Article:

Ehreke, I, Hess, S, Weis, C et al. (1 more author) (2015) Reliability in the German value of time study. Transportation Research Record, 2495 (1). pp. 14-22. ISSN 0361-1981

https://doi.org/10.3141/2495-02

Reliability in the German value of time study

Ilka Ehreke
ETH Zurich
Institute for Transport Planning and Systems (IVT)
Stefano-Franscini-Platz 5, CH-8093 Zurich
Tel: 0041-44-6333092; Fax: 0041-44-6331057; Email: ilka.ehreke@ivt.baug.ethz.ch

Stephane Hess
University of Leeds
Institute for Transport Studies (ITS)
34-40 University Road, University of Leeds, UK-Leeds LS2 9JT
Tel: 0044-113-3436611; Fax: 0113-3435334; Email: s.hess@its.leeds.ac.uk

Claude Weis
Transoptima GmbH
Technoparkstrasse 1, CH-8005 Zurich
Tel: 0041-44-2711163; Fax: 0041-76-3493139; Email: weis@transoptima.ch

Kay W. Axhausen
ETH Zurich
Institute for Transport Planning and Systems (IVT)
Stefano-Franscini-Platz 5, CH-8093 Zurich
Tel: 0041-44-6333943; Fax: 0041-44-6331057, Email: axhausen@ivt.baug.ethz.ch

Words: 5901 words + 2 tables + 4 figures = 7401 word equivalents

TRR Paper number: 15-2265

Submission Date 13/3/2015

ABSTRACT

The German Federal Ministry of Transport and Digital Infrastructure is currently preparing the 2015 Federal Transport Investment Plan. As part of this, it is updating the overall methodology of its cost-benefit analysis meaning values of both reliability (VOR) and travel time (VOT) for personal and business travel will be estimated. While the VOTs will replace a set of existing values, the VOR will be estimated for the first time as they are not incorporated in the standard appraisal yet. The data collection adopted a two-stage approach: first respondents reported about current trips (revealed preference), which were then systematically varied to be the basis for stated preference experiments. This paper presents the findings of estimating the VOR. In the SP experiments the reliability of the travel modes was presented with different formats. The final model formulation differs in the definition of reliability for private and public transport. For car trips, saving travel time is “worthier” to the respondents than reducing the variability. The calculated VOR for the mean expected unscheduled delay of public transport trips are slightly lower than the VOTs which means that the reliability is here less important to the respondents than relevant travel time saving. One minute of mean expected unscheduled delay and one minute of standard deviation are almost equivalent to one minute of travel time saving (reliability ratio). As this has been the first official estimation of the value of reliability and time for Germany, the values should be reconsidered and updated on a regular basis.

INTRODUCTION AND RELATED WORK

The effects of hundreds of infrastructure projects on transport policies and investments are to be evaluated with cost benefit analyses (CBA) in Germany. The German Federal Ministry of Transport and Digital Infrastructure (BMVI) is currently preparing the 2015 Federal Transport Investment Plan (Bundesverkehrswegeplan, BVWP) its medium to long-term investment strategy for the country's transport infrastructure serving longer distance travel. As part of this, it is updating and modernizing the overall methodology of its central evaluation tool: the cost-benefit analysis. One on-going project is focusing on the CBA as such, and a second is estimating and recommending values of reliability (VOR) and travel time (VOT) for personal and business travel. While the new VOTs will replace a set of existing values which were based on values from BVWP'92 and have not been verified independently since then, the VOR will be estimated for the first time as they are not yet incorporated in the standard appraisal yet. The aim of integrating reliability into the new BVWP is in line with practice and science to make transport systems not only faster but also more reliable (1) Another BMVI initiated ongoing project will calculate VOTs and VORs for freight (2) but this is not subject of the presented research.

Infrastructure projects evaluated with CBA and transport policies not only influence the mean travel time but also its distribution (3). The frequency of congestion and public transport schedule unreliability are growing in parallel with increasing transport demand. Travel time reliability has a significant impact on transport behavior (4) so should be understood and hence predicted together with travel behavior in project evaluation and demand forecasting.

International studies on the evaluation of value of time and travel time reliability are based on stated preference (SP) experiments, which estimate peoples' willingness to pay (WTP) for avoiding late or early arrival. (5) found, that avoiding delay is almost twice as important as arriving early and distinguish between two main impacts – a predictable and an unpredictable one - of unreliable travel time on travel behavior: as the average demand profile during the course of a day and a random incident. Numerous empirical examples of VOR determination are documented in recent research. For example for commuters in Barcelona, a willingness to pay for avoiding delayed arrival at work, which is up to three times higher than the VOT, is documented in (6) and (7). Beside empirical evidence, the presentation of reliability in the questionnaire is also an important element of various studies (8, 9, 10, 11, 12, 13, 14). Nevertheless, it is still a relatively new topic in the field of travel behavior research and therefore differences in study design (for example revealed ¨preference (RP) or stated preference (SP) experiments) or theoretical framework can lead to a wide range of reliability valuations compared to pure travel time (15). With a meta-analysis (16) show the methodological development of VOR determination during the recent decades. Like (15), they describe the differences between the two most common theoretical models of travel time reliability: the mean variance and the scheduling approach. (16) conclude that the ideal display of reliability is as proposed in (17). In (18), a model approach, which considers the so called probability weighting and additionally allows non-linear utility estimation for different influencing factors (especially for risk valuation), is presented. A more detailed review on international literature on VOR research can be found in (19) and (3).

This paper presents the findings estimating the value of travel time reliability in the German VOT study (19). It provides an overview of the study design, response rate and presentation of reliability in the different experiments. Several definitions of reliability also regarding compatibility to other projects in the BVWP 2015 framework were tested before the final theoretical model was determined. The paper concludes with the estimated values, the reliability ratios and an outlook on future research.

SURVEY DESIGN AND RESPONSE

Microeconomic models of time allocation have been used to derive the valuations of technologically constrained time use since the early work of (20), (21), (22), where the main focus is on the value of travel time. The current state of practice draws largely upon past British, Dutch and Scandinavian studies (23, 24) which over time moved from RP data to a growing reliance on personalized SP experiments to estimate the VOT and VOR by using suitably formulated discrete choice models of travel behavior, especially of route and mode choices.

