

Peer, Stefanie; Knockaert, Jasper; Koster, Paul; Verhoef, Erik

**Working Paper**

## Overreporting vs. Overreacting: Commuters' Perceptions of Travel Times

Tinbergen Institute Discussion Paper, No. 13-123/VIII

**Provided in Cooperation with:**

Tinbergen Institute, Amsterdam and Rotterdam

*Suggested Citation:* Peer, Stefanie; Knockaert, Jasper; Koster, Paul; Verhoef, Erik (2013) : Overreporting vs. Overreacting: Commuters' Perceptions of Travel Times, Tinbergen Institute Discussion Paper, No. 13-123/VIII, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:

<https://hdl.handle.net/10419/87390>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*

TI 2013-123/VIII  
Tinbergen Institute Discussion Paper



# Overreporting vs. Overreacting: Commuters' Perceptions of Travel Times

*Stefanie Peer*<sup>1,2,3</sup>

*Jasper Knockaert*<sup>1</sup>

*Paul Koster*<sup>1,2</sup>

*Erik Verhoef*<sup>1,2</sup>

<sup>1</sup> Faculty of Economics and Business Administration, VU University Amsterdam;

<sup>2</sup> Tinbergen Institute;

<sup>3</sup> Institute for the Environment and Regional Development, Vienna University of Economics and Business, Austria.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam  
Burg. Oudlaan 50  
3062 PA Rotterdam  
The Netherlands  
Tel.: +31(0)10 408 8900  
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance  
Gustav Mahlerplein 117  
1082 MS Amsterdam  
The Netherlands  
Tel.: +31(0)20 525 8579

# OVERREPORTING VS. OVERREACTING: COMMUTER'S PERCEPTIONS OF TRAVEL TIMES\*

Stefanie Peer<sup>†1,3</sup>, Jasper Knockaert<sup>1</sup>, Paul Koster<sup>1,2</sup>, and Erik T. Verhoef<sup>1,2</sup>

<sup>1</sup>*Department of Spatial Economics, VU University Amsterdam, De Boelelaan 1105, 1081 HV Amsterdam*

<sup>2</sup>*Tinbergen Institute, Gustav Mahlerplein 117, 1082 MS Amsterdam*

<sup>3</sup>*Institute for the Environment and Regional Development, Vienna University of Economics and Business, Nordbergstrasse 15, 1090 Vienna*

August 15, 2013

## Abstract

Participants of a large-scale, real-life peak avoidance experiment have been asked to provide estimates of their average in-vehicle travel times 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 test to which extent driver- and link-specific characteristics explain the overstating. Subsequently, 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. These results imply that reported travel times are neither an appropriate measure for representing actual nor perceived travel times, and are thus a strong caveat against the uncritical use of reported travel time data in research and policy.

*JEL codes:* C25, D83, D84, R41

---

\*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.

<sup>†</sup>Corresponding author (E-mail: s.peer@vu.nl)

## 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, they usually act as a substitute for actual travel times. Moreover, in the spirit of the literature on time perception (see for instance Grondin (2010) for a recent overview), 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 equal, whereas in the second case, reported and perceived travel times are assumed 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. Burnett, 1978; Henley et al., 1981; MVA Consultancy, 1987; O'Farrell and Markham, 1974; Rietveld et al., 1999; Van Exel and Rietveld, 2009), and, based on this evidence, 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 possible explanations for the overstated travel times. Consistent with the results obtained in these earlier studies, also the research presented in this paper found that travel times are strongly overreported, by a factor of 1.5 on average.

The implications of overstating for the use of reported travel time data depend strongly on whether the presence of overstating is a consequence of wrong reporting, 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 actual 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 in this case travel time misperceptions may lead to biased estimates of the value of travel time (VOT)<sup>1</sup> if the VOT is derived from stated preference (SP) data. SP data are collected from choice experiments where respondents are asked to decide between hypothetical travel alternatives, and are therefore distinct from revealed preference (RP) data that are collected under real-life circumstances.

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 overestimated as well, so that the response to a time difference stated in the experiment corresponds to the response to a smaller time difference in reality, resulting in a relatively low VOT. For instance, an individual who perceives a real travel time of 10 minutes as one

---

<sup>1</sup>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).

of 20 minutes probably reacts to a travel time of 20 minutes in an SP setting in the same way as he would react to a travel time of 10 minutes in reality. The VOT derived in an SP setting would therefore be half of the VOT that would be derived if RP data were analyzed. As a consequence, if the resulting SP-based VOT is 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 (or costs) due to changes in travel times constitute a major category in the appraisal of most transport policies. Hensher (2001), for instance, estimates them at 60% of total user benefits.

In line with the hypothesis of Brownstone and Small (2005), it is a frequent outcome that RP-based estimates of the VOT tend to be higher than SP estimates (e.g. Brownstone and Small, 2005; Ghosh, 2001; Hensher, 2001; Isacsson, 2007; Small et al., 2005). 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). So, it has been speculated that people are more sensitive towards monetary attributes in hypothetical choice situations than in real life. Recent research shows that the occurrence of hypothetical biases can be reduced by providing choice options to the respondents that are as close as possible to the choices they face in real life (e.g. Hensher, 2010).

In this study, we first analyze reported versus actual travel times. As mentioned, 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 do not find this to be the case, we can 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 travel behavior along this 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 this specific highway link, day- and time-of-day-specific door-to-door travel times can be approximated applying the method described in Peer et al. (2011), which takes GPS data as an input. Taken together, this combination of data sources 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 actual and hypothetical travel choice behavior.

To our best knowledge, this paper is the first study that investigates the difference between reported and actual travel times (reporting errors) and between reported and perceived travel times (misperceptions) using actual and reported travel times as well as SP and RP travel choices from the same set of commuters. Research that is probably closer to ours has been conducted by Ghosh (2001). Similar to this study, he tested whether the difference between reported and actual travel times can explain travel behavior. He finds some evidence that this is the case, however, he also shows that the VOT is barely affected as a consequence of including this difference variable in the choice model. While Ghosh (2001) uses only RP data in his analysis on travel time misperception, we are specifically interested in the explanatory power of the travel time ratio with respect to travel choices collected through SP experiments. If we can confirm the hypothesis brought forward by Brownstone and Small (2005), we can not only conclude that the reported travel times are a good indicator of the (time) misperceptions persons exhibit when filling in an SP choice experiment, but also why SP estimates of the VOT often diverge (strongly) from their RP counterparts, possibly leading to biased policy assessments.

