NEW VALUES OF TIME AND RELIABILITY IN PASSENGER TRANSPORT IN THE NETHERLANDS

Marco Kouwenhoven, Significance
Gerard C. de Jong, Significance and ITS Leeds
Paul Koster, VU University and Tinbergen Institute Amsterdam
Vincent A.C. van den Berg, VU University and Tinbergen Institute Amsterdam
Erik T. Verhoef, VU University and Tinbergen Institute Amsterdam
John Bates, John Bates Services
Pim Warffemius, KiM Netherlands Institute for Transport Policy Analysis, Ministry of Infrastructure and the Environment

Abstract
We have established new values of time (VOTs) and values of travel time reliability (VORs) for use in cost-benefit analysis (CBA) of transport projects in The Netherlands. This was the first national study in The Netherlands (and one of the first world-wide) to investigate these topics empirically in a joint framework.

Stated preference (SP) questionnaires were designed for interviewing travellers, where the hypothetical alternatives were described in terms of travel time, travel costs and travel time reliability, the latter being presented to the respondents in the form of five possible travel times having equal probability.

For passenger transport, we first collected interviews using an existing internet panel. Additional data collection recruitment was done by asking travellers at petrol stations/service areas, parking garages, stations, bus stops, airports and ports to participate in the survey. One important conclusion is that the SP survey using members of this internet panel leads to substantially lower VOTs than the SP survey with en-route recruitment, probably because of self-selection bias in the internet panel.

We estimated discrete choice models in which the values of time differ between trips with different time and costs levels, different time and costs changes offered in the SP, and different observed characteristics of the respondents (e.g. education, income, age, household composition). By using a panel latent class model, we also account for unobserved differences between respondents in the value of time and for repeated measurements/panel effects. The reference values of time and the reference reliability ratios were estimated on the 2011 sample only, but the effect of time and cost level, time and cost changes offered and socio-economic attributes was estimated on both the 2009 and 2011 samples.

Key words: value of time, value of reliability, stated preference, latent class model.

1. INTRODUCTION

In many countries, proposals for transport infrastructure projects and other transport policies such as road pricing or changes in maximum speed, are evaluated using cost-benefit analysis (CBA). Key benefits of such projects and policies for the travellers are often the travel time gains, as well as increases in travel time reliability. To include these benefits in the CBA, a conversion of travel time and travel time reliability into monetary units is required. These conversion factors are called Value of Time (VOT) and Value of Reliability (VOR).[1] Estimated VOTs and VORs can also be used in generalised cost functions in forecasting models, but the focus of this paper is on values for use in CBA.

Reviews of methods to obtain a VOT are provided in Hensher (2007) and Gunn (2008). Recently, a number of national VOT studies that have been completed have significantly advanced the methods used in survey design and especially in analysis of the data (Fosgerau, 2006; Axhausen et al., 2008; Ramjerdi et al., 2010; Börjesson and Eliasson, 2014). In terms of numerical results for the VOT, the international meta-analysis in Wardman et al. (2012) contains the latest evidence. Methods for establishing a VOR are described in Significance et al. (2012) and a recent review of outcomes is provided in Carrion and Levinson (2012).

Most countries that use CBA in transport have official VOTs, but official VORs are missing in almost every country. Both the valuation of reliability and the inclusion of reliability in transport forecasting models are challenging. As a result, the benefits of projects and policies that reduce travel time variability are likely to be underestimated – although VOT estimates might include reliability aspects if these are not specified explicitly in the choice model underlying the estimates. But even then the estimates are biased for the evaluation of projects which have non-proportional effects on travel times and their variability. In the Netherlands, there is a long history of estimating VOTs for passenger transport (Hague Consulting Group, 1990; Hague Consulting Group, 1998). However, VORs have never been measured in a formal valuation study; i.e. a study meant to produce values for actual policy making.[2]

The objective of the study for the Dutch Ministry of Infrastructure and the Environment reported in this paper is to update the official CBA values of time for both passenger and freight transport in The Netherlands and to deliver values of reliability based on primary data. This paper is restricted to passenger transport; the results for freight transport are reported in Significance et al. (2013) and de Jong et al. (2014).

Based on earlier projects (RAND Europe, 2004; Hamer et al., 2005; HEATCO, 2006), it was decided beforehand that the variability of transport time should be measured by the standard deviation of the travel time distribution. The main reason behind this choice was the assessment that including travel time variability in transport forecasting models would be quite difficult, and that using the standard deviation

[1] Other abbreviations used in the literature for the Value of Time are Value of Travel Time (VTT) and Value of Travel Time savings (VTTS). Other terms for the Value of Reliability are Value of Variability (VOV) or Value of Travel Time Variability (VTTV).

[2] Provisional values have been selected, however, based on an expert workshop (Hamer et al., 2005), and rules-of-the-thumb are used to include reliability in the CBA assessment of projects.

would be the easiest option. Any formulation that would go beyond the standard deviation of travel time (or the variance) would be asking too much of the national and regional models that are regularly used in CBA in The Netherlands.[3]

This study distinguishes three travel purposes: commuting, business travel (i.e. travelling on employer's business) and "other" travel. Furthermore, four modes are distinguished: car, public transport (bus, tram, metro and train[4]), airplane[5] and recreational navigation[6].

Specific targets were set for the sample sizes by purpose and mode. Web-based Stated Preference (SP) interviews were carried out both in 2009 and in 2011 among travellers, and various types of discrete choice models were estimated on the resulting SP data.

The plan of this paper is as follows. Section 2 introduces the SP questionnaire and the survey design. The descriptive statistics of the samples are presented in section 3. Section 4 presents the model specifications, as well as the estimation results. VoTs and VoRs from these models are presented in section 5. This section also includes a comparison with the values from the previous national value of time studies in The Netherlands (data collected in 1988 and 1997) and the international literature. Finally, section 6 contains conclusions and recommendations.

2. SP SURVEYS

2.1 Questionnaire

The questionnaire consisted of the following parts (Table 1 lists the attributes in the SP experiments):

1. Questions regarding the attributes of a trip recently carried out[7], e.g. travel time and costs. These values are used as the base levels for the attributes presented in the SP experiments.

2. Questions regarding the availability of another mode for this trip and what the attribute levels would be for that mode. (This, however, only produced a very limited number of RP choices and proved insufficient for the estimation of an

[3] When the travel time distribution is independent of the time of the day and known by the traveller, and when a traveller chooses his departure time optimally, it is possible to estimate a scheduling model (as in Small, 1982) and calculate a value of standard deviation of transport time from the estimated scheduling coefficients (Fosgerau and Karlström, 2010). Indeed, we have tried to estimate such "scheduling models" for our departure time experiment. However, models with a marginal utility for standard deviation performed better.

[4] The train mode includes conventional train services as well as high speed rail (we did not have enough observations to report separate high-speed rail VOTs).

[5] In the previous national VOT surveys of 1988-1990 and 1997-1998, airplane was not included.

[6] In the previous national VOT surveys of 1988 and 1997, recreational navigation was not included. However, a VOT for this mode is regularly needed in The Netherlands, especially for the appraisal of proposed locks and bridges.

[7] For the 2009 survey, respondents were asked to think back to the most recent trip they had made for a certain (preselected) purpose. For the 2011 survey, respondents were asked to think back to the trip they made when they were recruited.

RP model. Furthermore, such estimation results may be biased due to a systematic reporting bias in travel time (see Peer, 2013)).

3. SP experiment 1 with six choices between two route alternatives, each described by two attributes: transport time and transport cost.

4. Introduction of variable (unreliable) travel times.

5. SP experiment 2a with six choices between two route alternatives, each described by four attributes: transport time, transport cost, transport time reliability and most likely arrival time. A fifth attribute, departure time, was calculated from the other four. Respondents in the recreational navigation segment did not participate in this experiment.

6. SP experiment 2b with seven choices similar to experiment 2a but without the variation in the most likely arrival time. To allow for consistency checking, one of the choice pairs had a dominant alternative which means that its attributes are all better than or equal to the attributes of the other alternative.

7. Questions in which respondents were asked to evaluate the choices they made in the experiments.

8. Questions about the person (age, gender, etc.) and household (composition, income, etc.).

Table 1: list of attributes in the SP experiments (excluding recreational navigation)

Attribute	Experiment 1	Experiment 2a	Experiment 2b
Usual transport time	√	√	√
Transport cost	√	√	√
Reliability, i.e. five possible transport times		√	√
Five possible arrival time		√	√
Departure time		√	√

2.2 SP experiments (except recreational navigation)

The choice situations in all SP experiments are within-mode choices. Given a certain mode, each choice set consists of two generic alternatives and the respondent was asked to choose the most preferred one.

