TI 2013-123/VIII
Tinbergen Institute Discussion Paper

Overreporting vs. Overreacting: Commuters' Perceptions of Travel Times

Stefanie Peer[1,2,3]
Jasper Knockaert[1]
Paul Koster[1,2]
Erik Verhoef[1,2]

[1] *Faculty of Economics and Business Administration, VU University Amsterdam;*
[2] *Tinbergen Institute;*
[3] *Institute for the Environment and Regional Development, Vienna University of Economics and Business, Austria.*

OVER-REPORTING VS. OVERREACTING: COMMUTERS' PERCEPTIONS OF TRAVEL TIMES

August 2014

Stefanie Peer[a,c,*], Jasper Knockaert[a], Paul Koster[a,b], Erik T. Verhoef[a,b]

[a]*Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam*
[b]*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam*
[c]*Institute for the Environment and Regional Development, Vienna University of Economics and Business, Welthandelsplatz 1, 1020 Vienna*

Abstract

We asked participants of a large-scale, real-life peak avoidance experiment to provide estimates of their average in-vehicle travel time for their morning commute. Comparing these reported travel times to the corresponding actual travel times, we find that travel times are overstated by a factor of 1.5 on average. We show that driver- and link-specific characteristics partially explain the overstating. Using stated and revealed preference data, we investigate whether the driver-specific reporting errors are consistent with the drivers' scheduling behavior in reality as well as in hypothetical choice experiments. For neither case, we find robust evidence that drivers behave as if they misperceived travel times to a similar extent as they misreported them, implying that reported travel times do neither represent actual nor perceived travel times truthfully. The results presented in this paper are thus a strong caveat against the uncritical use of reported travel time data in transport research and policy.

Keywords: travel time perception, reported travel times, valuation of travel time, departure time choices, peak avoidance experiment, panel latent class models, revealed preference (RP) data, stated preference (SP) data

*Corresponding author: speer@wu.ac.at (Stefanie Peer)

1. Introduction

Reported travel times are mostly used in situations where it is difficult, expensive or even impossible to measure travel times directly. In such situations, reported travel times usually act as a substitute for actual travel times. Moreover, in the spirit of the literature on time perception[1] reported travel times are sometimes used as an indicator of perceived travel times. Clearly, these two forms of using and interpreting reported travel times rely on different assumptions. In the first case, reported and actual travel times are assumed to be equal, whereas in the second case, reported and perceived travel times are assumed to be equal. This paper investigates the validity of such assumptions, and gives recommendations for the usage of reported travel time data.

Various earlier studies have suggested that reported travel times are not an appropriate indicator of actual travel times. Most of them found that individuals substantially overstate travel times (e.g. O'Farrell and Markham, 1974; Burnett, 1978; Henley et al., 1981; MVA Consultancy, 1987; Rietveld et al., 1999; Van Exel and Rietveld, 2009). Based on this evidence, they usually conclude that travel times are perceived as longer than they actually are. They thus interpret reported travel times as perceived travel times. Only few studies mention reporting errors (e.g. Van Exel and Rietveld, 2009) or the possibility that actual travel times might have been measured wrongly (e.g. Rietveld et al., 1999) as plausible explanations for the overstated travel times. Consistent with the results obtained in these earlier studies, also we find that travel times are strongly overreported, by a factor of 1.5 on average.

The implications of overstating for the usage of reported travel time data depend strongly on whether overstating is a consequence of reporting errors, or whether travelers indeed perceive travel times as longer than they actually are. In the first case, the main conclusion would be that reported travel times are untrustworthy, and should therefore not be used as a representation of travel times. Consequently, this would call for further research on improved methods of data collection. If the second case applied, the misperceptions would be expected to affect actual as well as hypothetical travel decisions in situations when travel times differ between choice alternatives. Brownstone and Small (2005) speculate that this may lead to biased estimates of the value of travel time (VOT)[2] if the VOT is derived from stated preference data. The VOT is a core parameter for the evaluation of transport policies. Hensher (2001), for instance, estimates the benefits (or costs) due to changes in travel times at 60% of total user benefits. The VOT is usually derived from stated preference (SP) or revealed preference (RP) data, which differ by the character of the underlying choice situations: SP data are collected from choice experiments where respondents are asked to decide between hypothetical travel alternatives, whereas RP data are collected from real-life choice situations. Brownstone and Small (2005) argue that if travel times are perceived as longer than they are, it is likely that in an SP setting travelers react to stated travel times as if these were overstated. This means that their response to a time difference stated in the SP experiment corresponds to how they would respond to a relatively smaller time difference in reality. Deriving the VOT based on the observed SP choices would therefore result in a relatively low estimate of the VOT. To give an example: an individual who perceives a real travel time of 10 minutes as one of 20 minutes probably reacts to a stated travel time of 20 minutes in an SP setting in the

[1] See for instance Grondin (2010) for a recent overview.

[2] Recent meta-analyses of relevant empirical research on the value of (travel) time (VOT) can for instance be found in Zamparini and Reggiani (2007), Small and Verhoef (2007), Shires and De Jong (2009) and Abrantes and Wardman (2011).

same way as he would react to a travel time of 10 minutes in reality; the VOT derived in the SP setting would therefore be half of the VOT that would be derived if RP data[3] were analyzed. If the SP-based VOT is then used for the appraisal of transport policies and thus applied to objectively measured changes in travel times, substantial biases can arise, not least because the benefits from travel time gains constitute such a major category in the appraisal of most transport policies.

In line with the hypothesis of Brownstone and Small (2005), it is a frequently reported outcome in the relevant literature that RP-based estimates of the VOT are higher than SP-based ones (e.g. Ghosh, 2001; Hensher, 2001; Brownstone and Small, 2005; Small et al., 2005; Isacsson, 2007). However, besides the explanation that drivers misperceive travel times, also other theories have been put forward. Most of them can be summarized under the label of 'hypothetical biases', which arise if people behave differently in an SP setting than in an RP setting (see for instance Louviere et al. (2000) and Carlsson (2010) for overviews). In particular, it has been speculated that people are more sensitive towards monetary attributes in hypothetical choice situations than in real life. More specifically, Brownstone and Small (2005) argue that such an increased sensitivity might be caused by time inconsistencies in actual behavior (for instance, by failing to plan departure times such that the reward is obtained) that are not taken into account in hypothetical scheduling choices. Recent research shows that the occurrence of hypothetical biases can be reduced by providing choice options to the respondents that resemble closely the choices they face in real life (e.g. Hensher, 2010).

In this study, we first analyze reported versus actual travel times. As mentioned before, we find a clear tendency of travelers to overreport travel times. Various specifications of 'travel time ratios', which represent the ratio of reported and actual travel times, are computed. We test whether the travel time ratios can be explained by driver- and link-specific characteristics. In a further step, the variation of the travel time ratios across individual drivers is used to investigate whether drivers misperceive travel times in their SP and RP travel choices to a similar extent as they misreport them. If we cannot confirm this, we must conclude that reported travel times are neither an appropriate representation of actual nor of perceived travel times, but instead are subject to substantial reporting errors, which have a systematic upward bias. Possible causes of such reporting errors include strategic behavior, social desirability biases, inaccurate recall, or a misunderstanding of the reporting task.

All data used in this study have been collected in the context of a real-life peak avoidance experiment that took place in the Netherlands for a period of 4 months. It included approximately 2000 participants who were able to gain a monetary reward for not using a specific highway link during morning peak hours. Their actual travel behavior along the highway link was monitored using cameras capable of number plate detection. In addition to these RP data, reported travel times were gathered from a questionnaire, and an SP experiment was conducted among the same set of drivers. While the observed behavior was only measured directly along the highway link relevant for the rewards, day- and time-of-day-specific door-to-door travel times are approximated by applying the method described in Peer et al. (2013). Their model combines speed measurements from loop detectors along the highway with GPS-based speed measurements on links for which

[3] If the VOT is derived from RP data, travel time misperceptions generally do not induce biased estimates of the VOT, at least as long as in the model estimations objective rather than subjective travel time measurements are used for defining the travel time attributes of the choice alternatives. If objective travel time measurements are used, the estimated hourly VOT thus represents the travelers' willingness to pay for a reduction of one hour of objective travel time, which is the relevant measure in policy appraisals: after all, changes in travel times due to policy interventions are almost always predicted in objective terms, for instance, from traffic assignment models.

no continuous speed measurements are available, using geographically weighted regression. In a nutshell, this real-life experiment provides a unique opportunity to compare reported and actual travel times on a door-to-door basis, and to test whether differences between them also manifest themselves in real and hypothetical travel choice behavior.

This paper unfolds as follows. Section 2 contains a review of relevant literature on (travel) time perception, and highlights the contribution of this paper. Section 3 describes the peak avoidance experiment and the travel time data used in this paper. Section 4 introduces the travel ratio. Section 5 discusses the theoretical framework that allows for testing whether the observed real-life and hypothetical scheduling choices are consistent with the reported travel times. Section 6 gives an overview of the SP and RP data used in the analysis. Section 7 contains the econometric framework and the estimation results of the choice models. Section 8 concludes on whether reported travel time data may be used as representations of actual and/or perceived travel times. Finally, the appendix provides an overview of the relevant variables and abbreviations used in the main text.

2. Related literature on (travel) time perceptions

To our best knowledge, this paper is the first study that investigates the difference between reported and actual travel times (reporting errors) as well as between reported and perceived travel times (perception errors), using actual and reported travel times as well as SP and RP travel choices from the same set of travelers. Related research has been conducted by Ghosh (2001) who tested whether reported travel times can explain travel behavior, beyond the extent explained by the actual travel times. He finds some evidence that the difference between reported and actual travel times has explanatory power, but also shows that the VOT barely changes when this difference term is included in the choice model. While Ghosh (2001) uses only RP data in his analysis on travel time misperception, we also use SP data. The latter allow us to test the hypothesis brought forward by Brownstone and Small (2005), possibly leading to an explanation why SP estimates of the VOT are often found to diverge strongly from their RP counterparts.

Various other studies in the area of transport economics have attempted to draw conclusions on travel time misperceptions. For instance, Brownstone and Small (2005) speculate that the common result of the VOT being higher under congested and slow traffic conditions (e.g. Recarte and Nunes, 1996; Wardman, 2001; Hensher, 2001; Zhang et al., 2005; Wardman and Nicolás Ibáñez, 2012) might be an indication that the time spent under such traffic conditions seems to pass slower, supposedly due to the annoyance with heavy traffic. However, a definite conclusion on this subject is difficult to obtain, since perception and valuation of travel times can usually not be meaningfully disentangled if only travel choices and actual travel times are known. We try to circumvent this problem by considering reported travel times and explicitly testing whether these reflect travel time perceptions.

