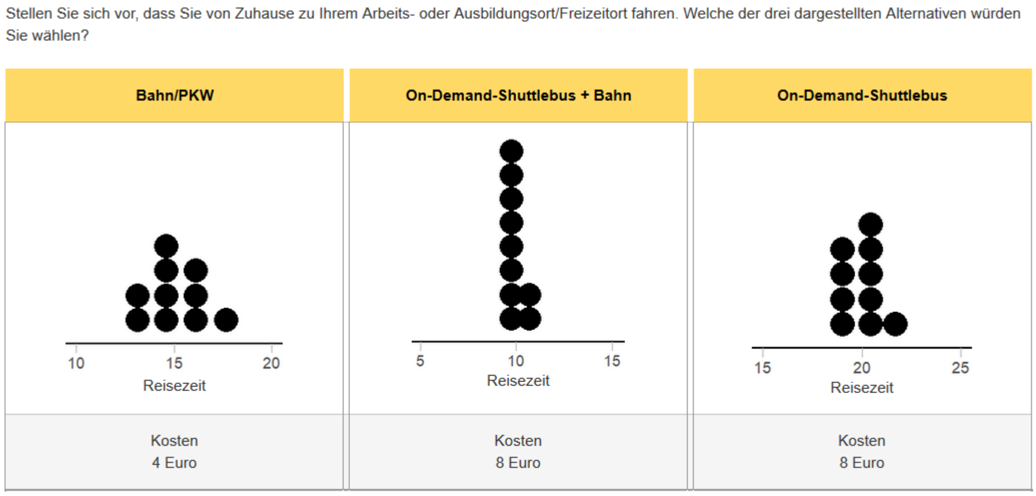


Figure 1: Example choice card with travel time distribution presented via quantile dotplots. Each quantile dot represents a 10% probability of travel time, which can be interpreted as one out ten possible travel times.

to indicate on a 5-point Likert scale their agreement to the following three statements on quantile dotplots: 1) "This way of depicting travel time in the choice experiment is understandable" 2) "Presenting travel time in this way is helpful" and 3) "Presenting travel time in this way would be helpful in real life". In addition, questions on socio-demographics were asked.

Model specification and estimation

In the transportation research literature the terms unpredictable travel time variability and travel time (un)reliability are used interchangeably, where high variability means high unreliability, and vice versa. This makes it natural to conceptualize travel time as a distribution in the probability theory sense (Carrion & Levinson, 2012). As a consequence, the travel time distribution can be described by its measures of central tendency. They are the mean, which constitutes expected travel time, and the standard deviation, which constitutes travel time (un)reliability. Then, assuming the mean-dispersion model (Jackson & Jucker, 1982; Small et al., 1999), the direct utility function for a given alternative is give by

$$U = \beta_{Car}\text{Car} + \beta_{PT}\text{PT} + \beta_{SD}SD(T) + \beta_M M(T) + \beta_C C \tag{1}$$

where $SD(T)$ and $M(T)$ are the standard deviation and mean of travel time T, C is travel cost, and β_M, β_{SD}, β_C are the respective marginal utility coefficients. Car and PT are indicators for status quo options, car or public transport, respectively. Coefficients β_{Car} and β_{PT} capture average utility of unobserved factors related to car and public transport relative to alternatives involving a shuttle. To estimate the utility function in (1) we specify the following indirect utility function. For individual i the utility of alternative j in choice set t is given by

$$v_{ijt} = X_{ijt}\beta_i + \varepsilon_{ijt} \tag{2}$$

where ε_{ijt} is iid Gumbel-distributed, β_i is a $(K \times 1)$ vector of marginal utility coefficients describing the preferences of individual i, X_{ijt} is a $(K \times 1)$ vector of explanatory variables describing attributes in alternative j in choice set t. Variables included in X_{ijt} are the mean of the travel time distribution in minutes, the standard deviation of the travel time distribution in minutes, travel costs in Euros, and indicators for car and public transport. To account for preference heterogeneity, utility parameters in (1) are allowed to vary as a function of socio-demographic characteristics which are reported in Table 1. Individual marginal utility parameters of respondent i for alternative j are distributed as follows

$$\beta_{ij} \sim N(\beta + \Gamma_j W_i, \Sigma) \tag{3}$$

where $\beta = (\beta_{car}, \beta_{pt}, \beta_{SD}, \beta_T, \beta_C)$ is a $K \times 1$ vector of means, Σ is a $K \times K$ covariance matrix; Γ_j is $K \times L$ matrix of parameters. W_i is a $(L \times 1)$ vector of variables as a function of which marginal utility parameters are allowed to vary across the population. Descriptive statistics of these variables are reported in Table 1). We allow for correlation of parameters of attributes that are common across alternatives by specifying the covariance matrix as follows