For car trips, the choice alternatives are presented as two available routes for a respondent's particular car journey, while for public transport (including air), the two alternatives were introduced as two possible services that differ in terms of cost, (timetable) travel time, and travel time variability.

An example of an actual screen shot of a choice situation in experiment 1 is shown, in Dutch, in Figure 1. The set-up with only two alternatives (A and B) and only two attributes (travel time and travel cost) was chosen to replicate the VOT surveys of 1988 and 1997 as much as possible (so that the outcomes can be compared over time). The statistical design used here (also see Significance et al., 2007) is the ‘Bradley design’, which is close to orthogonal but without creating dominant alternatives), and where all the possible alternatives occur (full design).

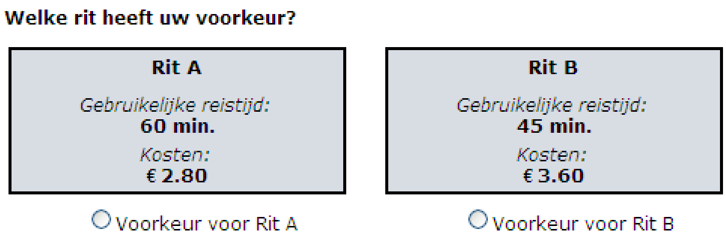

Figure 1: Example of SP question in experiment 1 for car respondents

In Figure 1, ‘Gebruikelijke reistijd’ (=usual travel time) refers to the amount of door-to-door journey time for a one-way trip. It is varied around the expected travel time at the moment of departure of the recent trip described by the respondent. ‘Kosten’ (=travel cost) refers to the total cost that a respondent has to pay for his one-way car journey.

Figure 2 provides an example of a choice situation in experiment 2a.

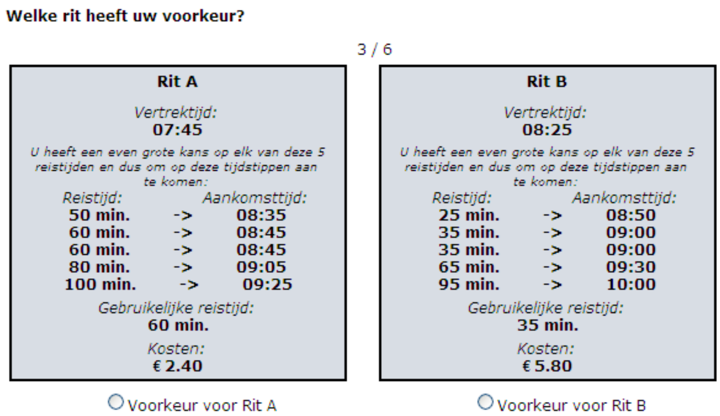

Figure 2: Example of SP question in experiment 2a for car respondents

Obviously, many respondents will not understand the concept of standard deviations, so it cannot be used directly to represent reliability in the SP experiments (though we use it later on in the modelling). Instead, reliability of travel time is presented by a series of five possible (equi-probable) travel times. Since this is the most challenging SP experiment, exploratory in-depth face-to-face interviews were carried earlier to determine the best concept and format for presenting reliability to respondents. The verbal presentation of five possible travel times turned out to work best in many respects (Significance et al., 2007; Tseng et al., 2009). The travel time distributions presented are asymmetric, which will better reflect reality than a symmetric distribution. Therefore, the mean and the median will not be equal. Furthermore, the second and third travel times (“reistijd”) are always the same, so that this time is both the mode and the median of the distribution. We estimated models both with a mean and with a median travel time, but here we will only present models based on the mean, since this is consistent with the transport models and CBA procedures used. Experiment 2a uses an orthogonal factorial design with four attributes, each with five attribute levels.

Figure 3 gives a choice situation for experiment 2b.

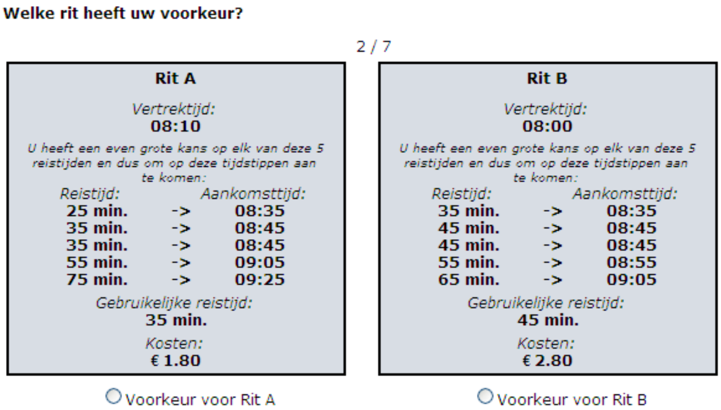

Figure 3: Example of SP question in experiment 2b for car respondents

The most likely arrival times (corresponding with the second and third travel times) for both alternatives in experiment 2b (08:45, see Figure 3) are the same, whereas they are different in experiment 2a (08:45 versus 09:00, see Figure 2). This is the difference between experiments 2a and 2b: in experiment 2a there is more scope for re-scheduling. Experiment 2b uses an extended Bradley design with three attributes, each with five levels.

2.2 SP experiments for recreational navigation

Since the purpose of a recreational navigation trip is usually not so much to travel from A to B, but to have an enjoyable trip, we expect the travel time to be valued positively instead of negatively (the longer the trip, the more preferred). However, in the Netherlands, the VOT for recreational navigation is not so much used to evaluate possible new canals or other boat routes, but to evaluate new bridges and locks.

Therefore, we have used a different setting for the recreational navigation experiments: respondents are asked to think of a situation where they have to wait for a bridge or a lock. Since we do not believe they have to depart or arrive at a certain time, we do not include departure and arrival times in the experiment. Hence, experiment 2a becomes identical to experiment 2b and only experiment 2b is presented to them. Screen shots from actual choice pairs are presented in Figure 4 (experiment 1) and Figure 5 (experiment 2b). The attributes are

- usual waiting time for a bridge/lock
- cost that one has to pay to pass the bridge or lock.
- reliability, presented as five possible and equally likely waiting times.

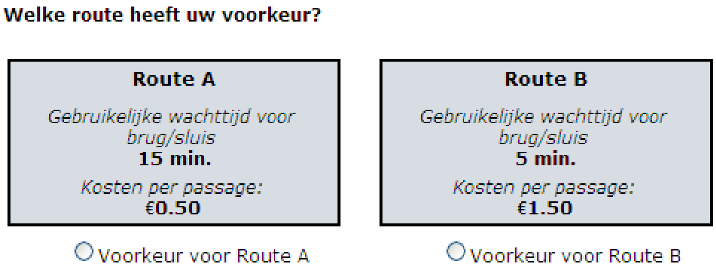

Figure 4: Example of SP question in experiment 1 for recreational navigation respondents

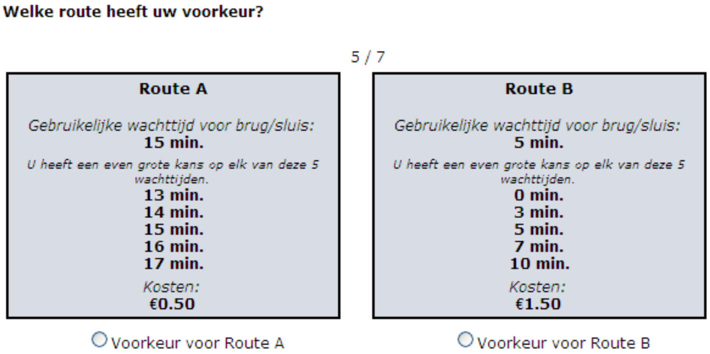

Figure 5: Example of SP question in experiment 2b for recreational navigation respondents

3. THE SAMPLES COLLECTED IN 2009 AND 2011

Two data sets were collected. The original plan was to use only data obtained from web-based interviews with members of an existing internet panel, but initial data analysis on this sample alone yielded questionable results: the VOTs from multinomial logit (MNL) models with time and cost as explanatory variables on this data set were 25-45% lower than the VOTs from 1997 (even without correcting these for inflation). Only a minor part of this difference could be explained by differences in the composition of the sample in terms of trip lengths, incomes and other socio-economic variables and the difference in the statistical SP design.

Various external experts[8] were consulted to advise on further steps. Both the external experts and the research team emphasised the possibility that the sample of respondents obtained from the panel was biased with respect to their value of time: within each segment (socio-economic, trip purpose, trip length, mode), the respondents that participate in such an online panel (which takes time, for a rather low monetary reward) may be expected to have a lower VOT than non-participants. Even after expansion, the resulting VOT would then be lower than the true VOT.