Various other studies in the area of transport economics have attempted to draw conclusions on travel time misperceptions based on the behavior of travelers in hypothetical and real-life choice situations. For instance, Brownstone and Small (2005) speculate that the common result of the VOT being higher under congested and slow traffic conditions (e.g. Hensher, 2001; Recarte and Nunes, 1996; Wardman, 2001; Wardman and Nicolás Ibáñez, 2012; Zhang et al., 2005) 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 higher VOTs can usually not be distinguished from an overestimation of the corresponding travel times. In this study, we circumvent this problem by explicitly testing whether the reported (on average overstated) travel times 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 (e.g. Noland and Small, 1995; Small, 1982; Vickrey, 1969) have gained an increasing amount of interest over the last few decades, as they are able to shed light on the effects of transport policies on departure time choices, and therefore, also on peak congestion as a whole. Empirically, departure time decisions are usually represented as choices between discrete departure time alternatives. The respective models can then be estimated using discrete choice analysis (McFadden, 1974; Train, 2003). Besides using standard multinomial logit models, in this paper 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. Börjesson, 2008; Brownstone and Small, 2005; Ghosh, 2001; Small et al., 2005). Moreover, most of them use fairly simplistic travel time measurements in their RP models (see Peer et al. (2011) for a discussion on possible biases), while the approximated door-to-door travel used in this study

allow for including 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 (hence, not only travel-time-related perceptions) (e.g. Grondin, 2010). Most studies in that field of research base their results on setting equal reported and perceived durations. This may be a reasonable assumption under controlled experimental conditions and for short time spans<sup>2</sup>, while it is much 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 differs substantially. 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 (e.g. 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. This effect might, however, be offset by another one: It has also been found that activities that are not well predictable, which also holds true for most car travel, are perceived as longer than they actually are (Y. Li, 2003; Van de Ven et al., 2011). Also the role of emotions is frequently pointed out. Most studies suggest 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 will seemingly pass faster (e.g. Block and Zakay, 1996), others report the opposite finding (e.g. Flaherty, 1999; Glicksohn, 2001).

This paper unfolds as follows. Section 2 discusses the specification of the travel time ratio and analyzes how this ratio can be explained by driver- and link-specific characteristics. Section 3 introduces the theoretical framework that allows for testing whether scheduling choice behavior both in reality and in hypothetical choice experiments is consistent with the reported travel times. Section 4 gives an overview of the SP and RP data used in the analysis, while Section 5 contains the econometric framework and the estimations of the choice models. Finally, Section 6 concludes on whether reported travel time data may be used as representations of actual and/or perceived travel times.

## 2 Specification of the travel time ratio

### 2.1 Data and definitions

#### Peak avoidance experiment

In this section of the paper individual- and link-specific travel time ratios, defined as the ratio of reported and actual travel times<sup>3</sup>, are computed. All data required to accomplish this comparison have been collected in the context of a peak avoidance experiment (*Spitsmijden*

---

<sup>2</sup>Some studies on time perception cover time intervals of only few seconds.

<sup>3</sup>Note that in the literature on time perception this ratio is also referred to as duration judgement ratio (e.g. Block et al., 2010).



in Dutch). The experiment took place in the Netherlands along a 9.21 km long stretch of the A12 highway leading to The Hague. As this link was confined by two cameras that were used for number plate detection, we 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, they were able to obtain a reward of 4 Euro for each avoided trip along the C1–C2 link during morning peak hours (6:30–9:30 a.m.), relative to their reference behavior. A more detailed description of the experiment can be found in Knockaert et al. (2012).

### 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 making up their door-to-door trip: the home–C1, C1–C2 and C2–work links. Respondents were asked to take into account their most recent 20 morning commute trips in their answers. Average home–work travel times are then computed as the sum of the travel times indicated for each of the three sub-links. To prompt respondents to give realistic answers, maps of the location of the C1–C2 link were shown next to the questions. Moreover, it was emphasized that only in-vehicle time should be reported. We label this definition ‘standard’, in order to distinguish it from an alternative measure of reported (home–work) travel times.

The alternative measure of reported (home–work) travel times is derived from yet another survey, which was conducted at the time when the participants entered the experiment. In this survey, drivers were asked to state their average home–work travel time. In contrast to the survey discussed in the preceding paragraph, respondents did not have to distinguish between any sub-links when filling in their answer. To emphasize this property, this definition is labelled ‘total’. We use this alternative measure to investigate whether the results of the ‘standard’ measure might be biased because they are based on the sum of travel times along the sub-links rather than total home–work travel times.

Reported travel times are denoted by  $T_{lz}^R$ , where  $l$  indicates the link  $l = \{\text{Home–C1, C1–C2, C2–Work, Home–Work}\}$  and  $z$  indicates the participant.

### Actual travel times

Multiple data sources are used in defining the actual travel times. Only the travel times through the C1–C2 link are observed individually for each driver, using number plate detection at the begin and at the end of the link. Home–C1 and C2–work travel times, however, need to be approximated. The method used is described in Peer et al. (2011). It

involves geographically weighted regression (GWR), which is used to measure the extent of correlation between speeds on the C1–C2 and home–C1 as well as C2–work links, respectively. 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. (2011) demonstrate that this method yields reliable predictions of travel times.

Actual travel times are denoted by  $T_{lzd}^A$ , where  $l$  again denotes the link and  $z$  the driver. Moreover,  $d = 1, \dots, D$  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  $D$ , which thus denotes the overall number of observed passages between September 1, 2009 and the date a specific driver has completed the questionnaire. For an observed passage to be considered in this list of past trips, certain conditions must be fulfilled. For the ‘standard’ definition, we assume that passages of the C1–C2 link between 6:00 and 11:00 are considered commuting trips. In an alternative definition, labelled ‘peak’, only those trips that result in a passage of the C1–C2 link between 6:30 and 9:30 (i.e. the period defined as peak in the experiment) are considered commuting trips. Although the latter definition reduces the number of observed trips that are available for the analysis, we use it in order to test whether the standard definition takes into account an excessively broad peak, which would in turn lead to a downward bias of actual travel times and thus to an upward bias of the ‘standard’ travel time ratio.