Study Design

The design of the present study builds on the experience of studies in Switzerland (25, 26, 27), which had further developed international practice by employing more complex stated choice experiments including multiple modes and multiple elements of the generalized costs of travel in a series of overlapping choice contexts (route, mode, departure-time-mode, departure time-route). The BVWP requires estimates for both private and business travel. As business travel is concentrated in a small share of the population, it was necessary to recruit a complementary sample of such travelers in addition to a population-based sample to obtain a sample large enough. Business travel was defined as all employment related travel, but excluding emergency services and driving as work (delivery, bus and coach drivers etc.). It includes various kinds of business travel from local craftsmen as well as lawyers and consultants.

The population-based sample was drawn from a dual frame of land line and mobile numbers (60% and 40%) to ensure that the growing share of mobile-only users are included. It was incrementally controlled over the survey period so as to ensure the spatial quotas in terms of the German federal states. This is based on an ITMS-Double frame and recommended by recent German market research studies (28). The additional sample of business travelers was recruited online with a TNS-Infratest access panel and these respondents received a small incentive for their participation.

In line with international practice, the study first collected data on recent trips performed by the respondents, then added the information about the non-chosen options, selected the reference trip and constructed the SP experiments around it. The RP trips are obtained as the 4 trips to the workplace and most important shopping and leisure destinations and the last long-distance trip over 100 km distance, where, if the latter was ground-based, data on the most recent air trip was also collected. The rationale behind the approach is based on the observation that the bulk of a person's everyday travel is to a very small number of destinations (29), so that a good range within a relatively short computer assisted telephone interview (CATI) can be obtained. The CATI made it possible to geocode the destinations and the route using Trip Tracer (30). This trip information was complemented with the usual socio-demographic information and information about the respondent's mobility tools. The business trip sample responded via a web-based survey system. Both samples received the SP experiments within a maximum of two weeks after participation in the CATI or the on-line version in paper-and-pencil form or again as a web-based survey. The non-chosen alternatives and their attributes were based on the information from a number of resources: door-to-door travel times from a MATSim (31) implementation based on an average link travel time reported by TomTom Stats® network for Germany. The travel times on public transport were obtained from relevant websites.

A focus trip was randomly assigned to the respondents. The origin and destination and the exact route of the focus trip were determined, as well as detailed information on their residence and workplace. Secondly, the stated choice questionnaires were constructed based on the focus trips from the RP experiments. Respondents received at least two SP game sets of either mode choice, route choice or reliability and a randomly assigned long term experiment (residential or workplace choice).

Participants in the business travel survey were preselected online and afterwards interviewed in a CATI. Each of the respondents reported their last three business trips from which the most recent one became the focus trip. Based on that trip, the SP experiments were conducted as described for the non-business survey but without a long-term experiment. The SP questionnaires for business trips were completed online. The process of the study is shown in the following FIGURE 1.

Business travelers are sometimes not free to choose the mode or even the route of their travel due to company policy, and this would invalidate the SP experiments with them. This was checked beforehand by conducting a small-scale qualitative interview survey. The 24 participating decision makers had been recruited to cover the main regions of Germany as well as the range of firm sizes. While many firms indeed had policies in place, the sample reported that their employees were free to choose their routes and in the vast majority also the mode of travel. The small share allowed to go ahead with the SP experiments without having to fear a major bias in the results.

It was also tested whether the size of the travel time differences offered to the respondents in the SP experiments had an impact on the valuations. After accounting for the other non-linearities, the models could not identify such size-effects. The empirical literature on short-term changes in travel behavior shows that small travel time changes are often ignored or not perceived by the travelers. Still, in the long-term logic of Cost-Benefit Analyses (CBA) this effect is irrelevant. To account for the effect would be inconsistent with assumptions of it and would open the chance to manipulate its results through dividing or aggregating projects into smaller or larger units. So it is recommend to follow international practice and to value all savings equally (19).

The reference trip was selected with a bias to the longer trips, given their rarity and the interest of the BVWP in intercity travel. This selection was corrected in the analysis through a reweighting to match the distance-purpose distribution observed in the most recent German national travel diary survey, Mobility in Germany 2008 (32). The SP survey asked respondents to choose between several different trip alternatives and travel costs. A total of four short-term SP experiments were performed: one for transport mode choice, one for automobile route selection, one for public transport without transferring to rail, and one for public transport including a transfer to rail.

Respondents received three different SP experiments in the general case, and two, if the reference trip was a business trip. In total, they were offered between 16 and 24 choice situations. The modes offered were walking, cycling, car, local public transport (PT) and the various long distance public transport modes (train, air and the newly deregulated coach option). The SP questionnaires also included attitudinal questions. The decision-making situations were developed for the specific individual based on path information added to their RP trips and routes. The SP survey provided: scheduled arrival time, travel time, travel costs, and, travel time reliability.

The reliability experiment was formulated as route-departure time choice with an expressed indication of travel time variability. Three formats of different complexity were tested, but each allowing to estimate the *mean-variance model of scheduling*. FIGURE 2 shows the three different presentation types of reliability using the example of public transport whereas each column represents one type of reliability experiment.

All three formats were retained since the pre-test indicated no clear preference between them, in spite of their growing complexity. The travel time reliability was varied by providing different congestion probability and average congestion times (delay) for automobile travel and by providing the probability of delays (in minutes) from scheduled arrival time for public transport travel (delays were a percentage of the specified tolerance from the RP survey). It was also possible to estimate a common model with non-linear variables (interaction terms based on distance or income). The model results are plausible and robust. In short, the survey successfully achieved its research objective, testing the difference between scheduled and delayed/early arrivals.