Therefore, the decision was taken to gather an additional data set, using a respondent recruitment procedure in line with the previous national VOT studies in

[8] Prof. dr. Peter Bonsall of ITS Leeds, Prof. dr. Jonas Eliasson of KTH Stockholm and Dr. Eric Molin of Delft University of Technology together with Prof. dr. Harry Timmermans of the Technical University of Eindhoven).

The Netherlands, i.e. at petrol stations/service areas along motorways, parking garages, train stations, bus stops, airports, recreational harbours and locks. These were then followed with interviews using internet technology (an almost identical web-based interview to that carried out in 2009[9]).

Table 2 and 3 list the number of successfully completed interviews for the 2009 survey and the number that was left after removing outliers (i.e. all observations on persons with implausible values on time, cost, speed, etc.). The respondents that had selected the dominated alternative in experiment 2b were also among the exclusions here. However, non-traders (respondents that always select the cheapest or the fastest alternative) have been kept. Please note that we tested the final models also including all excluded persons (both for the 2009 and the 2011 data). No systematic differences in the VOT were found, but the t-ratios are higher when excluding these respondents.

Table 2: Successfully completed interviews in the 2009 survey (internet panel)

			Purpose			
Year			**Commute**	**Business**	**Other**	**Total**
2009	Mode	Car	1341	523	790	2654
		Train	908	284	329	1521
		BTM	586	80	174	840
		Plane	0	157	374	531
		Recr.Nav.	0	0	214	214
	Total		**2835**	**1044**	**1881**	**5760**

Notes:
- BTM= bus, tram and metro
- Recr. nav. = recreational navigation.

Table 3: Interviews used in estimation: the 2009 survey (internet panel)

			Purpose			
Year			**Commute**	**Business**	**Other**	**Total**
2009	Mode	Car	1008	349	538	1895
		Train	699	235	249	1183
		BTM	469	61	136	666
		Plane	0	96	297	393
		Recr.Nav.	0	0	178	178
	Total		**2176**	**741**	**1398**	**4315**

In Tables 4 and 5 are the same statistics for the 2011 survey. We find a similar percentage of exclusions compared to 2009.

[9] A question was added as to whether the respondent was a member of an internet panel (and if so, which one); see section 5.

Table 4: Successfully completed interviews in the 2011 survey (recruitment en-route)

Year			Purpose			
			Commute	Business	Other	Total
2011	Mode	Car	184	305	125	614
		Train	131	52	103	286
		BTM	125	17	91	233
		Plane	9	29	163	201
		Recr.Nav.	0	0	95	95
	Total		**449**	**403**	**577**	**1429**

Table 5: Interviews used in estimation: the 2011 survey (recruitment en-route)

Year			Purpose			
			Commute	Business	Other	Total
2011	Mode	Car	150	235	93	478
		Train	105	41	79	225
		BTM	97	11	70	178
		Plane	7	23	152	182
		Recr.Nav.	0	0	81	81
	Total		**359**	**310**	**475**	**1144**

4. MODEL SPECIFICATIONS AND ESTIMATION RESULTS

4.1 MNL mean-dispersion models

In the literature on valuing reliability/variability of travel time in passenger transport, two main classes of models can be distinguished (see de Jong et al., 2004; Batley et al., 2008; OECD, 2010; Carrion and Levinson, 2012; Significance et al., 2012): the mean-dispersion approach and the scheduling approach. They differ in the terms that are included in the utility function[10]. First, the mean-dispersion model utility function includes travel time, cost, and a measure of travel time dispersion (usually the standard deviation, sometimes the variance). Second, the scheduling model utility function includes travel time, cost and schedule delay terms for the number of minutes that one will depart earlier or later than preferred. This specification can be based on the scheduling theory (as a departure time choice model) developed by Vickrey (1969) and Small (1982). A related alternative scheduling model that starts from the utility at the origin and the destination location over time was presented by Vickrey (1973) and Tseng and Verhoef (2008). Third, it is also possible to have both

[10] There is a theoretical equivalence relation (under certain assumptions, see footnote 3) between the Vickrey/Small scheduling approach and an approach using the mean and the standard deviation of travel time (Bates et al, 2001; Fosgerau and Karlström. 2010). There is also an equivalence relation between the Vickrey/Tseng/Verhoef scheduling model to a model with the mean and the variance of travel time (Fosgerau and Engelson, 2011). Therefore, it is theoretically possible to calculate a dispersion measure (and hence a VOR) from a departure time choice model. The best approach will depend on how one can obtain the best empirical data and which model would fit best in the transport forecasting model system that is used (Börjesson et al., 2011).

a dispersion term and schedule delay variables in the same utility function. In that case the dispersion term should pick up influences that are not related to the timing of departure and arrival per se, such as stress and re-planning cost when expected travel time does not come true, besides possibly reflecting non-linearities in the marginal utility of schedule delays.

We estimated all three specifications.[11] The mean-dispersion model (with the standard deviation) outperforms the scheduling model on the basis of the log-likelihood test. The combined mean-dispersion / scheduling model is slightly better than the mean-dispersion model. However, in that model the t-ratios for both the reliability ratio (the ratio of the value of reliability to the value of time) and the schedule delay (late) coefficient go down. In a pure mean-dispersion model, the reliability ratios pick up more of the unreliability effect. Since the difference in log-likelihood is small and since we might get into interpretation problems when using a model with both dispersion and scheduling terms (and will subsequently have to apply the results to predictions from static models), we prefer the mean-dispersion model, which can directly give us results for the monetary value of the standard deviation, as required for CBA (see section 1).

Within the mean-dispersion models, we started from a simple MNL utility function without interaction terms for socio-economic influences. Rather than estimating separate utility coefficients for each variable, we standardise the estimation in cost units, estimating a single scaling coefficient (β_C) on cost. This non-linear (in parameters) specification means that the VOT and VOR are estimated directly, as single coefficients, instead of inferred from the ratio of the estimated time (reliability) and cost coefficients.

In these models, we combined data from both surveys (2009 and 2011) and from all three SP experiments. Since we have data from two different recruitment methods, we estimate separate VOTs and reliability ratios for each survey. Further scale factors were introduced to capture possible differences in error levels between the two surveys (where the scale factor for the 2009 survey is constrained to 1) and between the experiments (where the scale factor for experiment 2a data is constrained to 1)[12].

The systematic utility function used is:

$$U = \left(Sc^{09}\cdot\delta^{09} + Sc^{11}\cdot\delta^{11}\right)\cdot\left(Sc^{\exp 1}\cdot\delta^{\exp 1} + Sc^{\exp 2a}\cdot\delta^{\exp 2a} + Sc^{\exp 2b}\cdot\delta^{\exp 2b}\right)\cdot \beta_C\cdot\left[C + \left(VOT^{09}\cdot\delta^{09} + VOT^{11}\cdot\delta^{11}\right)\cdot\left(T + \left(RR^{09}\cdot\delta^{09} + RR^{11}\cdot\delta^{11}\right)\cdot\sigma\right)\right] \qquad [1]$$

where:

- Sc^y = the scale factor for the survey in year y (2009 or 2011)
- δ^y = a dummy (0 or 1) that indicates whether an observation belongs to the survey in year y
- C = transport cost; all cost levels are corrected for inflation to 2010 levels
- T = mean transport time;

[11] All estimations were carried out using ALOGIT and/or BIOGEME.
[12] We tested the assumption whether experiment 1 would lead to higher VOTs (because they might include some of the VOR) than experiments 2a and b, but found no support for this.

σ = standard deviation of the transport time distribution;
β_c = cost parameter (coefficient to be estimated) (scale factor in money terms);
VOT^y = Value of time for year y survey (coefficient to be estimated);
RR^y = Reliability ratio (=VOR/VOT) for the year y survey (coefficient to be estimated).

Estimation results for these relatively standard MNL models are not reported here for the sake of brevity. The reader is referred to Significance et al. (2013).

4.2 Advanced MNL mean-dispersion models

Having successfully estimated standard MNL mean-dispersion models, we proceeded to estimate advanced MNL models that include sensitivity for higher base levels and, subsequently, diminishing values for smaller changes (such as lower VOT for small time savings). For this, we started from utility function [1] with the scale factors, the VOT and the RR, similar to the previous MNL models. For the purposes of exposition, we make use of a simplified version of Eq. [1] which ignores the various scale factors and dummy variables, though in practice these were retained throughout. For similar reasons we begin by ignoring the reliability component.

$$U = \beta_C \cdot \left[C + VOT \cdot T \right] \qquad [2]$$

Dependence on base (observed) cost and time
From earlier work (e.g. Gunn, 2001; Mackie et al., 2003; Stathopoulos and Hess, 2011) it is known that the VOT can be strongly dependent on the current level of the travel time and travel cost of the respondent, just as is claimed by prospect theory (Kahneman and Tversky, 1979, 1992; van de Kaa, 2008). These BaseCost (C_0) and BaseTime (T_0) values are used in the SP experiments around which the time and cost levels are varied. It can be expected that both levels are correlated. We therefore included the BaseTime and BaseCost dependency in the utility specification.