$$\Sigma = \text{diag}(\tau) \times \Omega \times \text{diag}(\tau) \tag{4}$$

where τ is a $k \times 1$ vector of parameter scales and Ω is a $k \times k$ correlation matrix.
Because we use Bayesian inference a full probability model is fitted to the observed data, y, as summarised by the probability distributions of the model parameters $\theta = \{\beta, \beta_i, \gamma, \tau, \Omega\}$. The probability model is specified in terms of the likelihood, $p(y|\theta)$ and the prior distributions of the parameters, $p(\theta)$, so as to obtain the posterior distributions $p(\theta|y)$ using Bayes' Theorem as follows

$$p(\theta|y) = \frac{p(y|\theta)p(\theta)}{p(y)} \propto p(y|\theta)p(\theta) \tag{5}$$

It implicitly also conditions on the observed values of the covariates x, which are omitted to simplify the notation. The likelihood is given by equations (2), (3) and (4). Prior distributions for all parameters in the likelihood are specified. The prior distributions for the parameter scales τ are positive Normal, which are weakly informative with a small scale

$$\tau^{-1} \sim N^{+}(0, 0.5) \tag{6}$$

and the prior distribution of the Cholesky factor for the correlation matrix Ω, L is distributed

$$LL' \sim \text{LKJCorr}(\eta) \tag{7}$$

where η is the shape parameter set to $\eta = 4$. The latter assumes that correlations between elements of β tends to be low. The prior distributions for the conditional mean of the marginal utility coefficients, β, and for γ are specified as follows:

$$\beta_{SQ} \sim N(0, s_0) \tag{8}$$

$$\beta_{SD}, \beta_T, \beta_C \sim N^{-}(b_0, s_0) \tag{9}$$

$$\gamma \sim N(0, g_0) \tag{10}$$

with hyperparameters $s_0 = 0.5$ and $g_0 = 1$, respectively. Thus the partworths of travel unreliability, time and cost are are assumed negative Normal to ensure correct signs of the resultant willingness to pay (WTP) estimates.
Estimation of the joint posterior distribution of all parameters is done using Markov Chain Monte Carlo Methods using the probabilistic programming language stan in R using the package RStan (Stan Development Team, 2021) with code based on Jim Savage's code (Savage, 2018) and adapted to allow for parameter restrictions. The model is run with four independent Markov chains, with 2,500 warmup iterations and 2,500 sampling iterations each. Convergence is checked via the split potential scale reduction factor (Rhat).

3 Results and discussion

This section presents the main results of this study. Figure 2 shows the answers to the questions on using quantile dotplots to present uncertain travel time. The majority of respondents agrees that quantile dotplots are understandable, helpful in the context of the choice experiment and potentially in real life. However, these results have to be interpreted cautiously as they capture *perceived* understanding and helpfulness only. Still, quantile dotplots appear to be a suitable format for presenting travel time distributions in SP experiments. Moreover, when designing their quantile dotplot presentation format, Kay et al. (2016) expressly considered small screens. This is relevant also for SP experiments, which are usually completed online and often on mobile phones. From a researcher's point of view, quantile dotplots are attractive because they convey the properties of the travel time distribution in terms of its measures of central tendency and spread, which can be directly included in the utility function, if the mean-dispersion model Small et al. (1999) is assumed.
Next, we focus on the VOR and VTT estimates. Their posterior distributions are obtained using all draws from the posterior distributions of the relevant marginal utility coefficients. Computation is done using $r = 1, ..., R$ draws from the posterior, where $R = 10,000$ and $\text{VOR}^r = \beta^r_{SD}/\beta^r_C$, $\text{VTT}^r = \beta^r_T/\beta^r_C$ and $\text{RR}^r = \text{VOR}^r/\text{VTT}^r$. These computed quantities also contain relevant