Response

The two-step survey was carried out in seven waves from July to September 2012 (after a pre-test in May). Before the beginning of the first wave the response rates for the paper-pencil and online business and non-business studies were predicted following (33). All three actual rates were in the expected range. The response behavior was better than the earlier Swiss experiences (33).
A recruitment rate of over 30% for the CATI interview and 75% completion rate for the first phases and response rates of 73% and 93% for the second phases in spite of the complexity of the instruments are indicative of the interest in the topic.
As described above in this research, a two-stage method was used to collect the necessary modeling data. In the RP survey almost 4,000 people completed the questionnaire providing socio-demographic characteristics and information on recent trips. The SP-surveys were controlled so that there was a sufficiently large sample of responses for all trip purposes.
TABLE 1 gives an overview about the distribution of the number of fully completed surveys by sort of the experiment and study. It shows that the response rate of the business study is overall higher than in the non-business study as participants were recruited in a business market research online panel. The required number of participants in the business study was already reached after wave six. Including the pretest data about 2,300 non-business and 790 business respondents completed the full questionnaire. The sample contains almost 64,500 choices in the different SP experiments. Besides experience from the pre-test, the main study confirmed that all three types of reliability presentation obtained equally high response rates (see also TABLE 1). Between the presentation types no clear pattern is recognizable. In the written paper-pencil non business survey, the reliability presentation type 2 got the most responses whereas respondents in the non-business online survey responded best to type three presentation of reliability. Type one turned out to gain most responses in the online business survey, whereas in total the difference between type 3 and type 1 is about 7%.
Non-traders in a SP survey always choose the same alternative across their choice sets regardless of the available alternatives' attributes. As it does not necessarily imply inconsistent responses they were included in the modeling process. In total, 34% of the respondents never varied their choices in the mode choice experiments. Another important issues for a survey is item non-response, which means that respondents do not answer a particular unit among the questions. The German VOT study showed only minor problems with item non-response, most of the shares of missing values were less than 3%. A more detailed analysis of the study design and response behavior can be found in (34).

METHOD

Individual models for each experiment were estimated but only the pooled results across all of the short-term SP experiments are reported, where the joint estimation was made possible by the presence of joint variables. Differences across experiments in terms of the relative influence of the unobserved utility components were accounted for by the estimation of experiment specific scale parameters. These vary between 1.5 and 3.3 with the mode choice experiment being the base indicating that the other experiments generated choices that are more deterministic.

It is well known that the VOT might change with distance or travel time. These non-linearities were tested with a set of formulations, including the elasticity continuous interaction terms suggested by (35) and various non-linear attribute specific transforms, ranging from simple log-transforms to the Tangens-Hyperbolicus. In the end, the best results were obtained with the following formulation:

$$U_i = \sum_j \dots \left(\beta_{i,j} \cdot x_{i,j} + \alpha_{i,j} \cdot ln\left(x_{i,j} + \gamma_{i,j}\right)\right) \cdot \left(\frac{z_j}{\mu(z_j)}\right)^{\lambda_{i,j,z_j}} \dots \quad (1)$$

U_i	Utility of alternative
i	Alternatives i = 1, ... n
$x_{i,j}$	Attribute *j* of alternative *i*
$(\beta,\alpha,\gamma)_{i,j}$	Parameters associated with $x_{i,j}$
$\lambda_{i,j}, z_j$	Elasticity of the sensitivity to attribute *j* for alternative *i* with respect to attribute z_j
$\mu(z_j)$	Mean of attribute z_j

The continuous interaction terms varied across attributes. In particular, income indexed as z_j for travel time and cost was used. This was divided by the sample mean value to normalize the values. For the other attributes, travel time as z_j was used, allowing sensitivities to change depending on travel time. Different specifications were tested for reliability, where the final specification used the variance of the travel times for private transport. For attribute specific non-linearity, a combined linear and logarithmic approach was used, with the additional positive offset term $\gamma_{(i,j)}$ to handle attribute values close to zero.

Value of reliability

Previous to the presented one, a feasibility study by (3) analyzed the possibility of a prospective integration of reliability in the BVWP's CBA. The report included an extended literature review as well as expert interviews with practitioners and researchers on the definition of travel time reliability for Germany. The findings of the study form the basis of the reliability definition in the German VOT study and are presented below. From a practical transport modeling point of view, the variance of travel time is easier to integrate and could even be used as an approximation for scheduling effects (36). Furthermore, the most recent German transport model for the BVWP does not include even a partial departure time model so it would be rather difficult to include scheduling variables (3).

The methodology is explained with the example of delay but the same holds for early arrival. Two kind of reliability definitions were specified for the estimation of the data. One is reliability defined as standard deviation of the travel time distribution. An unplanned delay is expressed through the standard deviation of the arithmetic mean, which implies that a decrease of the mean of the travel time distribution stands for the travel time savings and a decrease of the standard deviation can be interpreted as reliability. This is the *mean variance approach* (16). Especially for unscheduled transport modes like the car, this understanding serves well for reliability because car drivers not only consider their planned mean travel time but also the driver's sense of unreliability is taken into account (37). The VOR following this approach can be determined as follows:

$$VOR = \frac{dU/_{dstdev}}{dU/_{dcost}} \quad (2)$$

The partial derivative of the standard deviation can be written as:

$$\frac{dU}{dstdev} = \beta_{i,j} * \text{stdev}_{i,j} * \left(\frac{x_{i,j}}{60}\right)^{\lambda_{stdev_{i,j}}} \quad (3)$$

Whereas the partial derivative of the cost can be written as:

$$\frac{dU}{dcost} = \left(\beta_{i,j} + \frac{\alpha_{i,j}}{x_{i,j}+\gamma_{i,j}}\right) * \left(\frac{income_{i,j}}{\mu(income)}\right)^{\lambda_{income_{i,j}}} \quad (4)$$

Second, reliability is defined for modeling as the mean expected unscheduled delay. The probability of a late arrival is multiplied with the average delay on those journeys that are delayed. If more than one delay and the probability of it occur in the SP experiments, the average of those is taken. The VOR is as the arithmetic population weighted mean of all calculated mean unexpected delay values. For example a delay of 5 minutes in 15% of the trips and a 10 minute delay in 5 % of all trips leads to a delay of 1.25 minute or 1 minute and 25 seconds. This method serves for scheduled transport modes (public transport and flight). Respondents can react on the reliability in different ways: they can adjust their departure time, their route or change their mode of transport. The partial derivation of the mean expected unscheduled delay can be written as:

$$\frac{dU}{dmean\_expected\_delay} = \beta_{i,j} * \left(\frac{p_{delay_{i,j}} * \bar{x}_{delay_{i,j}}}{100}\right) \quad (5)$$

$p_{delay_{i,j}}$	Probability of delay
$\bar{x}_{delay_{i,j}}$	Average of delay

The determination of the VOR remains the same as well as the partial derivative of cost. Again the same holds for early arrival.

RESULTS

The values of reliability were derived as described above in the same manner as the VOT. Models are available to present the VOR as standard deviation and expected delay. TABLE 3 shows values of time and the values of reliability by trip purpose for a joint model with standard deviation for car and mean unexpected delay respectively early arrival for public transport and air. The parameters for expected early arrival were partially non-significant and not recommend for usage. The increasing availability of smart phones, tablets and other devices has made early arrival better usable in just about all situations. Especially the air high values due to not significant parameters seem not to be plausible. However they are shown for completeness.