In the analysis of travel choice behavior we emphasize the role of departure time decisions. Given that peak hour congestion poses a major problem in most urban areas, scheduling models have gained an increasing amount of interest over the last few decades, as they are able to shed light on how transport policies affect departure time choices, and therefore, also on peak congestion as a whole (e.g. Vickrey, 1969; Small, 1982; Noland and Small, 1995). Empirically, departure time decisions are usually represented as choices between discrete departure time alternatives. The respective models are then estimated using discrete choice analysis (McFadden, 1974; Train, 2003). Besides using standard multinomial logit (MNL) models, we also employ panel latent class models, which are able to account for unobserved heterogeneity between individuals as well as the panel nature of the underlying data. Moreover, we do not only estimate separate models on the SP and RP data, but also models that combine these two data sources. Only few studies on travel choice behavior have so far been undertaken with both data sources as input (e.g. Ghosh, 2001; Small et al., 2005; Brownstone and Small, 2005; Börjesson, 2008), mostly using fairly simplistic travel time measurements in their RP models.[4] Contrastingly, the estimated door-to-door travel times used in this study allow us to account for travel time stochasticity and driver-, day- and time-of-day specific travel time expectations.

This paper is also related to literature on time perception in general. Studies belonging to this field of research commonly assume that reported and perceived durations are equal. This may be a reasonable assumption under controlled experimental conditions and for short time spans[5], while it is more questionable in situations where the analyst has less control over the experiment and where durations are generally longer (as it is the case in this paper). Evidence on when durations tend to be over- or underestimated varies substantially across studies. One main consensus is that the perception of durations becomes increasingly inaccurate if the cognitive load during the time interval under question is high, supposedly because fewer mental resources are available for 'temporal processing' (Block et al., 2010). This finding has been confirmed also for the context of car travel (Baldauf et al., 2009). Another pattern relevant to travel is that familiar tasks (such as commuting) tend to be perceived as shorter than they are (e.g. Boltz et al., 1998). This effect might, however, be offset by another one: It has been found that activities that are not well predictable, which holds true for most car trips, are perceived as longer than they actually are (e.g. Boltz, 1998). Also the role of emotions is frequently pointed out. A main finding is that time passes slower under very stressful conditions (e.g. Droit-Volet and Meck, 2007). The evidence is more mixed when it comes to situations with very low mental arousal: While some studies claim that in such situations time seemingly passes faster (e.g. Block and Zakay, 1996), others report the opposite finding (e.g. Flaherty, 1999; Glicksohn, 2001). Based on these findings, we must conclude that the literature on time perception does not provide an unequivocal suggestion on whether travel time (specifically morning commutes) is expected to be perceived as passing slower or faster than it actually does.

3. Data

3.1. Peak avoidance experiment

All data used in this paper have been collected in the context of a peak avoidance experiment (*Spitsmijden* in Dutch). The experiment took place in the Netherlands along a 9.21 km long stretch of the A12 highway leading to The Hague. This link was confined by two cameras (C1, C2) that were used for number plate detection. We therefore refer to its start and end location as C1 and C2, respectively. While the entire experiment lasted for more than a year, the analysis conducted in this paper focuses on the time frame between September 2009 and December 2009 for reasons of data availability. Approximately 2000 commuters participated in the experiment during that time period. Before entering the experiment, their average number of weekly trips through the C1–C2 link was measured and defined as reference behavior. Between September and December, the participants were able obtain a reward of 4 Euro for each avoided trip along the C1–C2 link during morning peak hours (6:30-9:30 am), relative to their reference behavior. A more detailed description of the experiment can be found in Knockaert et al. (2012).

[4] See Peer et al. (2013) for a discussion on possible biases resulting from imprecise travel time measurements.
[5] Some studies on time perception cover time intervals of only few seconds.

3.2. Descriptives participants

In this paper, we only consider a subset of 443 out of the 2000 participants of the peak avoidance experiment, mainly because reported travel times and preferred arrival times at work (PAT) are only available for those who filled in a specific questionnaire (ca. 600). Moreover, we are only able to consider those participants in the analysis for whom we have sufficient data to approximate door-to-door travel times according to the methodology developed by Peer et al. (2013), which will be described more closely below.[6]

Table 1 shows some descriptive statistics for these 443 participants. The participants are mostly males between the age of 30 and 50 years, and tend live in households with a relatively high net income (the average Dutch household could dispose of 2760 Euro/month in 2008[7]). Moreover, Table 1 demonstrates that the distribution of the preferred arrival time (PAT) at work is strongly clustered between 7:30 and 9:00 am.

3.3. Reported travel times

The main data source for reported travel times is an online questionnaire, which was conducted among the participants of the peak avoidance experiment in November 2009. It resulted in a response rate of approximately 30%. Reported travel times are based on the answers provided by the participants when asked to state their average travel times on the three segments composing their door-to-door trip: the home–C1, C1–C2 and C2–work links. Reported travel times were thus collected at one time instance only. For their answers, respondents were asked to take into account their most recent 20 morning commute trips. To prompt them to provide realistic answers, maps of the location of the C1–C2 link were shown next to the questions. It was also emphasized that only in-vehicle times should be reported.

We define reported average home–work travel times as the sum of the average travel times indicated for each of the three sub-links. Reported travel times are denoted by T^R_{zl}, where l indicates the link $l = \{\text{Home–C1, C1–C2, C2–Work, Home–Work}\}$ and $z = 1, \ldots, Z$ indicates the participant.

In addition to the above 'standard' definition of reported travel time, we introduce an alternative measure to to investigate whether the 'standard' measure is biased because it is based on the sum of travel times along the sub-links rather than total home–work travel times. The alternative measure of reported (home–work) travel times, labelled 'total', is derived from a survey that was conducted at the time when the participants entered the experiment. They were asked to state their average home–work travel time, without distinguishing between sub-links.

3.4. Actual travel times

Multiple data sources are used for defining actual travel times. The travel times through the C1–C2 link are observed directly, and hence individually for each driver, using number plate detection at the beginning and at the end of the link. Home–C1 and C2–work travel times, however, need to be approximated, as no camera observations are available. The approximation method is described in detail in Peer et al. (2013). It involves geographically weighted regression (GWR), which is used

[6]The sample used for the choice models will be smaller (406), as we additionally exclude drivers who stated that they made their choices randomly in the stated preference experiment as well as those who indicate to have a preferred arrival time at work outside peak hours.

[7]http://statline.cbs.nl/StatWeb/publication/?DM=SLNL&PA=71511ned&D1=a&D2=a&D3=0&D4=a&VW=T

Table 1: Descriptive statistics of the subset of participants considered in the analysis

Variables	Value or Fraction of Sample
Nr. of individuals	443
Socio-economic characteristics	
Age	
$<$ 30 years	0.10
$>$ 50 years	0.11
Female	0.26
Monthly (net) household income	
$<$ 3500 Euro	0.33
$>$ 5000 Euro	0.15
unknown	0.22
Households with kids	
$0 - 5$ years	0.21
$6 - 10$ years	0.21
$11 - 15$ years	0.17
Commute-related variables	
Preferred arrival time (PAT)	
<7:30 am	0.16
>9:00 am	0.13

to measure the extent of correlation between speeds on the C1–C2 and home–C1 as well as C2–work links, respectively. Speeds on the C1–C2 link are measured continuously by loop detectors, whereas a limited number of GPS-based speed measurements is available along the home–C1 and C2–work links. As a result, day- and time-of-day-specific travel times can be estimated for almost all combinations of home and work locations in the dataset. Peer et al. (2013) demonstrate that this method yields reliable travel times predictions.

Actual travel times are denoted by T^A_{zld}, where again l denotes the link and z the driver. Moreover, $d = 1, \ldots, D_z$ is the index attached to the observed passages of the C1–C2 link, where $d = 1$ indicates the most recent trip before filling out the survey, $d = 2$ the second most recent trip, continuing up to D_z, which thus denotes the overall number of observed passages between September 1, 2009 and the date driver z has completed the questionnaire. For an observed passage to be considered in this list of past trips, certain conditions must be fulfilled. In the 'standard' definition, passages of the C1–C2 link between 6:00 and 11:00 are considered commuting trips. Besides that, we introduce an alternative definition of this time interval to test whether the 'standard' definition takes into account an excessively broad peak, which may lead to a downward bias of actual travel times (as travel times tend to be shorter at the margins of and outside the peak period) and thus to an upward bias of the travel time ratio. The alternative definition, which is labelled 'peak', only takes into account trips that result in a passage of the C1–C2 link between 6:30 and 9:30 am. This coincides with the interval during which rewards were granted in the experiment.

When reporting the average travel times in the questionnaire, participants were asked to consider their last 20 commuting trips. For 69% of the participants, more than 20 commuting trips have been been recorded before the date when the questionnaire was filled in.[8] For these partic-

[8]The average number of observed trips is 33.7.

ipants only the 20 most recent ones enter the analysis. For the remaining participants with 20 trips or less, we only take into account the available observations. The average actual link- and driver-specific travel time $\bar{T}^A_{zl}$ is thus defined as:

$$\bar{T}^A_{zl} = \frac{1}{\min[20, D_z]} \sum_{d=1}^{\min[20,D_z]} T^A_{zld} \qquad (1)$$

4. The travel time ratio

4.1. Specification

This section focuses on the computation of driver- and link-specific travel time ratios.[9] We define the travel time ratio τ_{zl} as the ratio between the reported average travel time and the average actual travel time. The travel time ratio is therefore larger than 1 for all cases of overstating, and between 0 and 1 for cases of understating. An identical formulation has for instance been used by Parthasarathi et al. (2013) for comparing reported and observed travel times.

$$\tau_{zl} = T^R_{zl} / \bar{T}^A_{zl} \qquad (2)$$

Some studies on time perceptions have adopted a power law function in order to control for nonlinearities in the relation between reported and actual durations (e.g. Eisler et al., 2007). However, most of them find that the exponential term is close to 1, implying that the non-linearities are small (e.g. Eisler, 1976; Allan, 1979). Leiser and Stern (1988) obtained similar results for travel times, suggesting that the linear formulation performs almost as well as the power law function. While we find some non-linearities to be present in our data (see the regression results in Table 4), we do not adopt a more complex specification, mainly because our main interest in this research is to use the travel time ratio as a simple input to departure time choice models.

4.2. Descriptives

Table 2 provides the descriptive statistics for the reported and observed average travel times as well as for the travel time ratios. The two bottom lines of the table show the alternative measures of the reported ('total') and observed ('peak only') home–work travel times. In total, valid measures for both reported and actual travel times for the 'standard' definitions are derived for 443 respondents.

Overstating is found to be a persistent phenomenon for all sub-links as well as for all definitions of the travel time ratio, with its average ranging from 1.55 to 1.77 (the median is slightly lower, ranging from 1.51 to 1.63). The first three rows of Table 2 demonstrate that the average travel time ratio exceeds 1 on all sub-links, hence, rejecting the hypothesis that (on average) drivers overreport travel times only on some of the sub-links (e.g. those affected most by congestion) and compensate this by understating travel times on the remaining links of the commute.[10] A comparison of the 'total' definition of the travel time ratio to the 'standard' one reveals that the extent of overstating is even higher if drivers are asked for the total average home–work travel time

[9]Note that in the literature on time perception this ratio is also referred to as duration judgement ratio (e.g. Block et al., 2010).