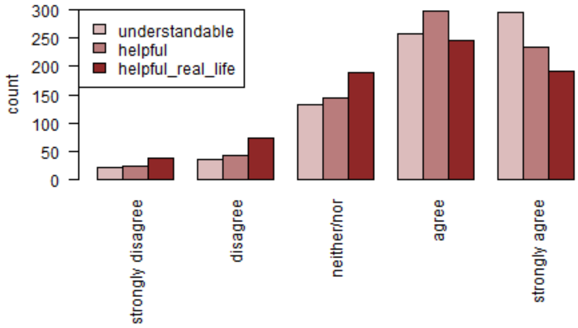


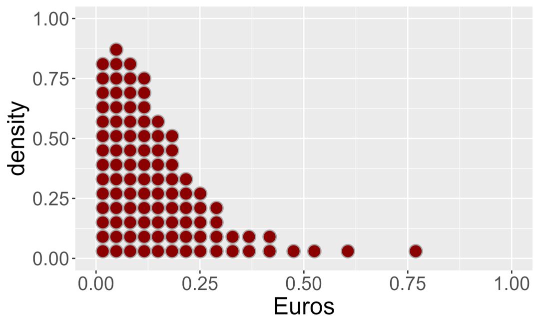
Figure 2: Frequency distribution of responses to statement on whether dotplots are understandable, helpful and helpful in real life when presenting travel time (n=741)

information on the marginal utility coefficients such covariance relationships and truncation to the negative space. The posterior mode in Table 3 is the most probable value and the 90% uncertainty interval boundaries indicate where the value will fall with 90% probability. VOR is the value of a one minute reduction of the standard deviation of the travel time distribution in Euros, and VTT is the amount of money travellers are willing to pay if travel time is reduced by one minute. The reliability ratio RR measures the marginal utility of a reduction of the standard deviation of the travel time distribution by one minute over the marginal utility of a reduction in average travel time by one minute. But because our aim is to highlight the importance of considering the entire posterior distribution of quantities of interest and therefore make use of all of the information that is available about the estimate, Figure 3 shows the entire posterior distributions of VOR^{SD} and VTT. Because they are based on fractions of marginal utility coefficients, which

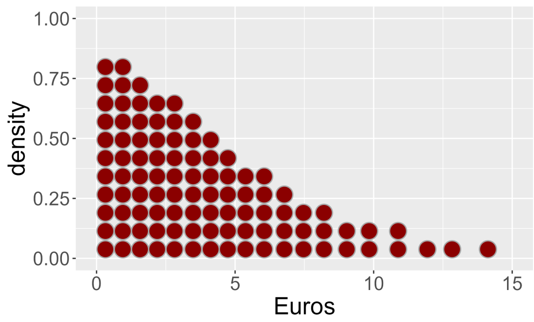
Table 3: Posterior modes, lower and upper bounds of 90% uncertainty intervals of marginal utility coefficients and their standard deviations, as well as of VOR, VTT and RR estimates.

Quantity	Mode	l-90% CI	u-90% CI
β_{car}	1.583	1.024	2.115
β_{pt}	-0.207	-0.499	0.116
β_{SD}	-0.246	-0.882	-0.058
β_{T}	-0.027	-0.029	-0.002
β_{C}	-0.110	-0.155	-0.052
τ_{car}	2.188	1.945	2.598
τ_{pt}	1.969	1.600	2.010
τ_{SD}	0.036	0.032	0.883
τ_{T}	0.013	0.002	0.037
τ_{C}	0.136	0.117	0.222
VOR	2.229	0.303	9.548
VTT	0.246	0.013	0.326
RR	9.052	2.285	159.372

are truncated to the negative space, they are both skewed. The VOR estimate varies widely with values up to 9.5 being possible. The reason for the large variation is partly explained by the fact that the posterior distribution of β_{SD} is wide as its variable, SD varies little (between 1.01 to 1.16 minutes) making it difficult for the estimator to estimate its effect on utility with precision. As a consequence, the VOR estimate varies widely. Also the RR estimates can take very large values because it is computed from a double fraction of the uncertain marginal utility coefficient estimates. This demonstrates how parameter uncertainty is propagated to quantities of interest and how it increases as more computational steps are involved. Something that is missed, if only point estimates are used in the computation. However, information about the full range of possible values is helpful to decision-makers. This is true in particular in situations where the tail of the distribution is important in order to avoid making costly decisions. In this application, if the



(a) Value of reducing the mean of the travel time distribution by one minute (VTT)



(b) Value of reducing the standard deviation of the travel time distribution by one (VOR)

Figure 3: Posterior distributions of VOR and VTT

costs of making small improvements to travel time reliability are low, it would still be advisable to make this investments because there is an - albeit small - chance that benefits could be high. On the other hand, if the costs of improving travel time reliability are high, even if there is small probability of achieving high benefits, it may be too risky to decide in favour of a large investment. This highlights the importance of quantifying uncertainty and providing all the information that is available about estimates of interest to support good decision-making.