For example a car driver on her or his way to work is willing to pay 3.45 €/h in reliability. The VOT for the way to work driving by car is 4.87 €/h (19). That means that saving travel time is “worthier” to the respondents than reducing its variability. For example, a trip of 30 minutes free flow travel time which, due to congestion, averages to 1 hour travel time and 5 minutes standard deviation. That means that 65% of all trips take between 55 and 65 minutes and 95% of all trips between 45 and 75 minutes. The following changes would be seen as the equivalent:

- Reduction of the standard deviation to 2.5 minutes which would mean that 95% of all trips take between 52.5 and 67.5 minutes

- Reduction of the average travel time to 58 minutes and 5 seconds with constant standard deviation which would mean that 95% of all trips take between 43 minutes and 5 seconds and 73 minutes and 5 seconds
- The WTP for the mean expected unscheduled delay means that the respondents are willing to pay for the reduction of the mean expected unscheduled delay and thereby increase the reliability. For example a value of 5.10 €/h means that the willingness to pay for a reduced delay which occurs in half of the trips and takes exactly one hour (so to say 30 minutes) is 2.55 €/h. The calculated VORs are slightly lower and in some cases higher than the VOTs (19), which means that the reliability is less important to the respondents than transit travel time saving except for business trips. To that effect it seems plausible that for business travel, arriving on time is essential. The lower value for shopping can be interpreted as that the reliability is less important to the respondents because shopping trips are in general not dependent on an on-time arrival at the destination. On the other hand the relatively low acceptance of a delay (FIGURE 1) is in contrast to that. Flight reliability is clearly more important than travel time saving to the respondents. However, the flight parameters were estimated with a high variation and are not significantly different from zero (19).
- Arriving early at a destination, as expected, seems to be less important to the respondents than arriving late or saving travel time. Again, the values for air are not plausible.

Beside the valuation of reliability it is interesting to know about people's tolerance for a delayed arrival at a destination. In the RP questionnaire the respondents were asked to state what they considered a noticeable delay in minutes for their focus trip. As respondents were randomly assigned with a focus trip it was possible to compare the acceptable delay of the different trip purposes. The cumulative shares of acceptable delay in minutes are shown in FIGURE 3.
As expected, the tolerance of being late for a business trip is very low. Surprisingly the acceptance of a delay when commuting is the highest. This may be due to more flexible working hours nowadays and can also be seen in the lower tolerance for educational commuting as education schedules are more bounded. Also as expected, the acceptance for a delayed leisure trip is quite high. All together, most of the tolerated delays are between five and 30 minutes and the steps in five minute intervals are also clearly visible.

Reliability ratio

The exchange rate of the value of reliability and the value of time is also called reliability ratio and can be computed as follows:

$$RR = \frac{\delta}{\gamma} = \frac{VOR}{VOT} \tag{6}$$

TABLE 3, which shows the reliability ratios calculated from the VOR and VOT of the presented study can be interpreted in the following way: One minute of mean expected unscheduled delay for commute public transport trips is almost equal to one minute of travel time saving (reliability ratio). This one minute of average delay can represent a delay of two minutes in 50% of the trips or a delay of 4 minutes in 25% of the trips, as well as a combination such as one minute in 50% of the trips and 2 minutes in 25% of the trips. For most of the trip purposes the equivalent valuation of the mean expected unscheduled delay is almost equal to one minute travel time saving or even less. Only the ratio for business trips is a bit higher. This result is somewhat unexpected and thus could be explained by respondents undervaluing the probability of the occurrence of undesirable events or even ignoring them as can be seen in other risk situations.

The interpretation of the reliability ratio of the standard deviation is similar: one minute standard deviation corresponds to 0.7 minutes of travel time saving. However, it means something else, since for public transport and air, a reduction of the value of the unreliability also causes a reduction of the average travel time, but for car driving the reduction of the standard deviation doesn't cause a reduction of the average travel time.

FIGURE 4 shows the socio-demographic characteristics of the respondents before the re-weighting. As seen in previous studies on travel behavior (for example 27) the share of highly educated, older participants with high income predominates the socio-economic distribution of participants in this study. Only the absolute share of male respondents is higher than usual in this study due to the higher share of 68 % of male participants in the business sample. In the second part of FIGURE 4 the calculated reliability ratios by mode of transport for the socio-demographic indicators are presented. It is obvious that the ratios are more or less evenly distributed among the characteristics for the single modes. Only fulltime employees value reliability twice as much when going by airplane. In contrast to findings of other studies (37) the values (and reliability ratios) for male and female participants do not differ much from each other. In the presented study men value reliability even slightly higher than women, at least with the used reliability definition.

CONCLUSION AND OUTLOOK

The German VOT study worked with different formats to present the reliability of the travel modes in the stated choice experiments. The final model formulation differs in the definition of reliability for private and public transport. This ultimately unsatisfactory situation arose from the different methods of the evaluation of transport policy reliability effects in the official transport models. A uniform procedure seems desirable both in the presentation of reliability in SP experiments and in the transport policy evaluation, as well as in the ongoing observation of the traffic situation. Likewise the parameter for air and early arrival do not differ significantly from zero, so the calculated values do not seem plausible in every point. A different model formulation for estimating those parameters should be tested.

There is less empirical evidence on the value of reliability than on the value of travel time savings in the international context, as the systematically estimation of the VOR only recently started. In particular the formulation of the mean expected unscheduled delay is rather difficult to compare to other values. The reliability ratio for the standard deviation for car trips and for public transport trips are in the range of the international values reported by (38) even though they are low in comparison. The same holds for the VOTs in the international comparison and deserves a more detailed investigation in the future.

However, the value of reliability is of great importance for future research, even with the lack of one single accepted formulation of reliability and a common presentation in SP experiments. Nevertheless, this has been the first official estimation of the value of reliability and the value of time for Germany. Now the values should be reconsidered and updated on a regular basis, ideally with every new or at least every second BVWP.

ACKNOWLEDGEMENTS

This research was funded by the German Federal Ministry of Transport and Digital Infrastructure as part of the project 96.0996/2011 within the framework of "Bundesverkehrswegeplan 2015“ and is a collaboration of IVT, ETH Zurich and TNS Infratest with the help of Prof. Stephane Hess and

Prof. Kai Nagel. We are grateful for the careful management of the project and substantial advice of Mrs. J, Monse and Mr. H. Hassheider of the BMVI, Berlin.