[10]The finding that the average home-work travel time ratio is smaller than the corresponding travel time ratios on the sub-links is driven by some outliers on the sub-links.

Table 2: Reported vs. observed travel times (in minutes)

			Reported		Observed		τ_{zl}		
Link	Label	Nr. Obs.	Mean	St. D.	Mean	St. D.	Mean	Median	St. D.
H–C1	–	443	25:56	13:38	17:24	11:01	1.75	1.57	0.71
C1–C2	–	443	15:51	6:03	9:54	2:38	1.66	1.56	0.66
C2–W	–	443	16:13	7:34	10:21	4:10	1.77	1.53	0.99
H–W	stand.	443	58:00	18:50	37:40	12:17	1.58	1.55	0.37
H–W	total	434	60:03	16:21	37:40	12:17	1.68	1.63	0.42
H–W	peak	398	58:26	19:00	38:35	12:06	1.55	1.51	0.37

rather than the average travel times on the sub-links (1.68 vs. 1.58). We can thus conclude that the observed extent of overstating for the 'standard' definition is not biased upwards as a consequence of letting participants report their travel times for the three sub-links rather than for their home–work journey. Finally, we find that the travel time ratio for the 'peak' definition is almost as high as for the 'standard' definition (1.55 vs. 1.58). The overstating found for the 'standard' definition can thus not be attributed to an excessively broad definition of the morning peak. We therefore conclude that the 'standard' formulation of the travel time ratio is an appropriate representation of the extent of travel time overstating; that is, it does not deviate in any substantial way from alternative measures that were available.

4.3. Explaining variation across participants and links

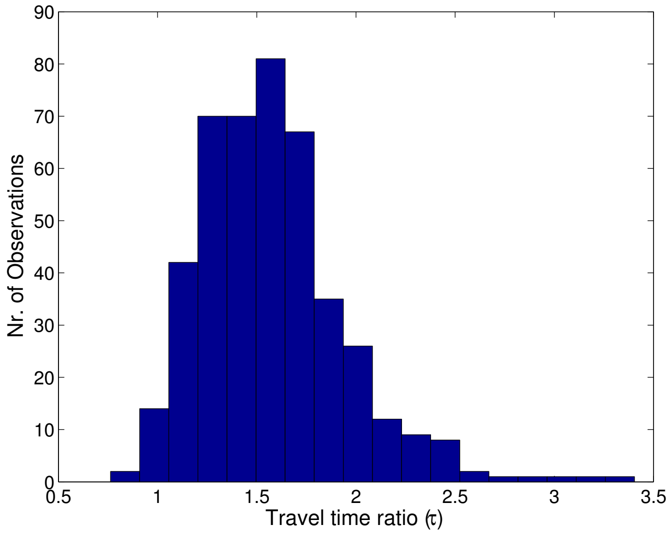
The travel time ratio varies considerably across persons, as Figure 1 illustrates. It gives a histogram of the person-specific home-work travel time ratio, based on the 'standard' definition. Only few respondents have a travel time ratio less than 1, and therefore understate travel times. The figure also demonstrates that the distribution of the travel time ratio is slightly skewed to the right, with a median of 1.55, a mean of 1.58 and few outliers towards the right end of the distribution.

We seek to explain the variation in the travel time ratio across drivers and links by regressing the 'standard' travel time ratios on driver and link-specific characteristics. The descriptive statistics of the variables considered in the regressions are provided in Table 3. In addition to these, we also tested the significance of various socio-economic variables (age, gender, education, income, children, flexibility of working hours), alternative measures of travel frequency (as a proxy for experience), measures of travel time variability (specifically, percentile differences and standard deviations), characteristics of the most recent trip, and person-specific outliers in terms of travel times. However, all these variables were found insignificant.[11] For various other variables that might affect the travel time ratio, such as road type or free-flow speed, no data were readily available.

The first model presented in Table 4 uses the travel time ratios for the home-work link as dependent variable, whereas the second model uses the travel time ratios of the three sub-links. The former includes person-specific explanatory variables as well as variables that characterize

[11] Our finding that the travel time ratio is independent of income differs from the result obtained by Burnett (1978), who concludes that individuals with low income overstate travel times the most, supposedly due to lower education levels or less experience in traveling.

Figure 1: Histogram travel time ratio (Home–Work)



home–work trips. In addition to those, the second model also accounts for variables that are defined at the level of the sub-links. A standard ordinary least squares (OLS) estimation is conducted for Model 1. For Model 2, a random effects (RE) regression is used in order to account for the panel structure of the dataset which follows from the three sub-links defined for each driver.

The estimations result in an R-square of 0.17 for the OLS and of 0.28 in the RE model, meaning that a considerable share of the travel time ratio can indeed be explained by driver- and link-specific characteristics. The first model shows that the home–work travel time ratio is relatively high for commuters with short commuting distances and high average speeds. Also, we find some evidence that those drivers who have little experience in commuting[12] tend to overstate travel times more, and that those with a relatively large share of trips outside the peak[13] tend to overstate travel times less. These results are confirmed at the link level, too.

The relatively high extent of overstating of travel times along the home–C1 and C2–work links might be explained by the fact that drivers consider also the time required to walk to the car and from the car to the office in their answers, although we asked them explicitly to only consider in-vehicle commuting time. This effect is probably more predominant on the home–C1 and the C2–work links than on the intermediate C1–C2 link. Since the home–C1 and the C2–work links tend to be longer than the C1–C2 link (see Table 3), this may also explain why longer link lengths tend to be associated with a higher travel time ratio. Moreover, since it is reasonable to assume that the time spent on 'out-of-vehicle commuting' is independent from the commuting distance, it is also a possible explanation for the higher travel time ratio for shorter commutes. The finding that travel times on sub-links with high speeds (relative to the average speed for this person) are overstated more strongly might result because drivers know the length of the sub-links (for instance from traffic signs installed along the roads). When prompted to indicate travel times on the sub-links, they might attribute their perceived total travel times to the sub-links more or less

[12]This variable is defined a dummy that assumes the value 1 if a driver has undertaken less than 4 trips since the start of the reward period in September 2009.

[13]Peak trips are defined as trips that lead to passages of the C1–C2 link between 6:30 and 9:30.

Table 3: Descriptive statistics of the variables considered in the regressions

Variable	Specification	Mean	St. Dev
Driver-specific variables			
Nr. of trips	within the last 20 working days	5.08	3.68
Share of trips outside peak	since September 1, 2009	0.20	0.29
Link- (& driver-) specific variables			
Mean speed home–work	in km/h	69.63	11.29
Mean speed home–C1	in km/h	72.93	17.36
Mean speed C1–C2	in km/h	66.96	13.71
Mean speed C2–work	in km/h	63.06	11.70
Distance home–work	in km	43.30	17.13
Distance home–C1	in km	22.88	15.92
Distance C1–C2	in km	9.21	–
Distance C2–work	in km	11.21	4.86

Table 4: Regression results: Explanation of the travel time ratio

	Dependent Variable			
	$\tau_{\text{Home–Work}}$		$\tau_{\text{Home–C1,C1–C2,C2–Work}}$	
Variable	Coefficient	t-stat.	Coefficient	t-stat.
Constant	1.33	12.24	2.10	9.38
Little experience	$5.30*10^{-2}$	1.51	$9.07*10^{-2}$	1.97
Share of trips outside peak	-0.28	-4.32	-0.37	-4.29
Mean speed home–work (MSHW)	$1.10*10^{-2}$	5.41	$9.90*10^{-3}$	3.67
Distance home–work (DHW)	$-1.11*10^{-2}$	-8.77	$-2.85*10^{-2}$	-11.08
Mean speed link/MSHW	–	–	0.91	5.76
Distance link	–	–	$4.30*10^{-2}$	7.31
Distance link/DHW	–	–	-4.47	14.70
dummy home–C1	–	–	0.36	6.49
dummy C2–work	–	–	0.18	4.09
Estimation Method	OLS		RE	
Nr. of Observations	443		1329	
R-square	0.17		0.28	

proportionally to their lengths, not taking into account the differences in speeds across links.

Based on comparing actual and reported travel times we are not able to draw conclusions on travel time perceptions, since we have not yet investigated whether reported travel times reflect travel time perceptions. For now, we can only conclude that some, but not all findings from the regressions are consistent with the literature on (travel) time perceptions. For instance, the result that drivers with little commuting experience overstate travel times more is in line with the research on time perception, which typically finds that the duration of familiar tasks is perceived as shorter than of unfamiliar ones (e.g. Boltz et al., 1998). Also, one might argue that the findings are consistent with some of the literature on the relation between mental arousal and time perception: the regression results support the hypothesis that time is perceived as passing more slowly during periods with presumably low mental arousal, namely when speeds are high, and congestion and interruptions (e.g. due to traffic lights) are thus limited. On the contrary, the finding that travel time variability turned out insignificant contradicts earlier studies: the relevant literature on time perception predicts a positive correlation between unpredictability and the perceived duration (e.g. Boltz, 1998).

5. Choice models: Theoretical framework

5.1. Introduction

This section introduces the theoretical framework required to test whether differences between reported and actual travel times influence travel choices. We model travel choices using the scheduling model of Vickrey (1969). He was the first to define departure time decisions as a result of trade-offs between travel times and schedule delays, which characterize the extent of earliness and lateness with respect to an exogenous preferred arrival time (PAT). His so-called 'bottleneck model' was later on extended along theoretical (e.g. Arnott et al., 1993, 1994) as well as empirical lines of research (e.g. Small, 1982; Noland and Small, 1995). Small (1982) was the first to estimate monetary valuations of travel time and schedule delays, whereas Noland and Small (1995) were the first to incorporate travel time variability in the scheduling model.