In the first place, we write the actual variables used in the SP (C, T) as the base value (C_0, T_0) plus an increment (ΔC, ΔT). Since the contribution to utility of the base values is the same for all alternatives, this allows us to re-write the utility function in terms of the increments, as in Eq. [3]

$$U = \beta_C \cdot \left[\Delta C + VOT \cdot \Delta T \right] \qquad [3]$$

Next, following earlier authors (e.g., Mackie et al., 2003; Stathopoulos and Hess, 2011), we used a power law dependence for both the BaseCost and BaseTime, multiplying the incremental cost and time terms by the base value raised to a power λ. Purely to stabilise the estimation, we divided the base value by an arbitrary "reference" value which was the same for the whole sample. Hence, the utility function becomes:

$$U = \beta_C \cdot \left(\left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \Delta C + VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T \right) \qquad [4]$$

where C_0 and T_0 are the individual's base cost and travel time respectively, and C_{ref} and T_{ref} are the reference values for base cost and travel time for the sample.

The VOT now depends on C_0 and T_0:

$$VOT(C_0, T_0) = \left. \frac{\partial U / \partial T}{\partial U / \partial C} \right|_{T=T_0, C=C_0} = VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} / \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \qquad [5]$$

From the utility definition in Eq. [4] it can be seen that choosing different reference points C_{ref} and $T_{ref,}$, will change the estimated values of β_C and VOT_{ref}, but the utility U itself remains indifferent to this choice. Therefore, the $VOT(C_0, T_0)$ in Eq. [5] is also independent of C_{ref} and T_{ref}, and we can choose them freely.[13]

Since in Eq. [4] we expect the sensitivity to a given change in cost and time to diminish, the higher the values of the base cost and time are, the expectation is that both λ_T and λ_C will be negative (see also Stathopoulos & Hess, 2011). How the VOT changes with higher values of base cost and time depends on the relative magnitudes of these two parameters, and on the relative change of the ratio of T_0/C_0, as can be seen from Eq. [5].

Dependence on the size of the time and cost changes offered in the SP
In addition to the possibility of dependency on the base levels, the most recent VOT studies (e.g. De Borger and Fosgerau (2008) for the Danish data, Ramjerdi et al. (2010) for the Norwegian data and Börjesson and Eliasson (2014) for the Swedish data) allow for different VOTs for small and large time savings offered in the SP (though they recommend using a single value for a large and a small time saving offered by transport projects). A general discussion can be found in Daly et al. (2014). We also want to correct for the influence of the size of the time and cost changes ΔC and ΔT, offered in the SP. We included another power function, of the ΔC and ΔT-terms, in the utility function to investigate whether there is any change in sensitivity further away from the base (current) values of cost and time. As before, we stabilise the estimation by dividing by an arbitrary base.

These two dependencies[14] lead to the following general specification of the utility function (again ignoring the scale factors and the differences between 2009 and 2011 coefficients from Eq. [1] for simplicity):

---

[13] The estimated parameter VOT_{ref} is the value of time when the base cost and time are equal to the arbitrarily assumed reference values. But as we will see in the next section, under our final utility function [6] set-up, the VOT when the base cost and time equals their references values is no longer equal to VOT_{ref} but proportional to it. Hence, in general, VOT_{ref} is just an intermediate coefficient without an interpretation.

[14] In line with other aspects of prospect theory, we also tested a specification with different values for gains and losses, but this did not significantly improve the model.

$$U = \beta_C \cdot \left(\left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} .\Delta C + VOT_{ref} \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \Delta T \right) \qquad [6]$$

where:

C_0,and T_0	=	the individual's base cost and time (as observed);
C_{ref},and T_{ref}	=	arbitrary reference values;[15]
ΔC and ΔT	=	cost and time changes offered in the SP;
ΔC_{ref} and ΔT_{ref}	=	arbitrary reference values;[16]
γ, λ, β_C and VOT_{ref}	=	coefficients to be estimated.

As noted, the reference values are merely included to stabilise the estimation.

With this specification, the VOT, obtained by differentiation, will depend on C_0 and T_0, and also on ΔC and ΔT:

$$VOT(C_0, T_0, \Delta C, \Delta T) = VOT_{ref} \cdot \frac{\gamma_T \cdot \left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T}}{\gamma_C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C}} \qquad [7]$$

As before, the $VOT(C_0,T_0,\Delta T,\Delta C)$ is independent of the chosen reference values for T_{ref}, C_{ref}, ΔT_{ref} and ΔC_{ref}. This is because the (estimated) VOT_{ref} is proportional to $\frac{(\Delta T_{ref})^{\gamma_T - 1} \cdot (T_{ref})^{\lambda_T}}{(\Delta C_{ref})^{\gamma_C - 1} \cdot (C_{ref})^{\lambda_C}}$, and this parameter will change to compensate for any assumptions made about the reference values.[17]

Since our assumption is that the response to a given change in cost and time may be proportionately larger, the larger the values of the change in cost and time, the expectation is that both γ_T and γ_C will be > 1 (though, as will be seen, this was not borne out in the estimation). As shown in Eq. [7], the impact on VOT will depend on the relative magnitudes of these two parameters, as well as those of the two λ parameters.

The reliability component was dealt with along similar lines. We began by replacing the term $VOT \cdot T$ in Eq. [2] by $VOT \cdot (T + RR \cdot \sigma)$. Then, along the lines of Eq. [4], we

---

[15] We chose to set C_{ref} = 3 euro, which is close to both the mean and median values for the BaseCost of the 2009 and 2011 car commute respondents, and we chose to set T_{ref} = 40 minutes, again close to the mean and median values for BaseTime of the 2009 and 2011 car commute respondents.

[16] Based on the mean and median values of ΔC and ΔT in our dataset, we chose ΔC_{ref} = 1 euro and ΔT_{ref} = 5 minutes.

[17] This time, however, when we set all the arguments of VOT to their reference values, we do not get VOT_{ref} but $(\gamma_T/\gamma_C) \cdot VOT_{ref}$. In this case. therefore, the estimated parameter VOT_{ref} can be interpreted as the value of time multiplied by the ratio (γ_C/γ_T) when all the arguments are equal to the arbitrarily assumed reference values.

allowed for a possible base reference for *RR*, replacing it by $RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R}$ where σ_0 is the individual's base standard deviation, and σ_{ref} was arbitrarily set at the level of reliability that corresponds to T_{ref} = 40 minutes). In contrast to the time and cost variables, the "gamma" variations associated with Eq. [6] were not used for the reliability variable, and the variations in the value of time resulting from different levels of ΔT were isolated from the *RR* term. The expanded version (but still excluding the scale effects and socio-economic interactions) is given as Eq. [8]:

$$U = \beta_C \cdot \left(\left(\frac{C_0}{C_{ref}} \right)^{\lambda_C} \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \Delta C + VOT_{ref} \cdot \left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \left(\left(\frac{|\Delta T|}{\Delta T_{ref}} \right)^{\gamma_T - 1} \cdot \Delta T + RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R} \cdot \Delta \sigma \right) \right) \quad [8]$$

From this, the VOR can be obtained by differentiation:

$$VOR(C_0, T_0, \sigma_0, \Delta C) = VOT_{ref} \cdot \frac{\left(\frac{T_0}{T_{ref}} \right)^{\lambda_T} \cdot \left(RR_{ref} \cdot \left(\frac{\sigma_0}{\sigma_{ref}} \right)^{\lambda_R} \right)}{\gamma_C \cdot \left(\frac{|\Delta C|}{\Delta C_{ref}} \right)^{\gamma_C - 1} \cdot \left(\frac{C_0}{C_{ref}} \right)^{\lambda_C}} \quad [9]$$

The VOT_{ref} and the RR_{ref} in Eq. [8] are estimated separately for both surveys. However, the γ and λ coefficients are jointly estimated on both surveys, since we believe that these are more intrinsic to the population and early estimates show that these coefficients are indeed similar in both data sets. By doing so, we can still make optimal use of the large 2009 data set to estimate some coefficients (such as these γ and λ factors, and also in the next section the socio-economic interaction coefficients), without having a bias in the VOT and RR.

The recent Danish, Norwegian and Swedish VOT surveys in passenger transport were estimated in logWTP space, i.e. they used logarithmic utility functions in their estimation processes (e.g. Fosgerau, 2006; Ramjerdi et al. 2010; Börjesson et al., 2012; Börjesson and Eliasson, 2014). In our standard mean-dispersion analyses for passenger transport (i.e. before the γ and λ coefficients were added) we also estimated models in logWTP space. These usually performed better than models in WTP space. However, after adding terms for the influence of the base levels and the size of the changes offered in the SP, the models in WTP space performed better than their counterparts in logWTP space.