Not for all drivers 20 commuting trips (which were suggested to be considered when stating the average travel time in the questionnaire) have been recorded before the date the questionnaire was filled in. We then take into account as many observations as are available. In the reverse case, when more than 20 commuting trips with passages of the C1–C2 link have been observed, the 20 most recent ones enter the analysis. The average actual link- and driver-specific travel time is therefore equal to the following expression:

$$\bar{T}_{lz}^A = \frac{1}{\min[20, D]} \sum_{d=1}^{\min[20, D]} T_{lzd}^A \quad (1)$$

### Travel time ratio

The link- and driver-specific travel time ratio  $\tau_{lz}$  is defined 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. (2011) for comparing reported and observed travel times.

$$\tau_{lz} = T_{lz}^R / \bar{T}_{lz}^A \quad (2)$$

Some studies related to time perceptions in general have adopted a power law function in order to control for nonlinearities in the relation between reported and actual durations or even for subjectivity in the point of time when an activity is experienced to start (A. Eisler et al., 2007), however, mostly with the conclusion that the exponential term is close to 1, and

non-linearities are therefore small (e.g. Allan, 1979; H. Eisler, 1976). Leiser and Stern (1988) obtained a similar result for the travel time context, suggesting that the linear formulation performs almost as well the power law function. While we find some non-linearities to be present in our data (see the regression results in Table 3, Section 2.3), we do not adopt a more complex specification, mainly because our main interests in this research is to use the travel time ratio as an input to departure time choice models.

## 2.2 Descriptives

Table 1 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 give the results for 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 can be derived for 540 respondents.

Table 1: Reported vs. observed travel times (in minutes)

Link	Label	Nr. Obs.	Reported		Observed		$\tau$		
			Mean	St. D.	Mean	St. D.	Mean	Median	St. D.
H-C1	–	540	31:07	19:28	22:11	16:42	1.67	1.48	0.68
C1-C2	–	540	15:40	5:55	9:50	2:43	1.66	1.56	0.67
C2-W	–	540	15:57	7:41	11:03	6:55	1.74	1.51	1.00
H-W	stand.	540	62:44	22:44	43:04	18:01	1.53	1.49	0.38
H-W	total	527	64:35	22:27	43:04	18.01	1.62	1.56	0.44
H-W	peak	486	62:44	22:44	43:55	17:36	1.50	1.45	0.39

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 1.50 to 1.74 (the median is slightly lower, ranging from 1.45 to 1.56).<sup>4</sup> The first three rows of Table 1 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 overreporting by understating travel times on the remaining links of the commute.<sup>5</sup> A comparison of the 'total' definition of the travel time

<sup>4</sup>It can be shown that the travel time ratio of reported and actual travel times is slightly higher for the group of travelers with less than 20 observed trips, possibly because they took into account also shorter pre-experimental trips (for which no data are available) when filling in the questionnaire. While one would expect the pre-experimental trips to be longer due to a larger share of trips taking place during the peak, another mechanism seems to be at work here: As travel time measurements from the experiment are available from September 2009 onwards, the pre-experimental trips are likely to have taken place during summer months, which in the Netherlands are traditionally characterized by relatively small extents of recurrent congestion.

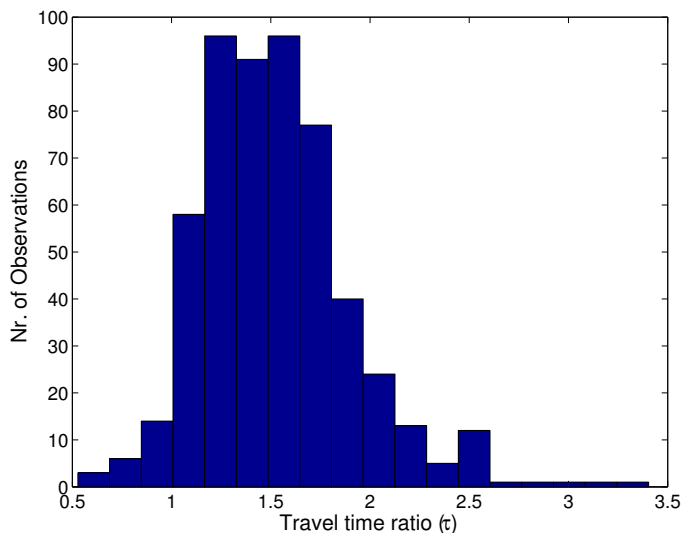
<sup>5</sup>The finding that the average home-work travel time ratio is smaller than the corresponding travel time ratios on the sub-links is driven by some outliers on the sub-links.

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 rather than the average travel times on the sub-links (1.62 vs. 1.53). The observed extent of overstating for the ‘standard’ definition is therefore not biased upwards as a result of using survey questions that refer to the travel times along the sub-links rather than the door-to-door travel time. Finally, we find that the travel time ratio for the ‘peak’ definition is almost as high as for the ‘standard’ definition (1.50 vs. 1.53). The overstating found for the ‘standard’ definition is thus not a consequence of an excessively broad definition of the morning peak. All in all, we can 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 we have tested.

## 2.3 Heterogeneity

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 a very minor share of respondents (4.44%) 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, although slightly skewed to the right, fairly symmetric, with a mean of ca. 1.5 and few outliers towards the right end of the distribution.

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



This section intends to explain the variation in the travel time ratio across drivers and across 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 2. 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 the frequency of travel (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 of these were found to be insignificant. 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. For various other variables that might affect the travel time ratio, such as road type or free-flow speed, no data were readily available.

Table 2: Descriptive statistics of driver- and link-specific variables

Variable	Specification	Mean	St. Dev
Nr. of trips	within the last 20 working days	13.01	6.76
Share of trips outside peak	since September 1, 2009	0.20	0.29
Mean speed home-work	in km/h	70.50	12.05
Mean speed home-C1	in km/h	74.89	18.65
Mean speed C1-C2	in km/h	66.91	13.82
Mean speed C2-work	in km/h	61.94	12.05
Distance home-work	in km	50.57	25.38
Distance home-C1	in km	29.82	24.60
Distance C1-C2	in km	9.21	0.00
Distance C2-work	in km	11.53	6.09

In the first model, the travel time ratios for the home-work link are used as dependent variable, while in the second model the travel time ratios of all three sub-links are considered. In the first model, the set of explanatory variables consists of person-specific variables as well as variables that characterize home-work trips. In addition to those variables, 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 having defined three sub-links  $t$  for each driver. Table 3 shows the according results.