Consistent with this literature, our model assumes that commuters choose their optimal departure time from home, trading-off expected travel times (T), schedule delays early (SDE) and late (SDL), and monetary trip costs (in this study, monetary rewards R). The schedule delays early and late are defined by:

$$\begin{aligned} \mathrm{SDE}(t) &= \max[\mathrm{PAT} - t - T(t),\ 0] \\ \mathrm{SDL}(t) &= \max[t + T(t) - \mathrm{PAT},\ 0], \end{aligned} \tag{3}$$

where $T(t)$ denotes the travel time associated with departure time from home t. For simplicity, we drop the indices related to individual z, choice alternative j (which is associated with a departure time t) and choice situation k. We allow for stochasticity of travel times, which is made explicit by using the expectation operator $\mathbb{E}$ with respect to the attribute values. The marginal utilities associated with the reward, travel time and schedule delay early and late are then denoted by $\beta_R, \beta_T, \beta_E$ and β_L. We assume an additive form of the expected utility function:

$$\mathbb{E}V(t) = \beta_R * \mathbb{E}R(t) + \beta_T * \mathbb{E}T(t) + \beta_E * \mathbb{E}SDE(t) + \beta_L * \mathbb{E}SDL(t) \tag{4}$$

The values of time (VOT) and schedule delay early (VSDE) and late (VSDL) can be defined as (-1) times the ratio of the time and schedule delay coefficients and the reward coefficient. The resulting values represent the willingness to pay for reducing expected travel time and schedule delays:

$$\text{VOT} = -\beta_T/\beta_R \qquad \text{VSDE} = -\beta_E/\beta_R \qquad \text{VSDL} = -\beta_L/\beta_R \tag{5}$$

5.2. RP model

The systematic part of the utility function associated with RP-based departure time choices includes the additive terms comprised in $\mathbb{E}V(t)$ (see Eq. 4). In the RP context we denote the corresponding utility function by $\mathbb{E}V^{RP}(t)$. In the model that is used to test whether the drivers behave in real life as if the travel times they take into account in their decision making are consistent with the person-specific travel time ratio, additional terms must be added to the utility function, which is then denoted by $\mathbb{E}V^{\tau,RP}(t)$: For all four attributes – rewards, travel times and schedule delays early and late – the difference between the attribute of the choice alternative that would result if drivers made a perception error consistent with their travel time ratio and the corresponding objectively measured attribute values must be included. We compute these terms for all attributes as it is implausible that travelers overestimate travel times when considering trip duration, but not when considering the other travel-time-dependent attributes, notably schedule delays and rewards. A driver who overestimates travel times is thus expected to underestimate schedule delays early and overestimate schedule delays late; for a given departure time this driver expects to arrive later than a person who departs at the same moment and does not overestimate travel times. The coefficients corresponding to the difference terms as well as the attribute values that would result if travel times were misperceived in accordance with the (individual-specific) travel time ratio are labelled with the superscript τ:

$$\begin{aligned}\mathbb{E}V^{\tau,RP}(t) =&\beta_R * \mathbb{E}R(t) + \beta_R^{\tau} * (\mathbb{E}R^{\tau}(t) - \mathbb{E}R(t)) + \\ &\beta_T * \mathbb{E}T(t) + \beta_T^{\tau} * (\mathbb{E}T^{\tau}(t) - \mathbb{E}T(t)) + \\ &\beta_E * \mathbb{E}SDE(t) + \beta_E^{\tau} * (\mathbb{E}SDE^{\tau}(t) - \mathbb{E}SDE(t)) + \\ &\beta_L * \mathbb{E}SDL(t) + \beta_L^{\tau} * (\mathbb{E}SDL^{\tau}(t) - \mathbb{E}SDL(t))\end{aligned} \tag{6}$$

The τ-related coefficients are expected to be insignificant if drivers take into account the objective attribute levels in their decision making rather than those attribute levels associated with overestimation (or, in few cases underestimation). On the contrary, if $\beta_R^{\tau}, \beta_T^{\tau}, \beta_E^{\tau}, \beta_L^{\tau}$ have the same sign and are roughly equal in size to $\beta_R, \beta_T, \beta_E, \beta_L$, respectively, we can conclude that the behavior of the individuals in real life is consistent with the travel times they reported, and, in particular, the deviations of the reported from the actual travel times. As a consequence, we could interpret reported travel times as a valid indicator of travel time perceptions.

5.3. SP model

Also for the SP model we distinguish between $\mathbb{E}V^{SP}(t)$ and $\mathbb{E}V^{\tau,SP}(t)$. The former again includes the additive terms comprised in $\mathbb{E}V(t)$ (Eq. 4), whereas the latter also accounts for the difference between the travel time attribute that would result if travel times were affected by a perception error consistent with the (individual-specific) travel time ratio and the objectively measured travel time attribute. In contrast to the RP model, this difference term is only considered for the travel time attribute, but not for the reward and scheduling attributes. The reason is that the SP experiment presents rewards as well as actual arrival times at work to the respondents.[14]

[14]The exact setup of the SP experiment will be discussed in Section 4.2, and is shown graphically in Figure 2.

Therefore, it is unlikely that respondents perceive the reward and the schedule delays differently from the ones that are presented to them, even if they misperceive travel times.

$$\begin{aligned}\mathbb{E}V^{\tau,SP}(t) =& \beta_R * \mathbb{E}R(t) + \beta_T * \mathbb{E}T(t) + \beta_T^{\tau} * (\mathbb{E}T^{\tau}(t) - \mathbb{E}T(t)) + \\ & \beta_E * \mathbb{E}SDE(t) + \beta_L * \mathbb{E}SDL(t)\end{aligned} \tag{7}$$

The interpretation of β_T^{τ} is similar to its interpretation in the RP utility function, with the difference that here we expect a positive β_T^{τ} in the case of overestimation (whereas β_T is expected to be negative, as in all models): the more one overestimates travel times, the smaller the 'real' duration that corresponds with a stated X amount of time, and hence the less negative the impact of that X amount of time on utility should be (for instance, respondents with a travel time ratio of 1.5 are expected to react to a 15-minute delay in an SP setting in the same way as they would react to a 10-minute delay in reality). If this is found true, the monetary value attached to the difference between the travel times multiplied by the travel time ratio and the objectively measured ones (i.e. $\text{VOT}^{\tau} = -\beta_T^{\tau}/\beta_R$) would be negative, hence, lowering the overall value attached to travel time ($\text{VOT}+\text{VOT}^{\tau}$) as suggested by Brownstone and Small (2005).

6. Choice models: Data

6.1. RP data

RP data have the advantage that they are based on real choices rather than hypothetical ones. They are therefore exempt from the suspicion to suffer from hypothetical biases, which arise if people choose differently in a laboratory setting than in real life. This benefit comes with the disadvantage that RP data are usually more difficult and expensive to collect than SP data and that they may suffer from strong correlations between the model variables. Moreover, the attribute values of RP choices as well as the choice set itself are often ambiguous, and thus difficult to identify. Clearly, in RP studies, the ranges of the attribute values are limited to the values that exist in reality, implying that responses to non-existent alternatives (e.g. new transport modes) cannot be measured (e.g. Swait et al., 1994).

RP data are defined in a similar way as the actual travel times used to compute the travel time ratio. The main difference is that for the choice analysis not only the attributes of the chosen departure time alternatives but also of the unchosen alternatives need to be known. Hence, for each departure time alternative j for driver z and choice situation k, expected rewards, travel times and schedule delays need to be derived. For the estimation of door-to-door travel times we use the model developed by Peer et al. (2013). We also use this model to approximate the chosen departure time at the home location. Choice situations refer here to days during which drivers have been observed to pass the C1–C2 link, along which the reward experiment took place. We therefore describe the departure time choice conditional on making a commute trip and passing the C1–C2 link, and the overall number of RP choice situations is therefore driver-specific. Departure time is formulated as a discrete variable. The choice set consists of 17 distinct 15-minute intervals during the morning peak (ranging from 5:30 to 9:45 am). Observed departures outside this time frame are dropped. This is because trips that take place outside this interval are likely to be undertaken for purposes other than regular commuting, meaning that other scheduling preferences may apply than for commuting trips.

The (person-specific) preferred arrival time (PAT) for regular commuting trips is derived from the same questionnaire as the reported travel times. It is defined as the preferred moment of arrival at work if there was no congestion (ever). Only drivers with a PAT between 6:30 and 9:30 are taken into account, since drivers with very early or late PATs barely face any trade-off between schedule delays and travel times.

Rewards vary by time of the day. They are equal to 4 Euro if a driver passes C2 before 6:30 or C1 after 9:30, and 0 otherwise. Moreover, the maximum number of rewarded trips per 2-week-period cannot exceed the biweekly number of peak trips that a driver has undertaken before the start of the experiment.[15] Once the maximum number of rewards has been reached, no further rewards are distributed during the rest of the week. Holidays and weekends are excluded from the analysis, mainly because trips observed on these days frequently do not constitute commuting trips.

Using the PAT, the rewards, and the driver-, day- and time-of-day-specific door-to-door travel times, the attribute levels specific to the choice situation k, individual z and choice alternative j can be derived. Expected levels of the attributes $A \in \{R, T, SDE, SDL\}$, $\mathbb{E}A$, are determined based on a compound measure of actual, time-of-day-specific attribute levels on the day of travel, A_{zkj}, and the time-of-day-specific average attribute levels, $\bar{A}_{zj}$ over the duration of the experiment.[16] Differences between A_{zkj} and $\bar{A}_{zj}$ are exclusively driven by fluctuations of travel times across days (for a given departure time). While actual travel time realizations are in reality unknown at the moment when the departure time decision is taken, they represent an upper benchmark for the maximum extent of information potentially available to drivers. The average travel times represent the long-run travel time pattern over the time of the day, which participants of the experiment are likely to be aware of.

We estimate the weight θ that drivers attach to the actual realizations relative to the averages; θ will be equal to 0 if travel time expectations are based on the averages only, while it will be equal to 1 if drivers are perfectly informed about actual realizations and do not consider the averages in their decisions. The expected attribute of the utility function for driver z, choice situation k and choice alternative j, $\mathbb{E}A_{zkj}$, is thus given as follows:

$$\mathbb{E}A_{zkj} = \theta * A_{zkj} + (1 - \theta) * \bar{A}_{zj} \tag{8}$$

Next, we apply the 'standard' definition of the travel time ratio for the home–work link to compute the attribute values that would result if drivers misperceived travel times to an equal extent as they misreported them: $\mathbb{E}A^{\tau}_{zkj}$. For this purpose, we first multiply the time-of-day- and day-specific actual travel times on each of the sub-links (home–C1, C1–C2, C2–work) by the person-specific 'standard' travel time ratio.[17] Based on the resulting travel times we compute the corresponding rewards and schedule delays. We denote the resulting time-of-day- and day-specific attribute levels by A^{τ}_{zkj}. By averaging $A^{\tau}_{zkj} = \{R^{\tau}_{zkj}, T^{\tau}_{zkj}, SDE^{\tau}_{zkj}, SDL^{\tau}_{zkj}\}$ over the relevant working days, we obtain the remaining component, $\bar{A}^{\tau}_{zj}$, of the compound attribute expectation measure EA^{τ}_{zkj} similarly. We can then compute $\mathbb{E}A^{\tau}_{zkj}$ similarly as in Eq. 8:

$$\mathbb{E}A^{\tau}_{zkj} = \theta * A^{\tau}_{zkj} + (1 - \theta) * \bar{A}^{\tau}_{zj} \tag{9}$$

[15] The reference behavior was measured without knowledge of the participants.