4.3 Advanced MNL mean-dispersion models with socio-economic interaction terms

To the above mean-dispersion terms we added socio-economic terms, by interacting these with VOT_{ref} in Eq. [8]. Since we estimate a reliability ratio (RR), the VOR will go up and down together with the VOT and so the VOR is interacted with the same factors as the VOT. For segments with a higher VOT, there will also be a higher

VOR. Significant interaction coefficients were found for age class, level of education, gender, household composition, income and also for trips in the peak and for mode.

These model estimations revealed an almost linear pattern for the income interactions with the VOT (again, detailed results are in Significance et al., 2013). Only for bus/tram/metro for commuting and train for other, did we find a significant (in these cases negative) modal influence (relative to car). For all other purposes, a distinction between modes in the estimated coefficients is no longer needed: other coefficients (e.g. income) pick up differences in behaviour between modes. However, for producing the recommended VOTs, we will be using sample enumeration and expansion to national mobility figures for each mode and purpose. As a result of this, differences in, for instance, trip length between modes (e.g. longer trips by train than for other modes) can still lead to differences in the final VOTs between modes.

4.4 Latent Class mean-dispersion models

We tested several mixed logit models, where the VOT_{ref} and RR_{ref} were drawn from a continuous distribution, but obtained more stable results when estimating latent class models (Hess et al., 2011; Hensher et al., 2012) to account for unobserved differences in preferences between respondents. The latter models assume discrete distributions for certain coefficients, but without imposing a particular shape on the distribution of preferences. The result is a 'histogram' with class probabilities, and corresponding estimated values for the coefficients. A latent class model with one class is equivalent to a standard MNL model. We therefore can apply statistical tests to test for heterogeneity and determine the number of latent classes in the data. Latent class models are thus a special case of mixed logit models: they are mixed logit models with a discrete distribution for one or more of its coefficients.

The SP data includes repeated measurements for the same individual. The MNL models estimated in sections 4.1 – 4.3 assume that each choice is made in isolation. Such models are called 'cross-sectional' since they do not account for the fact that every respondent makes a sequence of choices. The latent class models we estimated account for unobserved preference heterogeneity by keeping the tastes of an individual constant over a series of choices.[18]

A discrete distribution was used for the VOT_{ref}, which takes 2 to 5 different discrete values (depending on the travel purpose). To take advantage of the sample size of the 2009 data, it was assumed that the shape of the distribution is the same for the 2009 and 2011 data, whereas the values for the VOT (VOT_{ref09_0} to VOT_{ref11_4} in Table 6) for a given (class) probability may be different. Furthermore, it was assumed that covariates have the same proportional effect on the VOT and VOR in 2009 and 2011.

To save on the number of parameters, we interacted the covariates with the VOT_{ref} variables which means that the shape of the distribution is the same for each

[18] The latent class models we used do not extend to autocorrelation because simulation of an error component is very demanding in terms of computation time. Extending the latent class model with error components is an interesting direction for future research.

combination of covariates but the mean VOT_{ref} is not.[19] We estimated these models for commuting, business and other travel separately where we also included the covariates that were found to be significant in the earlier MNL estimation. For each of the estimations we optimise the number of classes using the Bayesian Information Criterion (BIC). The optimal number of classes is 5 for commuting, 4 for business, other and air travel and 3 for recreational navigation. The results are given in Table 6. The t-ratios in this table are so-called ‘robust’ t-ratios, which allow for non-severe misspecification errors (Bierlaire, 2008). Note that for recreational navigation no stable estimation of the gammas could be obtained, therefore they were constrained to one.

Table 6: Estimated coefficients and t-ratios for combined advanced latent class mean-dispersion models with socio-economic interaction terms

		Commute		Business		Other	
Car / Train / BTM	Observ. (resp.)	45186 (2528)		16476 (932)		20658 (1165)	
	Final log (L)	-22103.134		-7925.771		-9471.395	
	Rho^2(0)	0.33		0.362		0.156	
		Value	(T-ratio)	Value	(T-ratio)	Value	(T-ratio)
	BetaCost	-1.16	(-35.2)	-1.05	(-20.6)	-1.45	(-22.6)
	VOTref09_0	4.06	(19.8)	38.1	(6.4)	0.168	(0.8)
	VOTref09_1	69.3	(5.1)	9.37	(12.6)	10.5	(14.0)
	VOTref09_2	0	(0.0)	4.18	(10.9)	5.01	(15.1)
	VOTref09_3	15.3	(9.2)	0	(0.0)	163	(3.1)
	VOTref09_4	7.75	(19.8)	-		-	
	VOTref11_0	5.57	(19.3)	44.1	(5.9)	6.82	(10.3)
	VOTref11_1	57.3	(2.0)	1.31	(1.3)	39.1	(4.5)
	VOTref11_2	11.9	(10.7)	5.19	(12.3)	1.68	(1.3)
	VOTref11_3	46.3	(2.6)	12.2	(7.1)	12.4	(2.3)
	VOTref11_4	0	(0.0)	-		-	
	Group_0	0	(*)	0	(*)	0	(*)
	Group_1	-2.55	(-21.6)	1.17	(4.5)	-1.5	(-5.4)
	Group_2	-0.33	(-3.8)	2.28	(14.1)	-0.13	(-0.8)
	Group_3	-2.42	(-8.8)	1.58	(9.5)	-3.06	(-9.0)
	Group_4	-0.775	(-6.5)	-		-	
	RR09	1.17	(11.5)	1.51	(10.4)	1.2	(3.4)
	RR11	0.408	(2.2)	1.15	(6.8)	0.624	(1.3)
	facTrain					-0.106	(-1.9)
	facBTM	-0.0891	(-2.5)				
	fac3650	-0.107	(-4.1)			-0.0396	(-0.7)
	fac51pl	-0.186	(-6.5)	-0.104	(-1.9)	-0.233	(-4.8)
	facEdu1	-0.331	(-1.8)				
	facEdu2	-0.11	(-2.5)				
	facEdu12			-0.284	(-3.8)	-0.0826	(-1.0)
	facEdu34	-0.0409	(-1.5)			-0.0911	(-2.5)
	facFem			0.0062	(0.1)	0.0465	(0.8)
	facHH1	0.138	(3.4)				
	facHH12			0.165	(2.0)		
	facInc	0.0761	(8.1)	0.109	(6.3)	0.0273	(1.8)
	facPeak	0.0789	(2.8)	0.188	(3.0)	0.123	(2.4)
	gammaC	0.523	(40.9)	0.548	(28.4)	0.537	(28.6)
	gammaT	1.06	(48.8)	1.01	(31.2)	1.05	(18.8)
	lambdaC	-0.386	(-22.1)	-0.473	(-16.7)	-0.382	(-16.0)
	lambdaT	-0.526	(-14.8)	-0.515	(-10.5)	-0.559	(-7.3)
	lambdaR	-1.05	(-7.2)	-1.19	(-15.4)	-0.86	(-4.1)
	Sc11	0.723	(20.6)	0.683	(16.8)	0.636	(16.8)
	ScEx1	4.02	(24.7)	4.34	(15.1)	3.52	(14.1)
	ScEx2b	1.33	(25.1)	1.35	(15.3)	1.37	(18.6

19 Class membership models will be tested in the future but they need many more parameters since for each covariate one needs to add N-1 variables to the model, where N is the number of classes.

		Commute	Business	Other
Plane	Observ. (resp.)		9750 (575)	
	Final log (L)		-4952.262	
	Rho² (0)		0.164	
			Value	(T-ratio)
	BetaCost		-3.32	(-5.1)
	VOTref09_0		3.74	(5.9)
	VOTref09_1		0	(0.0)
	VOTref09_2		1.79	(5.7)
	VOTref09_3		9.83	(4.9)
	VOTref11_0		0.593	(1.0)
	VOTref11_1		6.68	(5.7)
	VOTref11_2		2.36	(4.8)
	VOTref11_3		35	(0.6)
	Group_0		0	(*)
	Group_1		-0.0667	(-0.2)
	Group_2		0.762	(1.9)
	Group_3		-1.66	(-4.9)
	RR09		1.35	(1.1)
	RR11		0.653	(1.4)
	facBus		0.0295	(0.4)
	gammaC		0.64	(22.6)
	gammaT		1.01	(24.9)
	lambdaC		-0.722	(-13.2)
	lambdaT		-0.796	(-11.6)
	lambdaR		-1.62	(-3.3)
	Sc11		0.72	(12.8)
	ScEx1		4.16	(10.7)
	ScEx2b		1.28	(12.1)
Recreational navigation	Observ. (resp.)			3102 (259)
	Final log (L)			-1660.0
	Rho² (0)			0.25
				Value (T-ratio)
	BetaCost			-0.336 (-9.5)
	VOTref09_0			15.9 (8.1)
	VOTref09_1			5.15 (19.1)
				1.39 (4.2)
	VOTref09_2			
	VOTref11_0			0.291 (0.4)
	VOTref11_1			6.39 (13.0)
	VOTref11_2			40.8 (4.5)
	Group_0			0 (*)
	Group_1			1.54 (6.2)
				-0.632 (-1.7)
	Sc11			1.01 (6.6)
	ScEx1			5.75 (6.9)