Both model estimations result in an R-square of 0.22 to 0.23. This is an indication 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 drivers with short commuting distances and high average speeds on their commutes. Also, we find some evidence that those drivers who have little experience in commuting<sup>6</sup> tend to overstate travel times more. The same is true for individuals who travel relatively often during the peak<sup>7</sup>. These results are confirmed at the link level, too.

<sup>6</sup>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.

<sup>7</sup>Peak trips are defined as trips that lead to passages of the C1-C2 link between 6:30 and 9:30.

Table 3: Regression results: Explanation of the travel time ratio

Variable	Dependent Variable			
	$\tau_{\text{Home-Work}}$		$\tau_{\text{Home-C1,C1-C2,C2-Work}}$	
	Coefficient	t-stats.	Coefficient	t-tstats.
Constant	1.25	13.55	1.52	8.23
Little experience	$5.59 \cdot 10^{-2}$	1.78	$3.54 \cdot 10^{-2}$	0.56
Share of trips outside peak	-0.27	-4.71	-0.32	-4.18
Mean speed home-work (MSHW)	$1.07 \cdot 10^{-2}$	6.71	$5.82 \cdot 10^{-3}$	2.64
Distance home-work (DHW)	$-8.87 \cdot 10^{-3}$	-12.26	$-6.83 \cdot 10^{-3}$	-6.93
Mean speed link/MSHW	-	-	0.69	5.29
Distance link/DHW	-	-	-2.25	-18.83
dummy home-C1	-	-	0.55	11.48
dummy C2-work	-	-	0.18	4.40
Estimation Method	OLS		RE	
Nr. of Observations	540		1620	
R-square	0.23		0.22	

The relatively high extent of overstating of travel times along the home-C1 and C2-work links might be explainable by drivers mistakenly not only considering in-vehicle commuting time, but also the time they need to walk to the car and from the car to the office, although they were explicitly asked to only consider in-vehicle commuting time when answering the questions. This effect is probably more predominant on the home-C1 and the C2-work links than on the intermediate C1-C2 link. 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 relatively high travel time ratio for shorter commutes. The finding that travel times on sub-links with high speeds 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 proportionally to their lengths, not taking into account the differences in speeds across links.

Only in the next sections of the paper, when we test whether the travel time ratios are consistent with the traveler’s behavior in hypothetical and real-life scheduling choice situations, we will be able to draw conclusions on travel time perceptions. Here, based on comparing actual and reported travel times, we can only conclude that some, but not all findings are consistent with the literature on time perception. For instance, the result that those drivers tend to overstate travel times more who have little experience in commuting is in line with the research on time perception in general (e.g. Boltz et al., 1998). On the contrary, the finding that (different indicators of) travel time variability turned out insignificant in the regressions contradicts the suggestion of the relevant literature that activities that are not well predictable are perceived as longer than they are (Y. Li, 2003; Van de Ven et al., 2011).

### 3 Choice models: Theoretical framework

#### 3.1 Introduction

In this section, the theoretical framework is introduced that is used to test whether the differences between reported and actual travel times, as described in the previous section, are also reflected in travel choice behavior. We model travel choice behavior in the form of scheduling decisions. The workhorse model for the analysis of scheduling choices has been introduced by 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 the 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. Noland and Small, 1995; Small, 1982). Small (1982) was the first to estimate monetary valuations of travel time and schedule delays, while Noland and Small (1995) were the first to incorporate stochasticity of travel times in the scheduling model.

In line with this literature, our model assumes that commuters choose their optimal daily 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 as follows:

$$\begin{aligned} \text{SDE} &= \max[\text{PAT} - t_h - T(t_h), 0] \\ \text{SDL} &= \max[t_h + T(t_h) - \text{PAT}, 0], \end{aligned} \quad (3)$$

where  $T(t_h)$  denotes the travel time associated with the departure time from home,  $t_h$ . Note that for simplicity, we drop the indices related to individual  $z$ , alternative  $j$  and choice situation  $k$ . We account for stochastic travel times by taking into account the expectations  $E[\cdot]$  of the attribute values. The coefficients associated with the reward, travel time and schedule delays early and late are then denoted by  $\beta_R, \beta_T, \beta_E$  and  $\beta_L$ . The expectation of the systematic part of the utility function,  $E[V]$  assumes the following additive form:

$$E[V] = \beta_R * E[R] + \beta_T * E[T] + \beta_E * E[\text{SDE}] + \beta_L * E[\text{SDL}] \quad (4)$$

As the reward coefficient, multiplied by a factor (-1), represents the marginal utility of income, 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 times and schedule delays, respectively, by one hour.

$$\text{VOT} = -\beta_T/\beta_R \quad \text{VSDE} = -\beta_E/\beta_R \quad \text{VSDL} = -\beta_L/\beta_R \quad (5)$$

#### 3.2 RP model

The systematic part of the utility function associated with RP-based departure time choices includes the additive terms comprised in  $E[V]$  (see Eq. 4). 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 (with the expected utility denoted by  $E[V_{RP}^T]$ ), additional terms must be added to the model: 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 difference terms for all attributes as it would be strange, if travelers overestimated travel times when considering trip duration, but not when considering other time-dependent attributes, notably schedule delays and rewards. So, for instance, a driver who overestimates travel times is expected to underestimate schedule delays early and overestimate schedule delays late. For a given departure time he expects to arrive later than a person who departs at the same moment and does not misperceive travel times. The coefficients corresponding to these difference terms as well as the attribute values that would result if travel times were misperceived in accordance with travel time ratio  $\tau$  are labelled with the superscript  $\tau$ .

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

The coefficients related to the differences 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.