4 Conclusions

This study demonstrates how a Bayesian approach can obtain value estimates from marginal utility coefficients with truncation and covariance relationships in a straightforward manner. Our results find that the VOR estimate is highly uncertain, as is the RR estimate. These uncertainties are completely missed, if only point estimates are used in the computation, but knowing the range of plausible values and not just the best estimate is important for decision-makers in particular when the costs of making a transport investment decisions are high. This study is first to use quantile dotplots for the representation of the travel time distribution in an SP experiment. We find that respondents perceive them as helpful and understandable, although a direct comparison with other representation formats and actual tests of respondents' understanding is beyond the scope this study, opening an interesting avenue for future research. We conclude that uncertainty is an important consideration both for the elicitation and the estimation of travel time and and travel time unreliability value estimates.

Acknowledgements

The authors gratefully acknowledge funding by the German Federal Ministry for Digital and Transport as part of the project RealLabHH.

References

Alonso-Gonzalez, M. J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S., & Hoogendoorn, S. (2020). Value of time and reliability for urban pooled on-demand services. *Transportation Research Part C: Emerging Technologies*, *115*, 102621. doi: https://doi.org/10.1016/j.trc.2020.102621

Bates, J., Polak, J., Jones, P., & Cook, A. (2001). The valuation of reliability for personal travel. *Transportation Research Part E: Logistics and Transportation Review*, *37*(2), 191-229. doi: https://doi.org/10.1016/S1366-5545(00)00011-9

Black, I., & Towriss, J. (1993). *Demand effects of travel time reliability.* Centre for Transport Studies, Cranfield Institute of Technology.

Carrion, C., & Levinson, D. (2012). Value of travel time reliability: A review of current evidence. *Transportation Research Part A: Policy and Practice*, *46*(4), 720-741. Retrieved from https://www.sciencedirect.com/science/article/pii/S0965856412000043 doi: https://doi.org/10.1016/j.tra.2012.01.003

Castro, S. C., Quinan, P. S., Hosseinpour, H., & Padilla, L. (2022). Examining effort in 1d uncertainty communication using individual differences in working memory and nasa-tlx. *IEEE transactions on visualization and computer graphics*, *28*(1), 411-421. doi: 10.1109/TVCG.2021.3114803

Copley, G., Murphy, P., & Pearce, D. (2002). *Understanding and valuing journey time variability* (Tech. Rep.). Cambridge. PTRC. (Proceedings of the European Transport Conference note)

Frans, N., Hummelen, B., Albers, C. J., & Paap, M. C. (2023). Visualizing uncertainty to promote clinicians' understanding of measurement error. *Assessment*, *0*(0), 10731911221147042. doi: 10.1177/10731911221147042

Jackson, B., & Jucker, J. (1982). An empirical study of travel time variability and travel choice behavior. *Transportation Science*, *16*(4), 460-475.

Kay, M., Kola, T., Hullman, J. R., & Munson, S. A. (2016). When (ish) is my bus? user-centered visualizations of uncertainty in everyday, mobile predictive systems. In *Proceedings of the 2016 chi conference on human factors in computing systems* (p. 5092–5103). New York, NY, USA: Association for Computing Machinery. Retrieved from https://doi.org/10.1145/2858036.2858558 doi: 10.1145/2858036.2858558

Li, Z., Hensher, D. A., & Rose, J. M. (2010). Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. *Transportation Research Part E: Logistics and Transportation Review*, *46*(3), 384-403. Retrieved from https://www.sciencedirect.com/science/article/pii/S1366554509001550 doi: https://doi.org/10.1016/j.tre.2009.12.005

Rose, J. M., & Bliemer, M. C. (2009). Constructing efficient stated choice experimental designs. *Transport Reviews*, *29*(5), 587–617.

Savage, J. (2018). *Combined conditional/mnl in stan.* (http://khakieconomics.github.io/2018/03/13/Combined-conditional-and-multinomial-logit-in-stan.html)

Small, K., R, N., Chu, X., & Lewis, D. (1999). *Valuation of travel time savings and predictability in congested conditions for highway user-cost estimation* (Tech. Rep. No. 431). Transportation Research Board.

Stan Development Team. (2021). *RStan: the R interface to Stan.* Retrieved from http://mc-stan.org/ (R package version 2.21.1)

Tilahun, N., & Levinson, D. (2010). A moment of time: Reliability in route choice using stated preference. *Journal of Intelligent Transportation Systems*, *14*(3), 179-187. doi: 10.1080/15472450.2010.484751

Wilkinson, L. (1999). Dot plots. *The American Stastistician*, *53*(3), 276-281.