In which:

- BTM: bus, tram and metro.
- VOT_{ref}: VOT value for one of the latent classes.
- Group: coefficient that governs the latent class membership probabilities.
- fac3650: travellers in the age class 36-50 have a lower commuting and other VOT than the younger age classes, which form the base or reference category.
- fac51pl: travellers in the age class 51 and older have a lower VOT than those younger than 36 (and also than those in the age class 36-50), for all three purposes.
- facEdu1: travellers with primary school as highest education have a lower commuting VOT than those with high education levels (College University)
- facEdu2: travellers with lower secondary school as highest education have a lower commuting VOT than those with high education levels (College/University)

- facEdu34: travellers with medium/higher secondary school as highest education have a lower commuting and other VOT than those with high education levels (College/University)
- facEdu12: travellers with primary school or lower secondary school as highest education have a lower other VOT than those with high education levels (College/University)
- facFem: females have a higher business and other VOT (possibly since they are often involved more in multi-tasking).
- facHH1: households with only one member have a higher commuting VOT than other households (we expect that these households cannot share other tasks with other members, so they have higher opportunity costs for travel).
- facHH12: households with one adult with or without children have a higher business VOT than other households (we expect that these households have more difficulty in sharing other tasks with other members than other households, so they have higher opportunity costs for travel).
- facInc: linear income: higher incomes have a higher VOT for all purposes.
- facPeak: trips in the peak (midpoint of trip falls within 7-9 hours or 16-18 hours) have a higher VOT for al purposes; this may have to do with the additional nuisance of travelling in congested/crowded conditions.
- facBTM: bus/tram/metro for commuting has a lower VOT than car.
- facTrain: train travellers have a lower other VOT (possibly because they can use their time in the train more pleasantly and productively than in other modes, using information technology) than car.

All the covariates have the intuitive signs although some of the covariates which were significant in MNL are not significant any more.

From these estimation results we can conclude that the parameters for reference dependence are significant, except for the γ (size) coefficient for time, which is not significantly different from 1 in the models for business and other travel. The "distance" effect λ is between -0.3 and -0.6 for both time and cost, so they are both negative as was expected. γ is between 0.4 and 0.6 for cost, and between 0.85 and 0.95 for time. We find that longer trips will have a higher VOT and that the *cost* changes offered in the SP have a considerable impact on the VOT (less so for the time changes).

In the estimation results for the latent class models, the VOT_{ref} for 2009 is 30% lower than the VOT_{ref} for 2011. Moreover, when we ran additional models (that are not presented her) on the 2011 data distinguishing between members of an internet panel and non-members, we got 10-30% lower VOTs for the members. This strongly suggests that, indeed, Internet panels can be selective with respect to the VOT.

The reliability ratios for the 2009 and 2011 surveys are significantly different from each other and also significantly different from 1. For the recommended VOT and VOR, we use the VOT_{ref} and RR_{ref} of 2011, which are much more in line with the previous Dutch surveys and international evidence than those for 2009, and do not have the bias that we believe is present in our estimates from the internet panel.

5. RECOMMENDED VOTS AND VORS

5.1 The method used

For the recommended VOTs we used the estimation results from the latent class models as presented in section 4. In order to get from these model estimation results to recommended VOTs and VORs a number of steps were taken. First, for business travel we used the so-called ‘Hensher equation’ (see Hensher 1977, Fowkes et al. 1986 and Wardman et al. 2013), which decomposes the VOT into an employee and an employer component. The former comes from the latent class model on the SP data among business travellers, the latter from survey data (from the 2011 survey) on the fraction of journey time spent working, the relative productivity of travel versus work, the percentage of saved time that would be spent working and the productive value of work time. Second, for the employee components and for all other purposes we used the 2011 sample and calculated the VOT and VOR for each individual in the sample, depending on his/her socio-economic and trip characteristics (respondent-specific base time and cost). Since it is believed that the dependence on the ΔT and ΔC is (partly) an SP artefact and may also lead to difficulties in a CBA (since a 2 € change would no longer be equal to twice a 1 € change), we used $\gamma_C = \gamma_T = 1$ to compute the VOT and VOR. Eq. [6] now no longer explicitly depends on ΔT and ΔC. However, the value of VOT_{ref} does depend on the chosen values of ΔT_{ref} and ΔC_{ref}, so a sensible choice for these reference values is required. We solved this by calculating for each respondent the mean time and cost differences between the two alternatives over the 18 non-dominant choices that he has been asked to make in SP experiments[20]. Third, in a subsequent step, we made our survey representative for the mobility of the Dutch population. For this, we divided all trips in the OViN (“Onderzoek Verplaatsingen in Nederland 2010”, the National Travel Survey by Statistics Netherlands) based on five population variables (gender, age, income, household composition and education) and two trip variables (period of the day combined with travel mode and travel duration category combined with travel mode). The OViN survey contains approximately 136,000 records. Every record in the OViN survey also has a weight factor in order to make the OViN survey representative for all trips of the Dutch population in one year. For this, we only considered people older than 16 who used car, train or bus/tram/metro as method of transportation for their trip. The distribution of the trips in our survey over the seven variables is different from the whole Dutch population (e.g. many more commuters are present in the VOTVOR survey than one would expect in a typical sample). An Iterative Proportional Fitting method was used to calculate new weights for our survey such that the weighted distributions for the seven variables match the weighted distributions of the OViN survey. In calculating the final values of time, the value of time for each respondent was weighted with the factor as determined by this expansion procedure, and with the travel time.

[20] We validated this method by also estimating advanced MNL models with γ_C and γ_T constrained to 1. The average VOT over the sample remained similar (within 10%) to the average VOT calculated using the method described above.

5.2 Outcomes for the VOT

The new VOTs that were obtained in this way are presented in Table 7. In Table 8 (final two columns) they are compared against the values previously used in CBA. Most of the new VOTs are not very different from the previous official Dutch values, which are based on the national VOT survey of 1997-1998 (Hague Consulting Group, 1998), corrected for inflation and real income growth. Overall, there is a decrease for car and an increase for train (where the BTM overall value remains basically the same).

We also looked at the average VOT for different segments of respondents. From this we concluded that the VOT increases with income and with trip length, as was found in earlier studies. The dependency of the VOT on the trip length cannot be directly derived from the (relative) values of λ, since the (relative) dependency of base time and cost on trip length also needs to be taken into account (see Eq. [7]). Using a similar approach, we also confirm that smaller time differences are valued less per minute than larger time differences. This has been found before in several studies (see for instance the review by Daly et al. 2014).

Table 7: New values of time (in 2010 euros per hour per person, including VAT) for car driver, train, bus/tram/metro, air and recreational navigation from this survey

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	9.25	11.50	7.75	9.75		
Business employee	*12.75*	*15.50*	*10.50*	*13.50*	*85.75*	
Business employer	*13.50*	*4.25*	*8.50*	*10.50*	-	
Business	26.25	19.75	19.00	24.00	85.75	
Other	7.50	7.00	6.00	7.00	47.00	8.25
All purposes	9.00	9.25	6.75	8.75	51.75	8.25

Note: all values are rounded off to the nearest multiple of € 0.25

Table 8 also presents a comparison against the international literature. The first column of this table gives the outcomes of an application of the estimation results of the international meta-analysis of Shires and de Jong (2009) to The Netherlands (e.g. using the Dutch GDP per capita, etc.). Here we adjusted for price changes since 2003 (the year that Shires and de Jong used to express their VOTs), but no correction for real income growth on top of that was applied. Income change-compensated values from the meta-analysis would be slightly higher than the values in this column.

The most recent national VOT studies are those of Sweden (Börjesson and Eliasson, 2014) and Norway (Ramjerdi et al., 2010). The study teams in these countries estimated non-parametric models, which account for the sign and size of the travel time changes offered, observed heterogeneity and unobserved heterogeneity between respondents. In Table 8 for Sweden and Norway we sometimes give two values per cell: the lower values for Norway refer to short distances, the higher values refer to long distances (>100 km). For Sweden the lower values are valid outside Stockholm and the higher values in Stockholm. All Swedish VOTs in this table refer to short distances; for longer distances, the Swedish VOTs are higher than presented (maximally 14.9 euro per hour), but for these there is no distinction between travel purposes. In Norway, train and BTM are not separate categories, whereas the Swedish study did not include business travel.