[16] Tseng et al. (2013) use a similar definition of expected travel times. However, instead of applying the compound measure to all attributes of the utility function, they apply it to travel times only.

[17] Note that we do not differentiate between the different magnitudes of the travel time ratio on the various sub-links, mainly because of outliers on the sub-links, which are less evident when considering the entire home–work stretch.

6.2. SP data

Most SP experiments on scheduling behavior ask respondents to make hypothetical choices between alternatives with varying departure times, which may differ in terms of costs, travel time and travel time variability. Because the researcher determines the attribute levels, problems of collinearity between attribute levels can be more easily avoided in SP settings than in RP settings. Furthermore, there is no ambiguity with respect to the attribute values and the definition of the choice set, as both are provided to the respondents directly in the SP experiment. On the negative side, SP-based estimates may be affected by the attribute values presented to the respondents as well as by the format and complexity of the choice task. And most importantly, they may be biased due to the hypothetical character of the choices (e.g. Swait et al., 1994). However, recent research has shown that these biases can be reduced by designing the SP experiments such that the realism of the choices and attribute levels is enhanced; for instance, by pivoting the design values around the status quo behavior of respondents (Hensher, 2010). This strategy has been adopted also in the SP experiment considered in this paper. All travel times shown to the respondents in the SP experiment have been designed such that they are always situated between the minimum and maximum travel time reported by the driver, and the preferred arrival time shown to them is identical to their reported PAT.

The SP choice experiment consists of 10 choices between 2 departure time alternatives. Travel time variability was considered by stating two possible travel time realizations for each departure time alternative, each of them occurring at random with a presented probability. For a given departure time, the variation in travel times also induces variations in schedule delays and, in some cases, in the reward. A reward of up to 8 Euro applies to passages of the C1–C2 link between 6:30 and 9:30. In the estimation of the SP choice models, the attribute values of each alternative are defined as weighted average across the two possible travel time realizations. Figure 2 shows an example of the choice screen.

Figure 2: Screenshot

Your preferred arrival time if there is no delay is: 8:40.

	Alternative A		**Alternative B**	
Departure time from home	**6:05**		**6:50**	
Probability	***80%***	***20%***	***90%***	***10%***
Total travel time	**30 min**	**40 min**	**20 min**	**35 min**
Travel time from home to camera A	15 min	15 min	10 min	10 min
Travel time from camera A to camera B	5 min	10 min	5 min	15 min
Travel time from camera B to work	10 min	15 min	5 min	10 min
Arrival time at work	**6:35**	**6:45**	**7:10**	**7:25**
Reward	**4 euro**	**4 euro**	**0 euro**	**0 euro**

After filling in the choice experiment, respondents were asked to indicate whether they answered some or all choice questions at random. Those who confirmed they did (2% of the respondents), were excluded from the data. The efficiency of the choice experimental design was thoroughly tested using extensive simulation in order to assure that a broad range of parameters can be reproduced (e.g. Koster and Tseng, 2010).[18] A detailed description of the design of the SP experiment can be found in Knockaert et al. (2012).

[18]Moreover, the layout and the phrasing was tested using focus groups and an online test.

Finally, to compute the travel time attributes that would result if drivers misperceived travel times to the same extent as they misreported them, the travel times shown in the choice experiment are multiplied by the 'standard' definition of the person-specific travel time ratio for the home-work link.

7. Choice models: Estimations

7.1. Econometric framework

We estimate standard multinomial logit (MNL) models and more advanced panel latent class models that take into account heterogeneity between individuals as well as the panel nature of the data. We estimate MNL models for SP and RP data separately, and a latent class model that pools the two data sources. For notational convenience, this section introduces the pooled version of the models.

Usually, the underlying notion of a joint analysis of SP and RP data is that SP data are able to correct for deficiencies in the RP data such as correlations between attributes or the lack of identification of some attributes or attribute ranges, while keeping the realism inherent to the RP data (e.g. Louviere et al., 2000). Estimations that pool RP and SP data are most advantageous when the RP and SP choice situations are similar and are also perceived similarly by the participants (e.g. Börjesson, 2008). In that case, common coefficients for both SP- and RP-based observations may be possible if individuals are found to react in consistent ways to trade-offs in RP and SP choice situations.

We define the expected utility $\mathbb{E}U_{zkj}$ as the utility associated with choice situation $k = 1, \ldots, K_z$, individual $z = 1, \ldots, Z$ and choice alternative $j = 1, \ldots, J_{zk}$. It consists of the systematic component as defined in Eqs. 4 (and in Eqs. 6 and 7 for the cases when the travel time ratio is considered) and a stochastic component ϵ_{zkj}. The panel is unbalanced since the number of RP choices differs across drivers. Also, the number of available alternatives, J_{zk}, differs across choices. It equals 2 for SP choices and 17 for RP choices. The indicator function $\mathbf{1}^{RP}$ is used to distinguish RP and SP choices, and hence the corresponding systematic parts of the utility function. It is equal to 0 for SP choices ($k \in \{1, \ldots, 10\}$) and equal to 1 for RP choices ($k \in \{11, \ldots, K_z\}$). Moreover, any difference in the variance of the error term between SP and RP observations is taken into account by defining a multiplicative scale factor λ that is relevant for SP observations (while the scale of RP observations is fixed to 1).[19] The random utility function to be maximized is then given as follows (note that $\mathbb{E}V_{zkj}^{RP}$ may be replaced by $\mathbb{E}V_{zkj}^{\tau,RP}$, and $\mathbb{E}V_{zkj}^{SP}$ by $\mathbb{E}V_{zkj}^{\tau,SP}$, depending on the model that is estimated):

$$\mathbb{E}U_{zkj} = \mathbf{1}^{RP} * \mathbb{E}V_{zkj}^{RP} + (1 - \mathbf{1}^{RP}) * \lambda * \mathbb{E}V_{zkj}^{SP} + \epsilon_{zkj} \tag{10}$$

For the MNL models, the random component ϵ_{zkj} is assumed to follow a Gumbel distribution, with errors assumed distributed identically and independently (iid) across observations. While the parameter estimates obtained from the MNL models ignore the panel nature of the data, we do account for it in the computation of the standard errors by using the panel specification of the sandwich estimator (e.g Daly and Hess, 2011).

[19] Differences in scale are only relevant if common coefficients for the SP and RP observations are estimated.

The probability of driver z choosing alternative $\breve{j} \in J_{zk}$ in choice k, $P_{zk\breve{j}}$, is then defined as follows for the RP domain (an identical formulation holds for SP domain, with the superscripts differing accordingly):

$$P_{zk\breve{j}} = \frac{\exp(V_{zk\breve{j}}^{RP})}{\sum_{j=1}^{J(zk)} \exp V_{zkj}^{RP}} \tag{11}$$

The corresponding loglikelihood function is given below. It is conditional on the choices made by driver z in choice situation k. For simplicity, we use the notation of $\breve{P}_{zk}$ to indicate the probability associated with the *chosen* alternatives:

$$\ln L = \sum_{z=1}^{Z} \sum_{k=1}^{K_z} \breve{P}_{zk} \tag{12}$$

Moreover, we use a panel latent class model, allowing for heterogeneity among drivers and taking into account the panel setup of the underlying data. Latent class models assume that drivers can be sorted into a set of $q = 1, \ldots, Q$ classes. Preferences can then vary between the classes, while they are assumed homogenous within each class. This means that (some or all) coefficients estimates will be class-specific. The term 'latent' derives from the fact that heterogeneity is unobserved by the analyst. Class membership probabilities as well as the size of each class are unknown in advance; the analyst has only control over the number of classes Q. Hence, in addition to the (class-specific) coefficients, also the driver-specific probabilities of being member of a specific class need to be estimated. In line with the notation used by Greene and Hensher (2003), these are denoted by H_{zq}. While various formulations are possible for the estimation of class membership, we adopt a simple multinomial logit form here (without additional explanatory variables[20] other than a constant), which ensures that relative class sizes sum up to 1. The log-likelihood function is then given by the following equation:

$$\ln L = \sum_{z=1}^{Z} \ln \left[\sum_{q=1}^{Q} H_{zq} \left(\prod_{k=1}^{K_z} \breve{P}_{zk|q} \right) \right], \tag{13}$$

where $\breve{P}_{zk|q}$ is equal to the probability associated with the alternative chosen by person z in choice situation k conditional on z being member of class q. The multiplicative term $\prod_{k=1}^{K_z} \breve{P}_{zk|q}$ therefore represents the probability of the sequence of choices $k = 1, \ldots, K_z$ made by driver z, again conditional on class membership.

In contrast to mixed logit models, which assume a continuous distribution of (some) parameters, latent class models do not require any assumptions regarding the shape of the distribution of a given parameter. Moreover, for a fairly small number of classes, it is usually not necessary to assume that specific coefficients are constant across classes, whereas such an assumption is frequently made in mixed logit models in order to ensure empirical identification of the models.

[20]We tested a panel latent class model with class membership being conditional on various socio-economic characteristics, however, most of the corresponding coefficients turned out insignificant. This is an indication that the between-class-differences of the coefficient estimates are mainly a result of unobserved heterogeneity.

7.2. Estimation results: Multinomial logit models

Table 5 provides the results for the multinomial logit (MNL) models.[21] For this table, RP and SP data have not been pooled. The first two models presented in Table 5 correspond to estimations of the standard scheduling model, while the third and the fourth model take into account the travel time ratio, using the specifications introduced in Eqs. 6 and 7 for RP and SP departure time choices, respectively. All monetary valuations[22] are provided in Euro per hour.

Table 5: MNL models

	Without τ				With τ			
	RP only		SP only		RP only		SP only	
	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
β_R	0.19	5.52	0.05	4.20	0.22	5.58	0.05	4.18
β_T	-7.01	-6.45	-2.37	-13.27	-7.43	-6.25	-2.90	-8.27
β_E	-1.93	-15.13	-1.59	-16.52	-1.90	-12.65	-1.59	-16.53
β_L	-1.55	-17.30	-1.14	-13.25	-1.56	-16.61	-1.14	-13.27
θ	0.14	5.28	–	–	0.14	5.07	–	–
β_R^τ	–	–	–	–	0.05	1.21	–	–
β_T^τ	–	–	–	–	–	–	0.79	1.92
β_E^τ	–	–	–	–	-1.60	-1.41	–	–
β_L^τ	–	–	–	–	1.74	1.50	–	–
VOT (Euro/h)	36.32	4.08	44.38	3.94	34.24	4.08	54.61	3.68
VSDE (Euro/h)	10.00	5.17	29.78	4.40	8.76	5.10	29.92	4.38
VSDL (Euro/h)	8.03	5.21	21.35	4.33	7.19	5.24	21.47	4.31
Nr. Obs.	8477		4060		8477		4060	
Nr. Individuals	406		406		406		406	
LogLik.	-20386		-2223		-20370		-2220	
Pseudo R^2 adj.	0.151		0.208		0.152		0.209	

In contrast to most earlier research that compares RP- and SP-based valuations of travel time, we find that VOT estimates derived from these two data sources for the standard scheduling model are fairly close.[23] This result might be a consequence of designing the SP choice experiment such that the choice situations presented to the respondents are personalized and closely resemble those they face in reality. However, it might also be a coincidental finding that disappears once heterogeneity across drivers is considered. The estimation results of the panel latent class model (as presented in the next section) will shed more light on this issue.