REFERENCES

1. BMVI, *Grundkonzeption für den Bundesverkehrswegeplan 2015*, Berlin, 2014.
2. BVU and TNS Infratest. *Entwicklung eines Modells zur Berechnung von modalen Verlagerungen im Güterverkehr für die Ableitung konsistenter Bewertungsansätze für die Bundesverkehrswegeplanung,* Draft final report, 96.1002/2012, Bundesministrium für Verkehr und digitale Infrastruktur, 2014.
3. Significance; Goudappel Coffeng; NEA. *Erfassung des Indikators Zuverlässigkeit des Verkehrsablaufs im Bewertungsverfahren der Bundesverkehrswegeplanung*. Final report to the BMVBS, 2012.
4. König, A. and K.W. Axhausen. Verlässlichkeit als Entscheidungsvariable. Final report for SVI 44/00. Swiss Federal Institute of Technology, Institute for Transport Planning and Systems, *Arbeitsberichte Verkehrs- und Raumplanung*, 110, 2002.
5. Bates, J.J., P. Polak, A. Jones und A. Cook. The valuation of reliability for personal travel, *Transportation Research E*, 37 (2), 2001, pp. 191-229.
6. Asensio, J. and A. Matas. An empirical estimation of the value of travel time reliability for commuters in Barcelona, paper presented at the *European Transport Conference*, Strasbourg, September 2006.
7. Asensio, J. and A. Matas. Commuters' valuation of travel time variability. *Transportation Research*, 44 (6), 2008, pp. 1074-1085.
8. De Jong, G., E. Kroes, R. Plasmeijer, P. Sanders, and P. Warffemius. The value of reliability. *Proceedings of the European Transport Conference*, 2004.
9. Hollander, Y. The attitudes of bus users to travel time reliability, paper presented at the *European Transport Conference*, Strasbourg, October 2005.
10. De Palma, A. and N. Picard. Route choice decision under travel time uncertainty. *Transportation Research*, 39 (4), 2005, pp. 295-324.
11. Bhat, C.R. and R. Sardesai. The impact of stop-making and travel time reliability on commute mode choice, *Transportation Research*, 40 (9), 2006, pp. 709-730.
12. Brownstone, D. and K.A. Small. Valuing time and reliability: Assessing the evidence from road pricing demonstrations, *Transportation Research*, 39 (4), 2005, pp. 279-293
13. Tseng, Y.-Y., E. Verhoef, G. de Jong, M. Kouwenhoven and T. van der Hoorn. A pilot study into the perception of unreliability of travel time using in-depth interview, *Journal of Choice Modelling*, 2 (1), 2009, pp. 8-28.
14. Fosgerau, M., K. Hjorth, C. Brems and D. Fukuda. *Travel time variability: Definition and valuation*, DTU Transport, 2008.
15. Carrion, C. and D. Levinson. Value of travel time reliability: A review of current evidence. *Transportation Research*, 46 (4), 2012, pp. 720-741.
16. Li, Z., D.A. Hensher and J.M. Rose. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence, *Transportation Research*, 46 (3), 2010, pp. 384-403.
17. Small, K.A., R.B. Noland, X. Chu and D. Lewis. *Valuation of travel-time savings and predictability in congested conditions for highway user-cost estimation*, NCHRP-Report, 431, Transportation Research Board, National Research Council, Washington, D.C, 1999.
18. Hensher, D.A. Hypothetical bias, choice experiments and willingness to pay, *Transportation Research*, 44 (6), 2010, pp. 735-752.

19. Axhausen, K.W., I. Ehreke, A. Glemser, S. Hess, C. Joedden, K. Nagel, A. Sauer and C. Weis. *Ermittlung von Bewertungsansätzen für Reisezeiten und Zuverlässigkeit auf der Basis eines Modells für modale Verlagerungen im nicht-gewerblichen und gewerblichen Personenverkehr für die Bundesverkehrswegeplanung.* Final report, BMVI, 2014.
20. Becker, G. S. A theory of the allocation of time, *Economic Journal*, 75 (299), 1965, pp. 493–517.
21. Beesley, M.E. The value of time spent in travelling: Some new evidence, *Economica*, 32 (126), 1965, pp. 174-185.
22. De Serpa, A.C. A theory of the economics of time, *Economic Journal*, 81 (324), 1971, pp. 828-846.
23. Wardman, M. The value of travel time: A Review of British evidence, *Journal of Transportation Economics and Policy*, 32 (3), 1998, pp. 285-316.
24. Abrantes, P.A.L. and M.R. Wardman. Meta-analysis of UK values of travel time: An update, *Transportation Research*, 45 (1), 2011, pp. 1-17.
25. Axhausen, K.W., A. König, G. Abay, J.J. Bates and M. Bierlaire. Swiss value of travel time savings, paper presented at the *European Transport Conference*, Strasbourg, October 2004.
26. Axhausen, K.W., S. Hess, A. König, G. Abay, J.J. Bates and M. Bierlaire. Income and distance elasticities of values of travel time savings: New Swiss results, *Transport Policy*, 15 (3), 2008, pp. 173–185.
27. Fröhlich, P., K.W. Axhausen, M. Vrtic, C. Weis and A. Erath, *SP-Befragung 2010 zum Verkehrsverhalten im Personenverkehr*, final report for the Bundesamt für Raumentwicklung, IG Modus, Zürich and IVT, ETH Zürich. Bern, 2012.
28. ADM – Arbeitskreis *Deutscher Markt- und Sozialforschngsinstitute. Dual-Frame-Ansätze 2011/2012.* http://www.adm-ev.de/fileadmin/user_upload/PDFS/ADM_Dual_Frame_Projekt_-_Forschungsbericht.pdf. Accessed, Jul. 31, 2014.
29. Schönfelder, S. and K.W. Axhausen. *Urban Rhythms and Travel Behavior: Spatial and temporal phenomena of daily travel.* Ashgate, Farnham, 2010.
30. DDS GmbH. TRIP TRACER – *Exakte Erfassung von Wegstrecken im Telefoninterview*, Software, http://www.ddsgeo.com/products/trip-tracer.html. Accessed, Jul. 31, 2014
31. MATSim. www.matsim.org. Accessed, Jul. 31, 2014.
32. Follmer, R., D. Gruschwitz, B. Jesske, S. Quandt, B. Lenz, C. Nobis, K. Köhler and M. Mehlin. MID 2008: Mobilität in Deutschland 2008: Struktur – Aufkommen – Emissionen – Trends, Final report to the BMVBS, Infras and DLR, Berlin and Berlin, 2010.
33. Axhausen, K.W. und C. Weis. Predicting response rate: A natural experiment, Survey Practice, 3 (2), 2010.
34. Ehreke, I., Axhausen, K.W and Claude Weis. Experience from the German value of time (VOT) and value of reliability (VOR) study, paper presented at the *10th International Conference on Transport Survey Methods*, Leura, November 2014.
35. Mackie, P.J., M. Wardman, A.S. Fowkes, G. Whelan, J. Nellthorp and J.J. Bates. Values of Travel Time Savings in the UK, report to Department for Transport, Institute for Transport Studies, University of Leeds and John Bates Services, Leeds, 2003.
36. Fosgerau, M. and A. Karlström, A. The value of reliability, Transportation Research B, 44 (1), 38-49, 2010.
37. Lam, T. C., and K. A. Small, The value of time and reliability: measurement from a value pricing experiment, Transportation Research E, 37(2), 231-251, 2001.