### 3.3 SP model

Similar to the RP model, the utility function of the SP model includes the additive terms comprised in  $E[V]$  (Eq. 4). To investigate whether reporting errors reflect misperceptions and thus affect behavior in an SP environment, also the difference between the travel time attribute that would result if travel times were affected by a perception error consistent with the travel time ratio and the objectively measured time attribute must be included in the regarding expected utility function  $E[V_{SP}^T]$ . In contrast to the RP model, this difference term is only included for the travel time attribute, but not for the reward and scheduling attributes. The reason is that in the SP experiment, rewards as well as actual arrival times at work are presented to the respondents. As a consequence, it is unlikely that respondents perceive the reward and the schedule delays differently from the ones presented to them,



even if they misperceive travel times.

$$E[V_{SP}^{\tau}] = \beta_R * E[R] + \beta_T * E[T] + \beta_T^{\tau} * (E[T^{\tau}] - E[T]) + \beta_E * E[SDE] + \beta_L * E[SDL] \quad (7)$$

The interpretation of  $\beta_T^{\tau}$  is similar to its interpretation in the RP utility function, with the difference that here we expect a positive  $\beta_T^{\tau}$  in the case of overestimation (hence, the reverse sign compared to  $\beta_T$ ): 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. If this was found true, the monetary value attached to the difference between the travel times multiplied by the travel time ratio and the objectively measured travel times ( $VOT^{\tau} = -\beta_T^{\tau}/\beta_R$ ) would be negative, hence, lowering the overall value attached to travel time ( $VOT + VOT^{\tau}$ ) as suggested by Brownstone and Small (2005).

## 4 Choice models: Data

### 4.1 RP data

Compared to SP data, RP data have the advantage that they are based on real choices rather than hypothetical ones, and 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 and which constitute the most prominent criticism against the use of SP data. 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)

In this study, 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 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. Again the model for the estimation of door-to-door travel times developed by Peer et al. (2011) is used. Moreover, we use it to approximate the chosen departure time at the home location. Choice situations here refer to days during which drivers have been observed to pass the C1–C2 link, along which the reward experiment took place. Our models thus describe the departure time choice conditional on making a commute trip, and the overall number of RP choice situations is therefore driver-specific. Departure time alternatives are defined as discrete intervals, each of them being equal to a 15-minute interval during the morning peak. Overall, there are 17 choice alternatives (ranging from 5:30 a.m. to 9:45 a.m.). The chosen departure times need

to be situated within this time frame, otherwise the according observations 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 also different scheduling preferences apply than for commuting trips.

The person-specific PAT that applies for regular commuting trips (and relative to which the schedule delays are defined) is derived from the same questionnaire as the reported travel times that are used for the computation of the ‘standard’ formulation of the travel time ratio. 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.<sup>8</sup>

Rewards vary by time of the day. They are equal to 4 Euro if a driver passes C2 before 6:30 a.m. or C1 after 9:30 a.m., and are equal to 0 otherwise. Moreover, the maximum number of rewarded trips per week cannot exceed the weekly number of peak trips that a driver has undertaken before the start of the experiment (reference behavior). Once the maximum number of rewards has been reached, no further rewards are distributed during the rest of the week. On these days, the reward is thus equal to 0 for all possible departure time choices. Holidays and weekends are excluded from the analysis, not only because all participants were unable to gain a reward on those days, but also because observed travel on these days frequently does not constitute commuting trips and then the PAT as defined above usually does not apply.

Given our knowledge of the PAT, the rewards and the driver-, day- and time-of-day-specific door-to-door travel times, which are again computed using the method presented in Peer et al. (2011), the attribute levels specific to the choice situation (i.e. day)  $k$ , individual  $z$  and choice alternative  $j$  can be derived. Expected attribute levels for attributes  $A = \{R, T, SDE, SDL\}$  are based on a compound measure of actual, time-of-day-specific travel times on the day of travel,  $A_{zkj}$ , and the time-of-day-specific average travel times,  $\bar{A}_{zj}$  over the duration of the experiment.<sup>9</sup> While actual realizations are in principle unknown in advance (hence, at the time 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 pattern of travel times over the time of the day, which participants of the experiment are likely to be aware of. We estimate the weight  $\theta$  that drivers attach to the actual realizations relative to the averages;  $\theta$  will be equal to 0 if travel time expectations are based exclusively on the averages, while it will be equal to 1 if they are perfectly informed about actual realizations and base their decisions on this knowledge. The expectation for attribute  $A$  of the utility function for driver  $z$ , choice situation  $k$  and departure time choice alternative  $j$  is thus given by:

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

---

<sup>8</sup>Only about 3% of all participants that filled in the questionnaire have a PAT that is either before 6:30 or after 9:30.

<sup>9</sup>Y. Tseng et al. (2011) use a similar definition of expected travel times, however, without applying the compound measure to all attributes of the utility function but to travel times only.

Finally, 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. 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.<sup>10</sup> Based on these we compute the other attribute values (rewards and schedule delays) that would be relevant if travel times were misperceived in a similar way as they were misreported ( $A_{z_kj}^\tau$ ). By averaging these over the number of working days (during the time the peak avoidance experiment took place), we obtain the remaining component,  $\bar{A}_{zj}^\tau$ , of the compound attribute expectation measure from Eq. 8.

## 4.2 SP data

In most SP experiments on scheduling behavior, respondents are asked to make hypothetical choices between alternatives with varying departure times, which may differ in terms of costs, travel time and 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. Also, 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 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 is also adopted in the SP experiment considered in this paper. All travel times shown to the respondents in the SP experiment are designed such that they were 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 each. 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 certain probability that was provided to the respondents. For a given departure time, the variation in travel times also induces variations in schedule delays and, in some cases, in the reward. Just as in the RP experiment, a reward of 4 Euro only 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.

<sup>10</sup>Note that we do not differentiate between the different magnitudes of the travel time ratio on the various sub-links. The main reasoning is the large number of outliers on the sub-links, which are less evident on the home–work stretch.

Figure 2: Screenshot

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

	Alternative A		Alternative B	
Departure time from home	<b>6:05</b>		<b>6:50</b>	
Probability	<b>80%</b>	<b>20%</b>	<b>90%</b>	<b>10%</b>
Total travel time	<b>30 min</b>	<b>40 min</b>	<b>20 min</b>	<b>35 min</b>
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	<b>6:35</b>	<b>6:45</b>	<b>7:10</b>	<b>7:25</b>
Reward	<b>4 euro</b>	<b>4 euro</b>	<b>0 euro</b>	<b>0 euro</b>

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, are not taken into account in the choice models. 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).<sup>11</sup> A detailed description of the design of the SP experiment can be found in Knockaert et al. (2012).