The finding that the VOT is relatively high in all models can probably be explained by the fact that participation in the peak avoidance experiment and the related questionnaire was voluntary, and the sample is thus not randomly selected. The self-selection of participants led to a sample consisting of individuals with fairly high household incomes (see Table 1), who are likely to have also a high VOT. Another possible explanation for the high VOTs may be that the travel time coefficient incorporates also the disutility from travel time variability, which is strongly correlated to travel

[21] All models are estimated using the Python BIOGEME software (Bierlaire, 2003).

[22] The corresponding t-statistics are computed using the Delta method (e.g. Small, 2012).

[23] The relevant literature often finds them to differ by a factor 2 (with the RP-based VOT typically being higher than the SP-based one)

time and therefore not estimated separately. Moreover, the VOTs derived in this study have been determined in the context of a reward experiment instead of the more common experimental setup where the monetary incentive is represented as a cost. In line with the concept of loss aversion, individuals may be more sensitive towards costs than rewards, providing an explanation for low reward coefficients, which in turn leads to higher valuations. And finally, we consider here fairly long trips (with an average duration of ca. 37 minutes: see Table 2), which tend to be associated with a higher VOT (Daly and Carrasco, 2009). Nevertheless, given the similarity between SP and RP estimates of the VOT and various (non-contradictory) possibilities to explain the rather high valuations, the VOTs still seem trustworthy. And again, the panel latent class model presented below will provide further insights.

In contrast to the valuation of travel time, the values of schedule delay early and late differ considerably, depending on whether RP or SP data are used. The VSDE and the VSDL are significantly higher for the SP-based models than for the RP-based ones.[24] A possible cause may be that the hypothetical character of the SP choices makes the scheduling restrictions seem more binding (and thus, more costly) compared to their real-life counterpart. Similarly, in the SP environment participants may assume that they will be able to re-schedule their activities in order to adapt to the delays, whereas they will not always be able to do so in reality (consequently inducing relatively lower RP scheduling valuations). Furthermore, the higher SP scheduling valuations may also follow from the explicit representation of travel time variability in the SP experiment. The fact that the random nature of schedule delays is explicitly stated may induce respondents to attach more weight to schedule delays in their SP choices than they would if travel time variability was expressed in a more implicit way. Correspondingly, recent research by Peer et al. (forthcoming) has shown that unexpected delays are valued higher than expected ones, supposedly because re-scheduling becomes more difficult when delays are unexpected.

In both the SP- and the RP-based standard scheduling models (without τ) we obtain the result that earliness with respect to the preferred arrival time (PAT) is more costly than lateness. This finding may be driven by participants with an early PAT who have a large disutility from switching to an even earlier departure time. Furthermore, we find that in the RP-based models, the relative weight of the actual travel time on the day of travel to the (time-of-day-specific) average travel times, θ, is equal to 0.14, regardless of whether the travel time ratio is taken into account or not. Travel time expectations are thus mainly based on long-run averages but get updated by day-specific information.

Also the effect of including non-linear terms for the travel time and scheduling attributes was tested for the standard scheduling models. While some of these terms turned out significant, we do not proceed with the non-linear models, since they make it difficult to test whether the travel time ratio affects choice behavior, especially when the travel time ratio is specified as a person-specific constant. For consistency reasons, the travel time ratio would have to be formulated as a (possibly also non-linear) function of the observed travel time. However, this is difficult to achieve, since only a limited number of trips is observed for each participant.

The third and fourth model presented in Table 5 take into account the travel time ratio for RP and SP data, respectively. In both models, we use the 'standard' (person-specific) definition for the travel time ratio from Table 2. As a robustness check, the models were re-estimated using the alternative definitions of the travel time ratio, but the results did not change significantly.

[24]Existing studies usually do not find a clear pattern in the relation of RP- and SP-based scheduling values (e.g. Tseng et al., 2005; Li et al., 2010).

Due to strong correlations in the attributes of the RP model when including the travel time ratio, we were not able to estimate the full model of Eq. 6. Instead we leave out the 'travel time ratio term' for the travel time attribute. If these terms for the remaining three attributes are significant, we can still conclude that the reported travel times are a good representation of the drivers' actual perceptions. However, from Table 5 we can see that this is not the case as all of these three coefficients are insignificant.[25] We thus do not find convincing evidence that drivers misperceive travel times to a similar extent as they over- (and in few cases under-) report them when the make their scheduling decisions in reality.

In the SP-based model that accounts for the travel time ratio, β_T^τ is found to be positive, as suggested by Brownstone and Small (2005). However, it is just at the verge of being significant at the 5% level, with a t-statistic of 1.92. Due to this inconclusiveness, in the next section we will present the estimation results of a joint SP–RP panel latent class model that accounts for possible misperceptions in the SP domain while allowing for heterogeneity across drivers.

A robustness check is performed for both MNL models that consider the travel time ratio. The test involves the truncation of the distribution of the travel time ratio at the right (at 1.5 and 2). The underlying reasoning is that many respondents have travel time ratios exceeding 1.5, while none of them has a travel time ratio smaller than 0.5 (see Figure 1). It is possible that the results presented in Table 5 are affected (or even driven) by these high travel time ratios, which seem too high to reflect true misperceptions (and are more likely to reflect reporting errors). Table 6 presents the relevant parameter estimates.

Table 6: Estimation results with truncated travel time ratios τ

			β_R^τ		β_T^τ		β_E^τ		β_L^τ	
Trunc.	Data	Nr. Obs.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.	Coeff.	t-stat.
2	RP	7697	0.06	1.23	–	–	-0.71	-0.49	0.89	0.58
1.5	RP	3820	0.17	2.31	–	–	3.18	0.96	-3.61	-1.04
2	SP	3580	–	–	1.16	1.49	–	–	–	–
1.5	SP	1810	–	–	-0.45	-0.22	–	–	–	–

The RP model, with truncation of the travel time ratio distribution at 2, yields similar results as the original estimations presented in Table 5. If all persons with travel time ratios exceeding 1.5 are excluded from the analysis, the scheduling coefficients consistent with a misperception similar to the travel time ratio (indicated by superscript τ) remain insignificant, whereas the reward coefficient becomes significant at the 5% level (t-statistic: 2.31). However, the positive and significant reward coefficient is insufficient to conclude that drivers misperceive travel times in the RP setting: on the contrary, the scheduling coefficients not being significant, and β_E^τ being positive (it should be negative if travel times were misperceived in a similar way as they were misreported) do not support this conclusion.

Also for the SP model, Table 6 shows that the relevant coefficients remain similar to the ones derived in the the original estimation when the distribution of the travel time ratio is truncated at 2: So, β_T^τ is still positive, however, has a lower t-statistic than in the original estimation (the

[25]Consistent with this finding, the reverse model that takes into account only the travel time ratio term for travel times does not yields a significant coefficient estimate for this term either.

1.49 compared to 1.92, respectively). If only travel time ratios below 1.5 are considered in the estimation, the β_T^τ becomes negative (whereas a positive sign would be expected based on the proposition of Brownstone and Small (2005)) and the t-statistic of decreases even further (to -0.22). While the decrease in the t-statistic may be partially explainable by the lower number of observations, the size of this drop together with the coefficient becoming closer to 0 suggest that participants do not react to the travel times presented to them in the SP setting in a way that is consistent with their misreporting. Nevertheless, this result does not necessarily mean that drivers react to the travel times presented to them as if they would react to them in reality. Instead, it may imply that our specification of the travel time ratio is not an appropriate representation of the drivers' travel time perceptions.

7.3. Estimation results: Panel latent class model

The MNL model presented in Table 5 and the robustness checks presented in Table 6 seem to suggest that the theory brought forward by Brownstone and Small (2005) to explain structural differences between SP- and RP-based valuations of travel time does not hold for our data (one reason may be that the reported travel times do not represent travel time perceptions; in that case we are not able to test their theory in a valid way). In order to gain further insights, we estimate a panel latent class model that takes into account the travel time ratio in the SP domain, while considering both SP and RP scheduling choices. The intuition is that by pooling the data sources, unobserved preference heterogeneity across participants can be better represented. Therefore, this estimation procedure can identify if there are groups of drivers who react to travel times presented to them in an SP setting as if they were overestimated in the case of overreporting, or underestimated in the case of underreporting.

We estimate a panel latent class model with three classes ($Q = 3$). Various statistical criteria such as the BIC or the AIC have been suggested as adequate criteria for the selection of the number classes (e.g. Greene and Hensher, 2003). However, as Scarpa and Thiene (2005) note, this selection *must also account for significance of parameter estimates and be tempered by the analyst's own judgment on the meaning of the parameter signs.* Here, statistical criteria would suggest a higher number of classes. Likely due to correlations between the attribute values, which are especially evident in the RP setting, several coefficients then assume the reverse sign of what one would expect, leading to negative valuations of time and schedule delays, which are clearly not plausible from an economic point of view. Nevertheless, the higher number of classes suggested by the statistical criteria is an indication of a substantial amount of heterogeneity in the participants' scheduling preferences. This claim is supported by the fact that the fit of the panel latent class model (as measured by the adjusted pseudo R^2) is substantially better than for the MNL models.

Table 7: Panel latent class model

	Class 1			*Class 2*			*Class 3*	
	Coeff.	t-stat.		Coeff.	t-stat.		Coeff.	t-stat.
Class Prob.	0.36			0.19			0.45	
β_R^{RP}	0.21	3.14		0.16	1.93		0	–
β_T^{RP}	-5.06	-2.43		-19.4	-5.73		-8.99	-6.01
β_E^{RP}	-2.60	-9.04		-1.39	-4.76		-3.45	-13.43
β_L^{RP}	-0.70	-6.25		-2.01	-5.66		-3.05	-13.84
θ	0.12	5.34	$=^a$	0.12	5.34	$=^a$	0.12	5.34
β_R^{SP}	0.19	6.99		0.048	0.69		0	–
β_T^{SP}	-4.26	-3.57		-2.18	-1.39		-3.33	-3.25
β_E^{SP}	-1.10	-5.04		-4.31	-5.15		-3.09	-9.93
β_L^{SP}	-0.94	-5.41		-9.16	-5.42		-1.06	-5.88
$\beta_T^{SP,\tau}$	1.36	0.99		-1.91	-0.85		0.96	1.16
VOT^{RP}	23.87	1.92		124.4	1.83		–	
VSDE^{RP}	12.26	2.97		8.91	1.79		–	
VSDL^{RP}	3.30	2.81		12.88	1.82		–	
VOT^{SP}	22.19	3.19		–			–	
VSDE^{SP}	5.73	4.09		–			–	
VSDL^{SP}	4.89	4.28		–			–	
Nr. Obs.				12537				
LogLik.				-20088				
Pseudo R^2 adj.				0.25				

[a] Restricted to be equal

In the model presented in Table 7, all coefficients are specific to one of the datasources, meaning that no separate scale parameters are identified. Although some coefficients do not significantly differ across datasources, we prefer to keep them separate, in order to identify the effect of the travel time ratio on the SP choices unambiguously (hence, reducing the risk that $\beta_T^{SP,\tau}$ picks up some difference between SP and RP coefficients). The SP and the RP choices are therefore connected mainly via class membership, as for a given person the same class membership probability applies to all his/her SP and RP choices.