38. Significance, VU University Amsterdam and & John Bates Services. Values of Time and Reliability in Passenger and Freight Transport in the Netherlands. Report to the Ministry of Infrastructure and the Environment, unpublished, 2012.

LIST OF TABLES

LIST OF FIGURES

TABLE 1: Response behavior of the samples in the main study

General sample	**Business travel sample**
	online recruitment
	recruited online: 1,112
RP Sample	**RP Sample**
9,491 contacts	1,112 contacts
3,151 completed CATI (33.2 %)	848 completed CATI (76.3 %)
thereof: 2,965 indicated willingness to participate written SP experiment 186 indicated willingness to participate online SP experiment	848 indicated willingness to participate online SP experiment
SP-Experiments	**SP-Experiments**
2,285 completed (72.5 %) thereof: 2,187 completed written (73.8 %) 98 completed online (52.7 %)	786 completed online (92.7 %)
SP-Experiment by SP type	**SP-Experiment by SP type**
mode choice: 1,631 (67.6 %) route choice: 748 (71.4 %) reliability: 1,938 (68.3 %) workplace: 1,225 (70.6 %) residence: 1,159 (66.9 %)	mode choice: 431 (90.8 %) route choice: 408 (91.1 %) reliability: 839 (90.9 %)
Reliability type (written/online)	**Reliability**
type 1: 66.8 % (67.8 % / 52.4 %) type 2: 72.9 % (74.0 % / 57.8 %) type 3: 67.3 % (67.7 % / 60.6 %)	type 1: 91.4 % type 2: 86.8 % type 3: 91.1 %

TABLE 2: Value of reliability (VOR) value of time (VOT) [€/h] and reliability ratio (RR) (population weighted)

Mode	Attribute	Purpose					
		Edu cation	Work	Shop ping	Leisure	Business travel	All
Car	Std. deviation	3.21	3.45	3.51	3.09	6.54	3.61
Car	Value of time	3.90	4.87	4.29	4.03	8.38	4.66
Car	Std.dev/VOT	0.7	0.7	0.7	0.7	0.7	0.7
PT	Mean expected unscheduled delay	4.66	5.10	4.28	4.82	15.97	5.48
PT	Mean expected unscheduled early arrival	1.81	1.98	1.67	1.88	6.22	2.13
PT	Value of time	4.39	4.47	5.11	4.35	7.01	4.83
PT	VOR_late/VOT	0.9	1.0	0.7	0.9	1.7	0.9
PT	VOR_early/VOT	0.3	0.4	0.3	0.3	0.7	0.4
Air	Mean expected unscheduled delay	--	--	--	38.44	51.27	46.60
Air	Mean expected unscheduled early arrival	--	--	--	90.16	120.25	109.30
Air	Value of time	--	--	--	25.45	38.76	33.67
Air	VOR_late/VOT	--	--	--	1.4	1.4	1.4
Air	VOR_early/VOT	--	--	--	3.3	3.2	3.2
-- Not apllicable							