Table 8: Values of time (in euro of 2010 per hour per person, including VAT) for car driver, train and BTM (bus/tram/metro) from various sources

	Value for the Netherlands from international meta-study (Shires and de Jong, 2009)	Norway (Ramjerdi et al. 2010)	Sweden (Börjesson and Eliasson, 2014)	Previous CBA value for The Netherlands	New value for The Netherlands
Commuting – car driver	11.05	12.13-26.95	9.2-12.1	9.55	9.25
Commuting – train	11.05		7.2	9.62	11.50
Commuting – BTM	9.14		5.3	8.93	7.75
Business – car driver	30.94	51.20		33.07	26.25
Business – train	30.94			20.36	19.75
Business – BTM	24.83			15.56	19.00
Other – car driver	8.85	10.37-19.67	5.9-7.8	6.59	7.50
Other – train	8.85		5.0	5.93	7.00
Other – BTM	6.21		2.8	5.65	6.00
Car – all purposes	-			10.67	9.00
Train – all purposes	-			7.58	9.25
BTM- all purposes	-			6.63	6.75

Notes:

- Business values include employee and employer components.
- All values from our new study are rounded off to the nearest multiple of € 0.25

We see in Table 8 that the recommended new values for commuting and other purposes provide a good match with the international literature (represented by the meta-analysis and the most recent national VOT studies). The meta-analysis is mainly based on studies that use MNL models, and we have found that in our study MNL estimates have a downward bias due to ignoring unobserved heterogeneity and panel effects. Recent studies that also take account of unobserved heterogeneity (Sweden and Norway) produce values which are often higher than the older literature, and our new values for commuting and other are within the bandwidths provided by the Norwegian and Swedish studies.

For business travel, our new values are somewhat lower than those of the meta-analysis. With the exception of car driver, the previous Dutch CBA values for business travel are also smaller than those from the meta-analysis. The difference between the previous CBA values and the new Dutch values on the one hand and those from the meta-analysis on the other hand may be caused by the latter being mainly based on countries that use the wage costs for the business VOT. The Dutch studies (1988-1990, 1997-1998 and this one) all took account of the fact that not all saved time on business trips is used for the employer and that travel time is not necessarily unproductive; this reduces the business VOT.

5.3 Outcomes for the RR and the VOR

In order to apply the valuation of reliability in CBA, one should multiply the RR by the corresponding VOT in euro per hour per person. For commuting and other travel this is rather straightforward. For business travel we have the situation that the RRs come from the individual travellers, but we have a VOT that consists of an employer and an employee component (though both were derived from interviewing the traveller). There is no information for calculating a separate employer component in the business VOR. We think it is best to assume that the business RR applies to the sum of the employer and employee component, i.e. to the total business VOT. All VORs are displayed in Table 9. Table 10 gives a comparison of the RRs against the international literature.

Table 9: VORs (in 2010 euros per hour per person, including VAT) as found in this study

	Car	Train	Bus, tram, metro	All surface modes	Air	Recr. navigation
Commute	3.75	4.75	3.25	4.00		
Business employee	*14.50*	*18.00*	*12.00*	*15.50*	*56.00*	
Business employer	*15.50*	*4.75*	*9.75*	*12.25*	*-*	
Business	30.00	22.75	21.75	27.75	56.00	
Other	4.75	4.50	3.75	4.50	30.75	0
All purposes	5.75	5.50	3.75	5.25	33.75	0

Notes:

- Business values include employee and employer components.
- All new values are rounded off to the nearest multiple of € 0.25

There have been a few more studies on the value of reliability that provided numerical outcomes than listed in Table 10, but these provide metrics other than the RR, and are therefore not comparable. We also included in Table 10 the outcomes of the 2004 expert workshop which provided provisional values for the RR for use in CBA in The Netherlands, even though these are not empirical findings but expert judgments (see Hamer et al., 2005). In a number of CBAs in the Netherlands the reliability benefits have been calculated as 25% of the travel time benefits (based on Besseling et al. (2004)). From Table 10 we conclude that the new RRs that we obtained fit quite well within the range of values provided by the international literature. All values we now get (except the one for business for car) are lower than the provisional values from the expert workshop of 2004, but many recent empirical values are also lower than the workshop values. For air transport, we found just one other study that provided an RR (Norway), and that value is clearly lower than our value (which is more comparable to the RRs for other modes).

Table 10: Comparison against the empirical literature on the reliability ratio (for the value of the standard deviation of travel time versus average travel time)

Study	Country	RR
Car		
MVA (1996)	UK	0.36 – 0.78
Copley et al. (2002)	UK	Pilot survey: 1.3
Hensher (2007)	Australia	0.3 – 0.4
Eliasson (2004)	Sweden	0.30 – 0.95
Mahmassani (2011)	USA	NCHRP 431: 0.80 – 1.10 SHRP 2 CO4: 0.40 – 0.90
Expert workshop of 2004	*The Netherlands*	*0.8*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Train		
ATOC (2002)	UK	0.6 – 1.5
Ramjerdi et al. (2010)	Norway	Short trips: 0.69 Long trips: 0.54
Expert workshop of 2004	*The Netherlands*	*1.4*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Bus/tram/metro		
MVA (2000)	France	0.24
Ramjerdi et al. (2010)	Norway	Short trips: 0.69 Long trips: 0.42
Expert workshop of 2004	*The Netherlands*	*1.4*
This study	The Netherlands	Commuting: 0.4 Business: 1.1 Other: 0.6
Air		
Ramjerdi et al. (2010)	Norway	0.20
This study	The Netherlands	Business: 0.7 Other: 0.7

6. CONCLUSIONS

New values of time (VOTs) and values of travel time reliability (VORs) have been established for use in cost-benefit analysis (CBA) of transport projects in The Netherlands. The previous Dutch passenger values of time are based on surveys carried out in 1997. While values of reliability also existed for use in CBA, these were only based on expert opinions. This was the first national study in The Netherlands that empirically investigated both topics in a joint framework.

As the operational measure for unreliability we use the standard deviation of travel time. The ratio of the value of the standard deviation to the value of travel time is called the reliability ratio. The main reason for choosing this definition was that all other possible measures of reliability would be much harder to incorporate in the national and regional transport models.

One important conclusion is that the 2009 SP survey using members of an internet panel leads to substantially lower VOTs than the 2011 SP survey (with en-route recruitment). The most likely interpretation is that the 2011 values are correct and that the 2009 values are biased downwards, mainly because persons with a lower value of time (in every socio-economic segment) have a higher probability of becoming a member of an internet panel. The shorter distances of the trips sampled in 2009 (and the corresponding smaller time savings offered in the SP) also played a role. The reference values of time (and reliability) in our estimates are therefore based on 2011 only.

We estimated discrete choice models in which the values of time differ between trips with different time and costs levels, different time and costs changes offered in the SP, and different characteristics of the respondents (e.g. education, income, age, household composition). By using a panel latent class model, we also account for unobserved differences between respondents in the value of time and for repeated measurements/panel effects. While the reference values of time and the reference reliability ratios were estimated on the 2011 sample only, the effect of time and cost level, time and cost changes offered and socio-economic attributes was estimated on both the 2009 and 2011 samples.

These are absolute models in willingness-to-pay space. When including the dependencies of the VOT and the VOR on the observed levels of time and cost and on the magnitude of the changes in the attributes offered in the SP, models in willingness-to-pay space perform better than models in log willingness-to-pay space, and are therefore preferred.

In our models we made the dependence of the VOT on base time and base cost explicit. In the calculation of recommended values for CBA, we used a weighting procedure for this. More worrying is the dependence on the time and cost changes offered in the SP, which was also made explicit in the new models. We found that the impact of the cost changes on the VOT was higher than that for the time changes. A consequence of this is that the choice of time and especially of cost values offered in the SP (the statistical SP design) has an impact on the resulting VOT. This has remained implicit in many previous studies, but now it has become explicit. How best to deal with this issue requires further research.

The recommended values of time were calculated by weighting the sampled respondents to represent the distribution of time travelled in the trips recorded in the Dutch national travel survey OViN.

It is in some sense encouraging that most of the new VOTs are not very different from the previous official Dutch values, and are within the range of the recent international literature, especially that for comparable models.

For the value of reliability, we estimated the reliability ratio for each modelling segment. The reliability ratio gives the monetary value of reliability (measured as standard deviation of transport time) divided by the value of time. We obtained values between 0.4 and 1.1 for the reliability ratio, depending on the travel purpose, reasonably in line with recent empirical studies in other countries..

Acknowledgements
This paper is based on a study carried out for the KiM Netherlands Institute for Transport Policy Analysis, Ministry of Infrastructure and the Environment. We also thank SURFsara (www.surfsara.nl) for the support in using the Lisa Computer Cluster and Jasper Knockaert for assistance in using PythonBiogeme.