Based on extensive testing of alternative model formulations, some coefficients have been restricted in the model presented in Table 7. So, we do not specify θ as class-specific, as it does not differ significantly across classes. More importantly, for one class ('Class 3') the reward coefficient has been fixed to 0 for both the SP and the RP domain. This is because the some of the reward coefficients are otherwise close to 0 or even negative if unrestricted. As a consequence, the point estimates of the travel time and schedule delay valuations would approach infinity or turn negative, respectively. The size of the class probability of 'Class 3' (0.45) indicates that a considerable share of participants do not consider the reward when deciding on their actual and hypothetical departure times. In fact, we find that also the reward coefficient in the SP domain in 'Class 2' is close to 0. A possible cause for these findings is that a reward of 4 Euro is too low for some participants to be willing to shift their departure times (actually or hypothetically) to off-peak periods.

Due to the reward coefficient in the SP domain of 'Class 2' being close to 0 and insignificant and due to the reward coefficients being restricted to 0 in 'Class 3', we are only able to derive willingness-to-pay estimates for 'Class 1' and the RP domain of 'Class 2'. We find that the resulting willingness-to-pay estimates are all significant at the 10% level. For 'Class 1' we find consistent results for the SP and RP domain, with especially the point estimates of the VOT resembling each other closely. But also the schedule delay valuations are in the same order of magnitude (unlike in the MNL model where the SP-based scheduling valuations were about three times as high as the RP-based ones). One can also observe that the willingness-to-pay estimates for 'Class 1' are generally lower than the ones found in the MNL models. This is because – unlike 'Classes 2 & 3' – 'Class 1' captures the behavior of individuals with high reward preferences. This is reflected by the high reward coefficients in 'Class 1' (higher than in the MNL models), which in turn imply lower willingness-to-pay values.

Only for 'Class 1', we are able to compare SP- and RP-based valuations. For the remaining two classes we need to rely on comparisons between the SP and RP coefficient values. Generally, although the coefficients vary between the two data types, there are some reassuring similarities between them. For instance, in 'Classes 1 & 3', the disutility of being early is found to be larger than of being late both in the SP and in the RP domain, while for 'Class 2' the opposite holds in both domains. The exception is the travel time coefficient in 'Class 2': the disutility from travel time is found to be very large in the RP domain, but much less so in the SP domain. In the latter, the disutility associated with X minutes of travel time is even lower than the disutility associated with X minutes of schedule delay.

Finally, consistent with the MNL results, the coefficients related to the objective SP travel time multiplied by the travel time ratio, $\beta_T^{SP,\tau}$, are again insignificant for all three classes. This is yet another indication that the theory put forward by Brownstone and Small (2005) cannot be confirmed here, either because drivers do not misperceive travel times in the SP setting, or because the travel time ratio does not reflect the actual misperceptions. For the same reasons, we cannot explain any remaining differences between SP and RP willingness-to-pay values by misperceptions that the travel time ratio is able to capture. The finding that the SP and RP results are fairly similar may actually reflect that travel time misperception plays little role in the context of the (RP and SP) choice experiments discussed here. At least it suggests a much lower extent of misperception compared to the misreporting we found.

8. Conclusions

This study compares reported and actual travel times of car commuters who participate in a large-scale peak avoidance experiment. A strong divergence between reported and actual travel times is observed at the level of the individual driver and expressed as ratio, which we refer to as 'travel time ratio'. It turns out that on average participants overstate travel times by more than 50% (a travel time ratio higher than 1.5). We are able to identify some main determinants of the travel time ratio, such as link length and average speed, and test some of the hypotheses brought forward in the literature regarding link- and driver-specific characteristics that may explain over-reporting. Next, we investigate whether the overstated travel times are the result of reporting errors or indeed reflect perception errors. We do so by testing whether the the person-specific travel time ratio affects departure time choices in revealed preference (RP) and stated preference (SP) settings. In both settings we find that the travel time ratio has only very limited explanatory power. We can therefore conclude that the travel time ratio is mainly a reporting error, and that reported travel times seem to be a poor indicator of both actual and perceived travel times, at least in the context of our research.

Reporting errors can arise for numerous reasons, some of which are briefly discussed here. First, even though the questionnaire from which the reported travel times were collected explicitly focused on in-vehicle time, drivers may also have taken into account the time they require to get from their home to their car as well as the time they require to get from their car to their work place.[26] Second, drivers might unconsciously add a safety margin when thinking about travel times.[27] We expect this reasoning to be mainly relevant for those with inflexible work starting hours, as they are more likely to consider a safety margin in their scheduling choices compared to travelers with flexible work starting hours. Third, the finding that the average travel time ratio is substantially higher than 1 might be driven by a social desirability bias. For instance, drivers might assume that it is considered socially desirable to drive slowly (i.e. being more responsible towards other road users, as well as towards the environment), and therefore state longer travel times than they actually experience. Fourth, also strategic behavior might play a role in the answers provided in the questionnaire. By overstating travel times, the respondents might hope to push forward the implementation of policies targeted at decreasing travel times. Finally, reporting errors may be induced by rounding errors or by asking drivers to state average, time-of-day-independent travel times, which are not necessarily relevant for their actual scheduling behavior. However, for both of the latter cases there is no clear reasoning why overstating (rather than understating) of travel times occurs; we therefore do not expect them to be the main drivers behind the results obtained in this paper.

[26] Drivers may have taken into account the time spent on cruising for parking in their reported travel times; however, no over-reporting is expected to arise from this, since cruising has not been removed from the underlying GPS dataset that is used for the estimation of door-to-door travel times. Generally, we do not expect cruising to play a big role in our data because only morning commute trips are considered in the analysis and most Dutch employers provide parking spaces for their employees.

[27] We thank the referee who brought this to our attention.

It might be possible to avoid or at least reduce reporting errors in travel time data if the data collection methods were improved. For instance, instead of asking respondents to state travel times, it might be beneficial to let them fill in a schedule, indicating at which time they leave their home, start their car, park their car again, and finally, arrive at their work location. Such a setup is likely to limit biases that arise from respondents not only considering in-vehicle time in their answers, and it would make it more difficult for them to engage in strategic or socially desirable answering patterns, compared to the situation when they only have to insert absolute travel times. This approach may be extended by asking for a set of previously experienced travel times rather than averages, with the aim to avoid reporting errors that are due to aggregating a set of commuting experiences. Clearly, these alternatives involve a trade-off between information gains for the researcher and additional time required from the respondents for filling in the questionnaire. Judging by the results of this paper, however, the benefits seem to outweigh the costs.

The main focus of the SP- and RP-based choice models in this paper was to test whether drivers misperceive travel times to a similar extent as they misreport them. Although this could not be confirmed for either of the two data sources, it is still possible that drivers misperceive travel times in a way that is not reproduced by the travel time ratio. However, such an alternative pattern of misperception is difficult to identify using observed (SP or RP) choices only, as it is usually not clear whether the estimated coefficient values are driven by misperception rather than reflecting actual preference structures. Further research should thus focus on the challenging task of identifying (instrumental) variables that affect the travel time misperception, but not the preferences that drive the scheduling choices.

The results of this paper were derived in a fairly specific context, namely a peak avoidance experiment along a highway link in the Netherlands. Participation in the experiment was voluntary. As a result, the participants are neither representative of the drivers using this highway link, nor of the Dutch population as a whole. Most prominently, participants tend to have a higher salary and more flexible working hours than non-participants. While we do not believe that this selection bias drives our results on over-reporting vs. overreacting, we are aware that the valuations derived in this paper are not representative for a larger population, and hence encourage follow-up research using alternative data sources.

To conclude, we found strong evidence of ‘overstating’, however, much less evidence of ‘overreacting’, neither in SP- nor in RP-based scheduling choices. The results should thus be taken as a strong note of caution regarding the use of reported travel time data as an indicator of actual as well as perceived travel times. While we show this for the case of travel times, this conclusion can probably be generalized for various contexts.

Acknowledgements

This study is financially supported by the Dutch Ministry of Infrastructure and the Environment, as well as by the TRANSUMO foundation. Also the financial support from the ERC (Advanced Grant OPTION #246969) for the research of Paul Koster and Erik Verhoef, and from IPRISM for the research of Jasper Knockaert is gratefully acknowledged. We wish to thank two anonymous reviewers for their constructive suggestions.

References

Abrantes, P. A., Wardman, M. R., 2011. Meta-analysis of UK values of travel time: An update. Transportation Research Part A: Policy and Practice 45 (1), 1–17.

Allan, L. G., 1979. The perception of time. Attention, Perception, & Psychophysics 26 (5), 340–354.

Arnott, R., De Palma, A., Lindsey, R., 1993. A structural model of peak-period congestion: A traffic bottleneck with elastic demand. American Economic Review 83 (1), 161–179.

Arnott, R., De Palma, A., Lindsey, R., 1994. The welfare effects of congestion tolls with heterogeneous commuters. Journal of Transport Economics and Policy 28 (2), 139–161.

Baldauf, D., Burgard, E., Wittmann, M., 2009. Time perception as a workload measure in simulated car driving. Applied ergonomics 40 (5), 929–935.

Bierlaire, M., 2003. Biogeme: A free package for the estimation of discrete choice models. Proceedings of the 3rd Swiss Transportation Research Conference, Ascona, Switzerland.

Block, R. A., Hancock, P. A., Zakay, D., 2010. How cognitive load affects duration judgments: A meta-analytic review. Acta psychologica 134 (3), 330–343.

Block, R. A., Zakay, D., 1996. Models of psychological time revisited. In: Helfrich, H. (Ed.), Time and mind. Hogrefe & Huber, Kirkland, WA, pp. 171–195.

Boltz, M. G., 1998. Task predictability and remembered duration. Perception & Psychophysics 60 (5), 768–784.

Boltz, M. G., Kupperman, C., Dunne, J., 1998. The role of learning in remembered duration. Memory & Cognition 26 (5), 903–921.