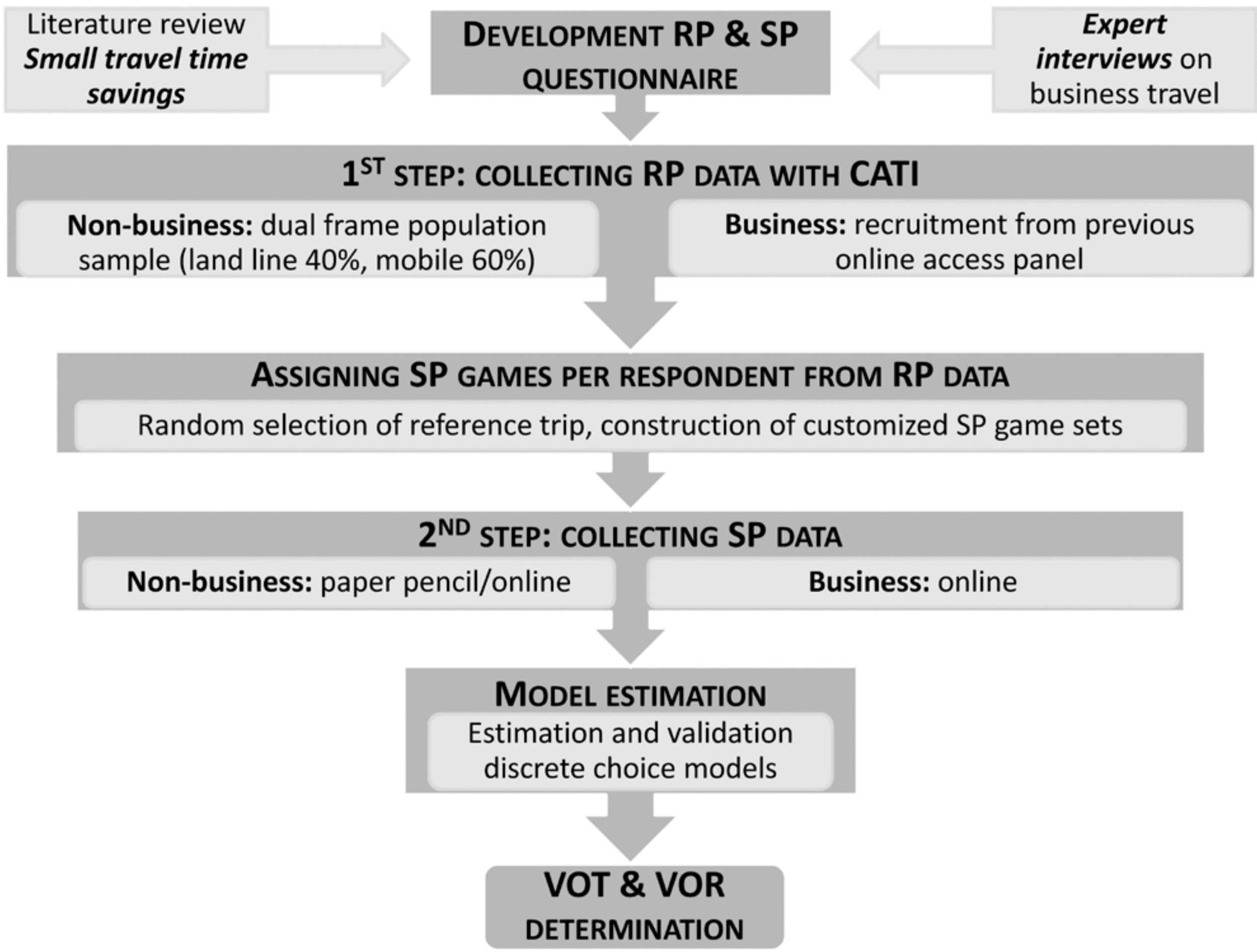

FIGURE 1: Process of the German VOT study

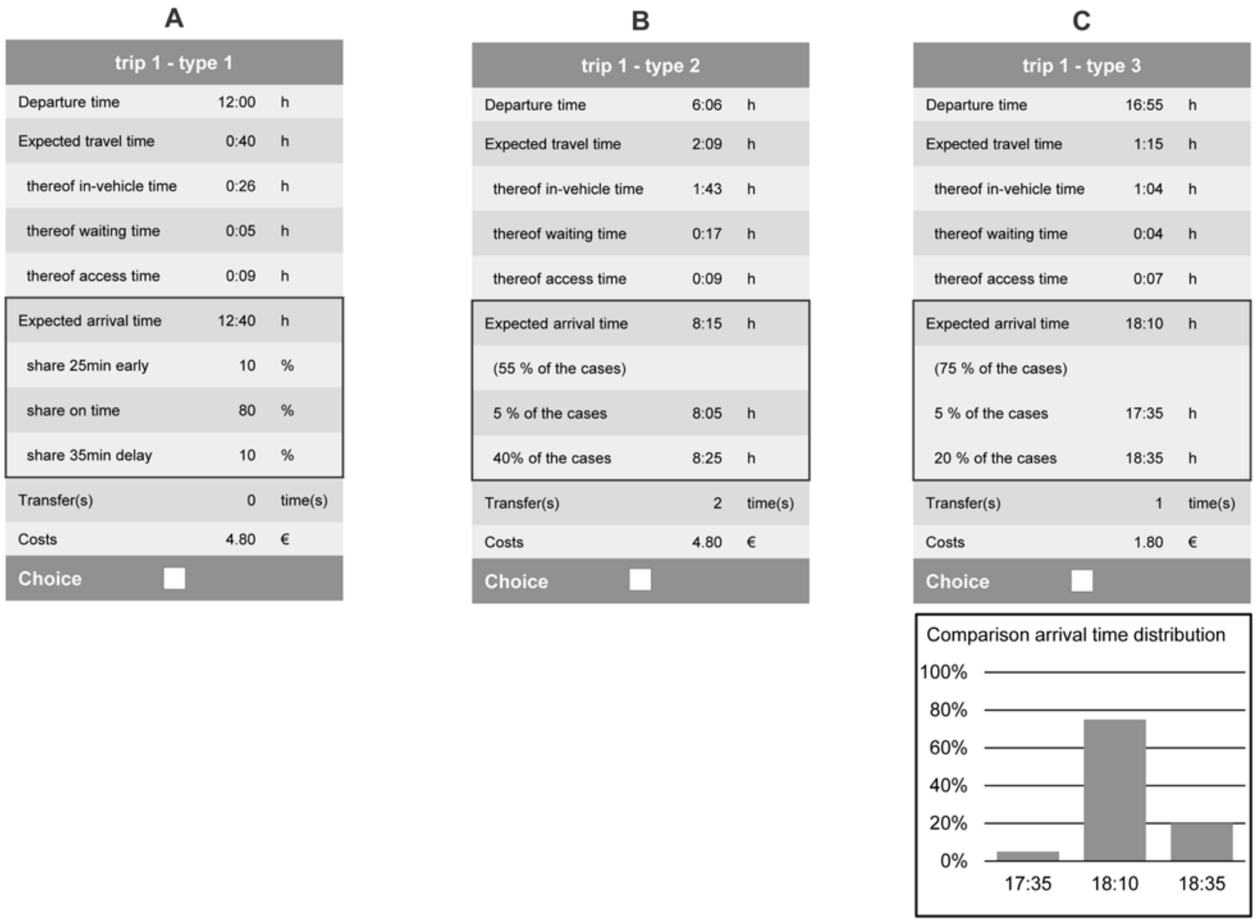

FIGURE 2: The three different types of reliability presentation