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 simply multiplied by the ‘standard’ definition of the person-specific travel time ratio.

### 4.3 Dataset descriptives

Table 4 shows some descriptive statistics for those participants who are taken into account in the SP and RP scheduling choice models. It shows that the majority of the participants is between 30 and 50 years old, around a quarter of the participants is female and, probably most notable, many of them have a fairly high income. The table also provides insights on the distribution of the PAT. It shows to be clustered between 7:30 and 8:30 for more than 60% of the participants. The travel time ratio is almost identical to the travel time ratio derived for the entire sample (using the ‘standard’ definition).<sup>12</sup>

<sup>11</sup>Moreover, the layout and the phrasing was tested using focus groups and an online test.

<sup>12</sup>The sample is here smaller than in the computations of travel time ratio, as we exclude drivers who stated that they made their choices randomly in the SP experiment as well as those who indicate a PAT during off-peak hours.

Table 4: Descriptive statistics

Variables	Value or Fraction of Sample
<i>Socio-economic characteristics</i>	
Age	
< 30 years	0.10
> 50 years	0.20
Female	0.27
Monthly (net) income	
< 3500 Euro	0.33
> 5000 Euro	0.15
unknown	0.22
Households with kids	
0 – 5 years	0.22
6 – 10 years	0.22
11 – 15 years	0.17
<i>Commute-related variables</i>	
Preferred arrival time (PAT)	
<7:30 a.m	0.16
>8:30 a.m.	0.21
Mean travel time ratio ( $\tau$ )	1.59
<i>Choices</i>	
Nr. of individuals	406
Total nr. SP choices	4060
Total nr. RP choices	8477

## 5 Choice models: Estimations

### 5.1 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. While we estimate the MNL models separately for SP and RP data, a latent class model is estimated that pools the two data sources. All models in this section are introduced in the more general, 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 (Louviere and Hensher, 2001; 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 utility  $U_{z kj}$  as the utility associated with choice  $k = 1, \dots, K_z$ , individual  $z = 1, \dots, Z$  and choice alternative  $j = 1, \dots, J_{zk}$ . It consists of the systematic component as defined in Eq. 4 and a stochastic component  $\epsilon_{z kj}$ . The definition of the systematic component differs between the data sources as long as not all coefficients are common across SP and RP observations. We denote the RP- and SP-based systematic components by  $V_{z kj}^{RP}$  and  $V_{z kj}^{SP}$ , respectively. 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 ( $k \in \{1, \dots, 10\}$ ) and equal to 1 for RP choices ( $k \in \{11, \dots, K_z\}$ ). 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  $\lambda$  that is relevant for SP observations (while the scale of RP observations is fixed to 1).<sup>13</sup> The random utility function to be maximized is then given by:

$$U_{z kj} = \mathbf{1}^{RP} * V_{z kj}^{RP} + (1 - \mathbf{1}^{RP}) * \lambda * V_{z kj}^{SP} + \epsilon_{z kj} \quad (9)$$

In the estimation of the MNL models, the random component  $\epsilon_{z kj}$  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).

The probability of driver  $z$  choosing alternative  $j$  in choice  $k$ ,  $P_{z kj}$ , is then defined as follows:

$$P_{z kj} = \mathbf{1}^{RP} * \frac{\exp(V_{z kj}^{RP})}{\sum_{j=1}^{J(zk)} \exp(V_{z kj}^{RP})} + (1 - \mathbf{1}^{RP}) * \frac{\exp(V_{z kj}^{SP})}{\sum_{j=1}^{J(zk)} \exp(V_{z kj}^{SP})} \quad (10)$$

The corresponding loglikelihood function is given below. For simplicity, we use the notation of  $\bar{P}_{zk}$  to indicate the chosen alternative in choice  $k$  by driver  $z$ :

$$\ln L = \sum_{z=1}^Z \sum_{k=1}^{K_z} \bar{P}_{zk} \quad (11)$$

Moreover, we use a panel latent class model<sup>14</sup>, 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  $q = 1, \dots, Q$  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 of a specific driver 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

<sup>13</sup>Differences in scale are only relevant if common coefficients for the SP and RP observations are estimated.

<sup>14</sup> A general discussion on latent class analysis can for instance be found in McCutcheon (1987).

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<sup>15</sup> other than a constant), which ensures that relative class sizes sum up to 1 without actively restricting the parameter values. The loglikelihood function is then given by the following equation:

$$\ln L = \sum_{z=1}^Z \ln \left[ \sum_{q=1}^Q H_{zq} \left( \prod_{k=1}^{K_z} \bar{P}_{zk|q} \right) \right], \quad (12)$$

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

Unlike for mixed logit models, which are the most common method to capture unobserved heterogeneity in discrete choice analysis, no assumptions regarding the shape of the distribution of a given parameter across individuals is required when estimating panel latent class models. Only the number of classes needs to be set by the analyst. Moreover, for a fairly small number of classes, it is usually not necessary to assume that specific coefficients are constant for all individuals. Such an assumption is frequently made in mixed logit models in order to ensure convergence of the regarding models.

## 5.2 Estimation results: Multinomial logit models

Table 5 provides the results for the multinomial logit (MNL) models. In contrast to the panel latent class model presented in the next section, RP and SP data have not been pooled in the MNL models. 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. All monetary valuations<sup>16</sup> are provided in Euro per hour.

Unlike in 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 – especially considering that 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). Our 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 rather coincidental finding that disappears once heterogeneity across drivers is considered. We will show some evidence for the latter in the next section of the paper.

<sup>15</sup>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 coefficient estimates are mainly a result of unobserved heterogeneity.

<sup>16</sup>The corresponding t-statistics are computed using the Delta method (e.g. Daly et al., 2012).