REFERENCES

Association of Train Operating Companies (ATOC) (2002) Passenger Demand Forecasting Handbook, London: ATOC.

Axhausen, K.W., S. Hess, A. König, G. Abay, J.J. Bates and M. Bierlaire (2008) Income and distance elasticities of values of travel time savings: New Swiss results, *Transport Policy*, 15(3), 173-185.

Bates, J., J. Polak, P. Jones and A. Cook (2001) The valuation of reliability for personal travel, *Transportation Research E (Logistics and Transportation Review)*, 37-2/3, 191-229.

Batley, R.P., S. Grant-Muller, J. Nellthorp, G.C. de Jong, D. Watling, J.J. Bates, S. Hess and J. Polak (2008) Multimodal Travel Time Variability, Final Report, report for the UK Department of Transport, ITS Leeds, John Bates and Imperial College.

Besseling, P., W. de Groot and A. Verrips (2004) Economische toets op de Nota Mobiliteit, *CPB document 65*, CPB, The Hague.

Bierlaire, M. (2008) An introduction to BIOGEME version 1.7, biogeme.epfl.ch.

Borger, B. de and M. Fosgerau (2008) The trade-off between money and travel time: A test of the theory of reference-dependent preferences, *Journal of Urban Economics*, 64, 101-115.

Börjesson, M., and J. Eliasson (2014) Experiences from the Swedish Value of Time study, *Transportation Research Part A: Policy and Practice*, *59*, 144-158.

Börjesson, M., J. Eliasson and J.P. Franklin (2012) Valuations of travel time variability in scheduling versus mean–variance models. *Transportation Research Part B: Methodological*, *46* (7), 855-873.

Carrion, C. and D. Levinson (2012) Value of travel time reliability: A review of current evidence, *Transportation Research A,* 46, 720-741.

Cirillo, C., A.J. Daly and K. Lindveld (2000) Eliminating bias due to the repeated measurements problem, in J. de D. Ortúzar (Ed.): *Stated Preference Modelling Techniques*, PTRC, London.

Copley, G., P. Murphy, and D. Pearce (2002) Understanding and valuing journey time variability; European Transport Conference – 2002, Cambridge.

Daly, A.J., F. Tsang and C. Rohr (2011) The value of small time savings for non-business travel, Paper presented at European Transport Conference 2011, Glasgow.

Eliasson, J. (2004) Car drivers' valuations of travel time variability, unexpected delays and queue driving, *Proceedings of the European Transport Conference*, 2004.

Fosgerau (2006) Investigating the distribution of the value of travel time savings, *Transportation Research Part B*, 40(8), 688-707.

Fosgerau, M. and A. Karlström (2010) The value of reliability, *Transportation Research B*, 44(1), 38-49.

Fosgerau, M. and L. Engelson (2011) The value of travel time variance, *Transportation Research B*, 45(1), p.1-8.

Fowkes, A.S., P. Marks and C.A. Nash (1986) The value of business travel time savings, Working Paper 214, Institute for Transport studies, University of Leeds.

Gunn, H.F. (2008) Valuation of travel time savings and losses, In D.A. Hensher and K.J. Button (Eds.): *Handbook of Transport Modelling*, Second Edition, Pergamon, Oxford.

Hague Consulting Group (1990) The Netherlands' value of time study: final report, Report for DVK, HCG, Den Haag.

Hague Consulting Group (1998) The second Netherlands' value of time study: final report, Report 6098-1 for AVV, HCG, Den Haag.

Hamer, R., G.C. De Jong, and E.P Kroes (2005) The value of reliability in Transport – Provisional values for the Netherlands based on expert opinion, RAND Technical Report Series, TR-240-AVV, Netherlands.

HEATCO (2006) Developing Harmonised European Approaches for Transport Costing and Project Assessment, Deliverable 5, Proposal for harmonized guidelines. IER, University of Stuttgart.

Hensher, D.A. (1977) Value of business travel time, Pergamon Press.

Hensher, D.A. (2007) Valuation of Travel Time Savings, in: A. de Palma, R. Lindsey, E. Quinet and R. Vickerman (Eds.): Handbook in Transport Economics, Edward Elgar Publisher.

Hensher, D.A., J.M. Rose and Z. Li (2012) Does the choice model method and/or the data matter? *Transportation*, 39(2), pp. 351-385.

Hess, S., M. Ben-Akiva, D. Gopinath and J. Walker (2011) Advantages of latent class over continuous mixture of logit models, Working paper, ITS, University of Leeds.

Jong, G.C. de, M. Kouwenhoven, E.P. Kroes, P. Rietveld and P. Warffemius (2009) Preliminary monetary values for the reliability of travel times in freight transport, in *European Journal of Transport and Infrastructure Research,* 9(2), pp. 83-99.

Jong, G.C. de, M. Kouwenhoven, J. Bates, P. Koster, E. Verhoef. L. Tavasszy and P. Warffemius (2014) New SP-values of time and reliability for freight transport in the Netherlands, *Transportation Research E*, 64, pages 71-87, 2014.

Kaa, E.J. van de (2008) Extended prospect theory, Findings on choice behaviour from economics and the behavioural sciences and their relevance for travel behaviour, PhD Thesis, Delft University of Technology.

Kahneman, D. and A. Tversky (1979) Prospect theory: an analysis of decision under risk, *Econometrica*, 47, 263-291.

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

Mahmassani, H.S. (2011) Application to the New York metropolitan region of an integrated model of user responses to pricing and reliability with a state of the art simulation-based dynamic traffic assignment tool, Presentation to NYMTC, New York City.

MVA (1996) Benefits of reduced travel time variability; report to DfT; MVA.

MVA (2000) Etude de l'impact des phénomènes d'irregularité des autobus – Analyse des resultats, MVA, Paris.

Peer, S. (2013) The economics of trip scheduling, travel time variability and traffic information. PhD Thesis, VU University Amsterdam.

Ramjerdi, F., S. Flügel, H. Samstad and M. Killi (2010) Value of time, safety and environment in passenger transport – Time. TØI report 1053B/2010, Institute of Transport Economics, Oslo.

RAND Europe (2004) De Waardering van kwaliteit en betrouwbaarheid in personen- en goederen vervoer (The valuation of quality and reliability in passenger and freight transport). AVV/RAND Europe, Rotterdam.

Significance, VU University Amsterdam and John Bates (2007) The value of travel time and travel time reliability, Survey design, Final Report prepared for the Netherlands Ministry of Transport, Public Works and Water Management, Significance. Leiden.

Significance, Goudappel Coffeng and NEA (2012) Erfassung des Indikators Zuverlässigkeit des Verkehrsablaufs im Bewertungsverfahren der Bundesverkehrswegeplanung: Schlussbericht, Report for BMVBS, Significance, The Hague (see:

http://www.bmvbs.de/SharedDocs/DE/Artikel/UI/bundesverkehrswegeplan-2015-methodische-weiterentwicklung-und-forschungsvorhaben.html).

Significance, VU University, John Bates Services, TNO, NEA, TNS NIPO and PanelClix (2013) Values of time and reliability in passenger and freight transport in The Netherlands, Report for the Ministry of Infrastructure and the Environment, Significance, The Hague.

Small, K.A. (1982) The Scheduling of Consumer Activities: Work Trips, *American Economic Review*, 72, 467-479.

Stathopoulos, A.I. and S. Hess (2011) Revisiting reference point formation, gains-losses asymmetry and non-linear sensitivities: one size does not fit all! Paper presented at ETC 201, Glasgow.

Tseng, Y.Y. and E.T. Verhoef (2008) Value of time by time of day: A stated-preference study, *Transportation Research B*, 42(7-8), 607-618.

Tseng, Y.Y., E.T. Verhoef, G.C. de Jong, M. Kouwenhoven and A.I.J.M. van der Hoorn (2009) A pilot study into the perception of unreliability of travel times using in-depth interviews, *Journal of Choice Modelling*, 2(1), 8-28.

Vickrey, W.S. (1973) Pricing, metering, and efficiently using urban transport facilities, *Highway Research Record* 476, 36-48.

Vickrey, W.S. (1969) Congestion theory and transport investment, *American Economic Review (Papers and Proceedings)* 59, 251-261.

Wardman, M., P. Chintakayala, G.C. de Jong and D. Ferrer (2012) European wide meta-analysis of values of travel time; Paper prepared for EIB, ITS Leeds,

Wardman, M., R. Batley, J, Laird, P. Mackie, A.S. Fowkes, G. Lyons, J. Bates and J. Eliasson (2013) Valuation of travel time savings for business travellers, Report prepared for the Department for Transport, Institute for Transport studies, University of Leeds.