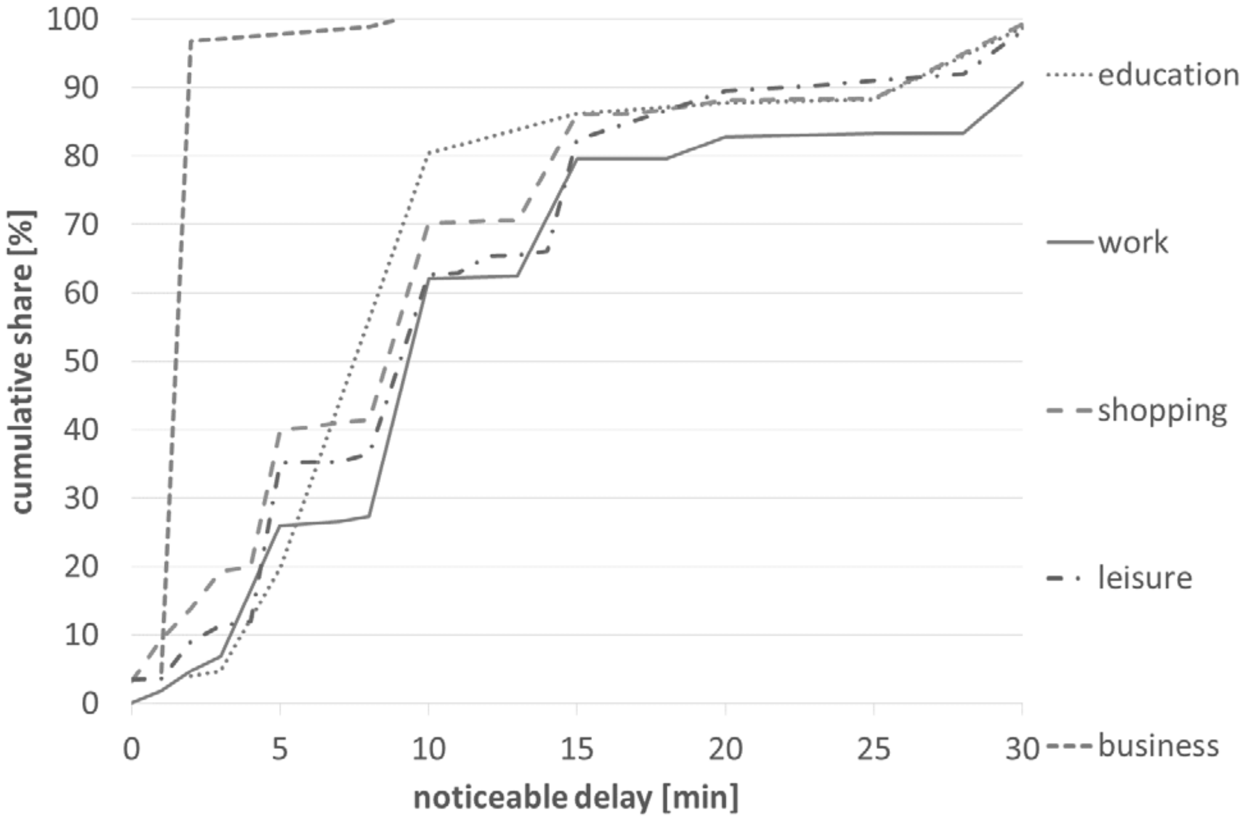

FIGURE 3: Noticeable delay (population weighted)

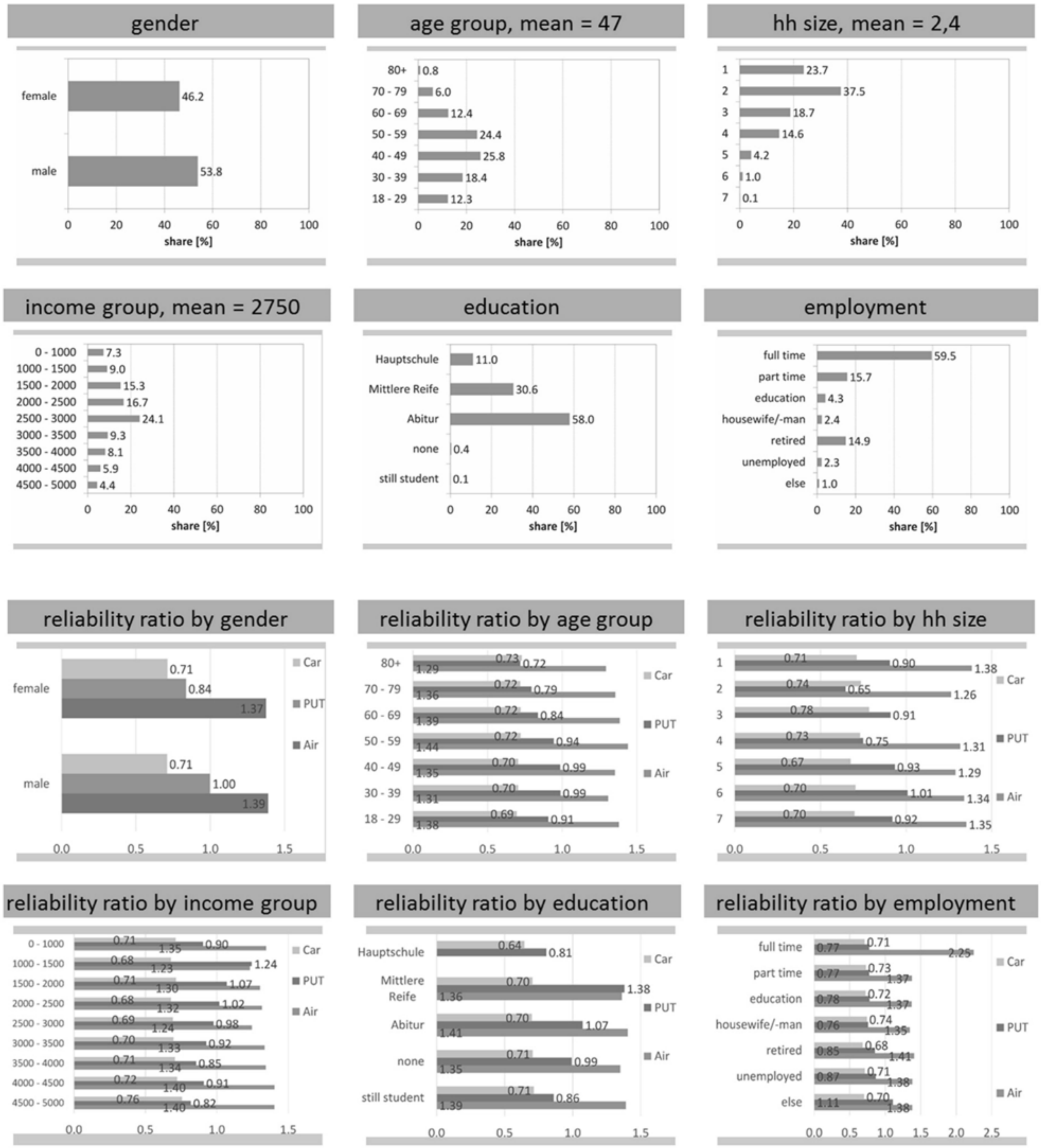

FIGURE 4: Reliability ratio by socio-demographic characteristics