Table 5: MNL models

Coefficient	Without $\tau$				With $\tau$			
	RP only		SP only		RP only		SP only	
	Value	t-Stat.	Value	t-Stat.	Value	t-Stat.	Value	t-Stat.
$\beta_R$	0.19	5.52	0.05	4.20	0.22	5.58	0.05	4.18
$\beta_T$	-7.01	-6.45	-2.37	-13.27	-7.43	-6.25	-2.90	-8.27
$\beta_E$	-1.93	-15.13	-1.59	-16.52	-1.90	-12.65	-1.59	-16.53
$\beta_L$	-1.55	-17.30	-1.14	-13.25	-1.56	-16.61	-1.14	-13.27
$\theta$	0.14	5.28	–	–	0.14	5.07	–	–
$\beta_R^\tau$	–	–	–	–	0.05	1.21	–	–
$\beta_T^\tau$	–	–	–	–	–	–	0.79	1.92
$\beta_E^\tau$	–	–	–	–	-1.60	-1.41	–	–
$\beta_L^\tau$	–	–	–	–	1.74	1.50	–	–
VOT	36.32	4.08	44.38	3.94	34.24	4.08	54.61	3.68
VSDE	10.00	5.17	29.78	4.40	8.76	5.10	29.92	4.38
VSDL	8.03	5.21	21.35	4.33	7.19	5.24	21.47	4.31
Nr. Obs.	8477		4060		8477		4060	
LogLik.	-20386		-2223		-20370		-2220	
Pseudo R <sup>2</sup> adj.	0.151		0.208		0.152		0.209	

The finding that the VOT is relatively high in all models can probably be explained by the fact that many participants belong to the upper income classes (see Table 4). Another possible explanation is that the travel time coefficient incorporates the disutility from travel time variability, which is strongly correlated to travel time, and therefore not estimated separately. And finally, we consider fairly long trips (with an average duration of ca. 43 minutes: see Table 1), which tend to be associated with a higher VOT (Daly and Carrasco, 2009).

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.<sup>17</sup> A possible cause is 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. In both the SP- and the RP-based standard scheduling models (without  $\tau$ ) 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,  $\theta$ , 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 are updated by day-specific information.

<sup>17</sup>This is in contrast to the results obtained in the meta-analysis by Y.-Y. Tseng et al. (2005). Other studies, however, do not find any pattern in the relation of RP- and SP-based scheduling values (e.g. Z. Li et al., 2009).



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 model, as it makes it difficult to test whether the travel time ratio affects choice behavior with the current formulation of the travel time ratio as a person-specific constant. For consistency reasons, the travel time ratio would then have to be formulated as a (possibly also non-linear) function of the observed travel time, whose derivation is not feasible for most drivers because of the rather low number of observed travel time observations.

The third and fourth model presented in Table 5 take into account the travel time ratio in the RP and SP setting, respectively. In both models, we use the ‘standard’ definition for the travel time ratio from Table 1. 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.

Due to correlations in the choice attributes of the RP model when the travel time ratio is taken into account, we are not able to estimate the full model as suggested in Eq. 6, which considers the differences between the attribute values that would apply if participants misperceived travel times to the same extent as they misreport them and the objectively measured attribute values for all attributes. Instead we leave out this difference term for the travel time attribute. If the difference terms for the remaining three attributes were significant, we would 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. This means that we 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 they make their scheduling decisions in reality.

In the SP-based model that accounts for the travel time ratio,  $\beta_T^\tau$  is found to be positive, as suggested by Brownstone and Small (2005).<sup>18</sup> 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.

Before proceeding with the panel latent class model, 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 worse, 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.

For the RP model, with the truncation of the travel time ratio distribution at 2, the results remain similar to the original estimates from Table 5. If all persons with travel time ratios exceeding 1.5 are excluded from the analysis, the scheduling coefficients consistent

---

<sup>18</sup>This also explains the higher VOT, as the VOT based on  $\beta_T^\tau$  would be negative due to  $\beta_T^\tau$  being positive.

Table 6: Estimation results with truncated travel time ratios  $\tau$ 

Trunc.	Data	Nr. Obs.	$\beta_R^\tau$		$\beta_T^\tau$		$\beta_E^\tau$		$\beta_L^\tau$	
			Value	t-Stat.	Value	t-Stat.	Value	t-Stat.	Value	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	–	–	–	–

with a misperception similar to the travel time ratio (indicated by superscript  $\tau$ ) 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 being significant only above the 30% level, and  $\beta_E^\tau$  being positive (although it should be negative if travel times were misperceived in a similar way as they were misreported), give quite a strong indication that this is not the case.

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,  $\beta_T^\tau$  is still positive, however, less close to significance at the 5% level than in the original estimation (the t-statistic equals 1.49 compared to 1.92 in the original estimation). If only travel time ratios below 1.5 are considered in the estimation, the t-statistic of  $\beta_T^\tau$  decreases even further (to -0.22) and the sign of the coefficient becomes negative (whereas a positive sign would be expected based on the proposition of Brownstone and Small (2005)). 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 imply that drivers react to the travel times presented to them as if they would when facing the same travel time in reality. Instead, it might simply imply that the travel time ratio is not an appropriate representation of travel time perceptions. We will further investigate this matter in the next section by estimating a panel latent class model using both SP and RP data as input.

### 5.3 Estimation results: Panel latent class model

The MNL model presented in Table 5 and the robustness check presented in Table 6 seem to suggest that the theory brought forward by Brownstone and Small (2005) does not hold for our data (provided that the misreported travel times are a good representation of travel time misperceptions). To further strengthen the evidence, 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 better be modeled. And therefore, this

estimation procedure should help identifying 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 the statistical criteria would suggest a higher number of classes. But probably due to the correlations in the attribute values, which are especially evident in the RP setting, several coefficients can then assume the reverse sign of what one would expect, leading to negative valuations of time and schedule delays, which are clearly not useful 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.