Börjesson, M., 2008. Joint RP-SP data in a mixed logit analysis of trip timing decisions. Transportation Research Part E: Logistics and Transportation Review 44 (6), 1025–1038.

Brownstone, D., Small, K., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transportation Research Part A: Policy and Practice 39 (4), 279–293.

Burnett, P., 1978. Time cognition and urban travel behavior. Geografiska Annaler. Series B. Human Geography 60 (2), 107–115.

Carlsson, F., 2010. Design of stated preference surveys: Is there more to learn from behavioral economics? Environmental and Resource Economics 46 (2), 167–177.

Daly, A., Carrasco, J., 2009. The influence of trip length on marginal time and money values. In: Kitamura, R., Yoshii, T., Yamamoto, T. (Eds.), The Expanding Sphere of Travel Behavior Research: Selected Papers from the 11th International Conference on Travel. Emerald, Bingley, UK.

Daly, A., Hess, S., 2011. Simple approaches for random utility modeling with panel data. Paper presented at the 90th Annual Meeting of the Transportation Research Board, Washington, DC.

Droit-Volet, S., Meck, W. H., 2007. How emotions colour our perception of time. Trends in Cognitive Sciences 11 (12), 504–513.

Eisler, A., Eisler, H., Montgomery, H., 2007. A quantitative model for retrospective subjective duration. NeuroQuantology 2 (4), 263–291.

Eisler, H., 1976. Experiments on subjective duration 1868-1975: A collection of power function exponents. Psychological Bulletin 83 (6), 1154–1171.

Flaherty, M., 1999. A Watched Pot: How we Experience Time. New York University Press, New York, NY.

Ghosh, A., 2001. Valuing time and reliability. Commuters' mode choice from a real time congestion pricing experiment. Ph.D. thesis, University of California at Irvine.

Glicksohn, J., 2001. Temporal cognition and the phenomenology of time: A multiplicative function for apparent duration. Consciousness and Cognition 10 (1), 1–25.

Greene, W., Hensher, D., 2003. A latent class model for discrete choice analysis: contrasts with mixed logit. Transportation Research Part B: Methodological 37 (8), 681–698.

Grondin, S., 2010. Timing and time perception: a review of recent behavioral and neuroscience findings and theoretical directions. Attention, Perception, & Psychophysics 72 (3), 561–582.

Henley, D., Levin, I., Louviere, J., Meyer, R., 1981. Changes in perceived travel cost and time for the work trip during a period of increasing gasoline costs. Transportation 10 (1), 23–34.

Hensher, D., 2010. Hypothetical bias, choice experiments and willingness to pay. Transportation Research Part B: Methodological 44 (6), 735–752.

Hensher, D. A., 2001. Measurement of the valuation of travel time savings. Journal of Transport Economics and Policy 35 (1), 71–98.

Isacsson, G., 2007. The trade off between time and money: Is there a difference between real and hypothetical choices? Working Paper, Swedish National Road & Transport Research Institute (VTI), Borlänge, Sweden.

Knockaert, J., Ettema, D., Verhoef, E., Koster, P., Peer, S., Tseng, Y.-Y., 2012. Spitsmijden Gouda-Zoetermeer. Scientific project report.
URL `http://www.spitsmijden.nl/downloads/SM2D_scientific_report_final.pdf`

Koster, P., Tseng, Y., 2010. Stated Choice Experimental Designs for Scheduling Models. In: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and the State-of-Practice: Proceedings from the Inaugural International

Choice Modelling Conference. Emerald, Bingley, UK.
Leiser, D., Stern, E., 1988. Determinants of subjective time estimates in simulated urban driving. Transportation Research Part A: General 22 (3), 175–182.
Li, Z., Hensher, D., Rose, J., 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. Transportation Research Part E: Logistics and Transportation Review 46 (3), 384–403.
Louviere, J., Hensher, D., Swait, J., 2000. Stated choice methods: analysis and applications. Cambridge University Press, Cambridge, UK.
McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. In: Zarembka, P. (Ed.), Frontiers in econometrics. Academic Press, New York, NY.
MVA Consultancy, 1987. The Value of Travel Time Savings. Policy Journals, University of Oxford, Newbury, Berkshire, UK.
Noland, R., Small, K., 1995. Travel-time uncertainty, departure time choice, and the cost of the morning commute. Paper presented at the 74th Annual Meeting of the Transportation Research Board, Washington, DC.
O'Farrell, P., Markham, J., 1974. Commuter perceptions of public transport work journeys. Environment and Planning A 6 (1), 79–100.
Parthasarathi, P., Hochmair, H., Levinson, D., 2013. Network structure and travel time perception. Paper presented at the 92th Annual Meeting of the Transportation Research Board, Washington, DC.
Peer, S., Knockaert, J., Koster, P., Tseng, Y., Verhoef, E., 2013. Door-to-door travel times in RP departure time choice models. An approximation method based on GPS data. Transportation Research Part B: Methodological 58, 134–150.
Peer, S., Verhoef, E., Knockaert, J., Koster, P., Tseng, Y.-Y., forthcoming. Long-run vs. short-run perspectives on consumer scheduling: Evidence from a revealed-preference experiment among peak-hour road commuters. International Economic Review.
Recarte, M., Nunes, L., 1996. Perception of speed in an automobile: Estimation and production. Journal of Experimental Psychology: Applied 2 (4), 291–304.
Rietveld, P., Zwart, B., Van Wee, B., Van den Hoorn, T., 1999. On the relationship between travel time and travel distance of commuters. The Annals of Regional Science 33 (3), 269–287.
Scarpa, R., Thiene, M., 2005. Destination choice models for rock climbing in the Northeastern Alps: a latent-class approach based on intensity of preferences. Land Economics 81 (3), 426–444.
Shires, J., De Jong, G., 2009. An international meta-analysis of values of travel time savings. Evaluation and program planning 32 (4), 315–325.
Small, K., 1982. The scheduling of consumer activities: work trips. American Economic Review 72 (3), 467–479.
Small, K., 2012. Valuation of travel time. Economics of Transportation 1 (1-2), 2–14.
Small, K., Verhoef, E., 2007. The economics of urban transportation. Routledge, London, UK.
Small, K., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. Econometrica 73 (4), 1367–1382.
Swait, J., Louviere, J. J., Williams, M., 1994. A sequential approach to exploiting the combined strengths of SP and RP data: application to freight shipper choice. Transportation 21 (2), 135–152.
Train, K., 2003. Discrete choice methods with simulation. Cambridge University Press, Cambridge, UK.
Tseng, Y., Knockaert, J., Verhoef, E., 2013. A revealed-preference study of behavioural impacts of real-time traffic information. Transportation Research Part C: Emerging Technologies 30, 196–209.
Tseng, Y.-Y., Rietveld, P., Verhoef, E., 2005. A meta-analysis of valuation of travel time reliability. In: Colloquium Vervoersplanologisch Speurwerk. Rotterdam, Netherlands.
Van Exel, N., Rietveld, P., 2009. Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travellers on the main travel corridors to the city of Amsterdam, The Netherlands. Transportation Research Part A: Policy and Practice 43 (4), 374–385.
Vickrey, W. S., 1969. Congestion theory and transport investment. American Economic Review 59 (2), 251–260.
Wardman, M., 2001. A review of British evidence on time and service quality valuations. Transportation Research Part E: Logistics and Transportation Review 37 (2-3), 107–128.
Wardman, M., Nicolás Ibáñez, J., 2012. The congestion multiplier: Variations in motorists valuations of travel time with traffic conditions. Transportation Research Part A: Policy and Practice 46 (1), 213–225.
Zamparini, L., Reggiani, A., 2007. Meta-analysis and the value of travel time savings: A transatlantic perspective in passenger transport. Networks and Spatial Economics 7 (4), 377–396.
Zhang, L., Xie, F., Levinson, D., 2005. Variation of the subjective value of travel time on freeways and ramp meters. Paper presented at the 84th Annual Meeting of the Transportation Research Board, Washington, DC.

Appendix A. List of Variables

Variable	Description
Data sources	
RP	Revealed preference (also used as superscript)
SP	Stated preference (also used as superscript)
Indices	
$d = 1, \ldots, D_z$	(Person-specific) trip index
$j = 1, \ldots, J_{zk}$	(Person- and choice-situation-specific) index of choice alternatives
$k = 1, \ldots, K_z$	(Person-specific) index of choice situations
l	Link index ∈{home–C1, C1–C2, C2–work, home–work}
$q = 1, \ldots, Q$	Class index for latent class models
$z = 1, \ldots, Z$	Person index
Travel time ratio	
T^R_{zl}	(Person- and link-specific) reported average travel time
T^A_{zld}	(Person-, link- and trip-specific) actual travel time
$\bar{T}^A_{zl}$	Average (person- and link-specific) actual travel time
τ_{zl}	(Person- and link-specific) travel time ratio
τ (superscript)	Reference to travel time ratio (e.g. $\beta^\tau_R, \beta^\tau_T, \beta^\tau_E, \beta^\tau_L, U^\tau_{zkj}, V^\tau_{zkj}$)
Utility function: Attributes	
$\mathbb{E}$	Expectation operator (see Eqs. 8 and 9 for the definition)
t	Departure time from home
PAT	(Driver-specific) preferred arrival time at work
$R(t), R_{zkj}, \bar{R}_{zj}$	Reward attribute (general, choice-specific, average)
$T(t), T_{zkj}, \bar{T}_{zj}$	Travel time attribute (general, choice-specific, average)
$SDE(t), SDE_{zkj}, S\bar{D}E_{zj}$	Schedule delay early attribute (general, choice-specific. average)
$SDL(t), SDL_{zkj}, S\bar{D}L_{zj}$	Schedule delay late attribute (general, choice-specific, average)
$A = \{R, T, SDE, SDL\}$	Attributes of the utility function (general)
$A_{zkj}, \bar{A}_{zj}$	Attributes of the utility function (choice-specific, average)
Utility function: Coefficients	
$\beta_R, \beta_T, \beta_E, \beta_L$	Reward, travel time, earliness and lateness coefficients
θ	Relative weight drivers attach to actual travel time realizations vs. averages
λ	Multiplicative scale factor for SP choices (scale for RP choices is fixed to 1)
Utility function: General	
U_{zkj}	(Person-, choice situation- and alternative-specific) utility
V_{zkj}	Deterministic component of the utility function
ϵ_{zkj}	Stochastic component of the utility function
$\ln L$	Loglikelihood
P_{zkj}	Probability of person z choosing alternative j in choice situation k
$\breve{P}_{zk}$	Probability associated with the alternative *chosen* by person z in choice situation k
$\breve{P}_{zk\|q}$	Probability associated with the alternative *chosen* by person z in choice situation k conditional on being member of class q
H_{zq}	Probability of person z being member of class q
Valuations	
VOT	Value of (travel) time
VSDE	Value of schedule delay early

VSDL	Value of schedule delay late