Table 7: Panel latent class model

Coefficients	Class 1			Class 2			Class 3	
	Value	t-Stats.		Value	t-Stats.		Value	t-Stats.
Class Prob.	0.24			0.32			0.44	
$\beta_R^{RP}$	0.18	3.61	= <sup>a</sup>	0.18	3.61		0	–
$\beta_T^{RP}$	-18.6	-5.95		-5.73	-3.83		-8.92	-6.03
$\beta_E^{RP}$	-1.39	-5.12		-2.62	-9.21		-3.45	-13.48
$\beta_L^{RP}$	-2.03	-6.34		-0.68	-7.00		-3.04	-5.23
$\theta$	0.12	5.30	= <sup>a</sup>	0.12	5.30	= <sup>a</sup>	0.12	5.30
$\beta_R^{SP}$	0.15	6.10	= <sup>a</sup>	0.15	6.10		0	–
$\beta_T^{SP}$	-5.77	-3.70		-2.63	-3.76	= <sup>a</sup>	-2.63	-3.76
$\beta_E^{SP}$	-0.57	-1.32		-3.07	-2.87		-2.65	-9.42
$\beta_L^{SP}$	-0.74	-3.34		-0.67	-1.79		-3.03	-5.23
$\beta_T^{SP,\tau}$	2.25	1.27		-0.50	-0.22		0.97	1.30
VOT <sup>RP</sup>	101.64	3.04		31.31	2.54		–	–
VSDE <sup>RP</sup>	7.60	2.96		14.32	3.36		–	–
VSDL <sup>RP</sup>	11.09	3.17		3.69	3.26		–	–
VOT <sup>SP</sup>	39.79	3.17		18.14	3.23		–	–
VSDE <sup>SP</sup>	3.92	1.33		21.17	2.70		–	–
VSDL <sup>SP</sup>	5.08	3.08		4.63	1.75		–	–
Nr. Obs.				12537				
LogLik.				-20110				
Pseudo R <sup>2</sup> adj.				0.122				

<sup>a</sup> Restricted as equal

In the model presented in Table 7, all coefficients are specific to one of the datasources, meaning that no scale parameter needs to be (and even can be) estimated. 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 (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 SP and RP choices.

Based on the results of alternative model formulations and the corresponding estimation results, some coefficients have been restricted.<sup>19</sup> So, 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 reward coefficients for this class is 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. The class probability of 'Class 3' (0.44) indicates that a considerable share of participants do not consider the reward when deciding on their actual and hypothetical departure times. A possible cause for this finding might be that a reward of 4 Euro is too low for some participants to shift their departure times to off-peak periods, or that their scheduling restrictions are too stringent. Furthermore, we restrict the SP and RP reward coefficients such that they assume the same value for 'Class 1' and 'Class 2', and also the SP time coefficient for the second and third class is restricted to be equal. The main reason for these two restrictions is that the corresponding coefficients would not be significant if estimated class-specific, meaning that we would not be able to derive statistically significant point estimates for the monetary valuations either. Finally, we do not specify  $\theta$  as class-specific, as it does not differ significantly across classes.

Given these restrictions, we find that all remaining willingness-to-pay estimates except for two SP-based scheduling values are significant at the 5% level. The results for the RP and SP valuations are consistent within and across classes, yet not (statistically) equal. So, in both domains members of 'Class 1' have a high VOT, and their scheduling valuations are moderate, with the VSDL being higher than the VSDE. The VOTs in 'Class 2' are substantially lower than in 'Class 1', for RP as well as SP choices. Moreover, unlike in the first class, in the second class the VSDE exceeds the VSDL (again for both data sources). Moreover, the RP- and SP-based VSDEs for 'Class 2' are higher than any of the scheduling valuations in 'Class 1'. Naturally, no willingness-to-pay estimates can be computed for the third class, where the reward coefficient has been fixed to 0.

Despite the consistency of SP- and RP-based attribute valuations within and across classes, the point estimates of the valuations differ considerably, in particular the VOT: It is more than double as high for the RP than for the SP choices in 'Class 1', and more than 2/3 higher in 'Class 2'. Using the panel latent class model, we therefore obtain results that are similar to what is usually found when comparing SP and RP estimates on the same type of choices (usually made by the same individuals), namely that the SP-based willingness-to-pay estimates are lower than their RP counterparts.<sup>20</sup> This result remains valid in models where the travel time ratio is not considered ( $\beta_T^{SP,\tau} = 0$ ).

<sup>19</sup>If the restrictions cover multiple classes of the same datasource, they are indicated by '=' in Table 7.

<sup>20</sup>At least in studies involving time valuation this is a frequent outcome (e.g. Abrantes and Wardman, 2011; Shires and De Jong, 2009).

The coefficient related to the objective (SP) travel time multiplied by the travel time ratio,  $\beta_T^{SP,\tau}$ , is 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 here the differences between SP and RP willingness-to-pay values by misperceptions that the travel time ratio is able to measure. The most prominent alternative explanation are hypothetical biases that affect the SP estimates. So, drivers might be more sensitive towards the reward in a hypothetical context than in a real-life setting. Brownstone and Small (2005) suggest that such an increased reward sensitivity in an SP environment 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 considered in the hypothetical choices.

Finally we find that the VSDE and VSDL are more similar between the SP and RP domain in the panel latent class model compared to the MNL models. We can observe that this is mainly due to the higher SP reward coefficient in ‘Class 1’ and ‘Class 2’ compared to the MNL model, which does not only lower the corresponding VOTs but also the schedule delay values. Based on this result, the scheduling preferences in SP and RP choice situations seem to be fairly comparable, once heterogeneity across drivers is taken into account.

## 6 Conclusions

In this study, we first compare reported and actual travel times of participants of a large-scale peak avoidance experiment by computing the person-specific ratio between reported and actual travel times, which we refer to as ‘travel time ratio’. We find very robust evidence that on average the participants have a travel time ratio of 1.5, meaning that they overstate travel times by 50%. We are able to identify some main determinants of the travel time ratio, such as link length and average speed, and to test some of the hypotheses brought forward in the literature regarding link- and driver-specific characteristics that may explain overreporting. Next, we investigate whether the overstated travel times are the result of reporting errors or indeed reflect travel time misperceptions. We do so by testing whether the departure time choice behavior of the participants in reality as well in a hypothetical choice experiment is affected by the person-specific travel time ratio. This does not seem to be the case, as the travel time ratio has only little explanatory power for the choice behavior in both revealed preference (RP) and stated preference (SP) settings. We can therefore conclude that the travel time ratio is mainly a reporting error, and that reported travel times may be a poor indicator of actual and perceived travel times.

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 might also have taken into account the time they take to get from their home to their car as well as the time they require to get from their car to their work place. Second, the average travel time ratio being 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. Third, also strategic behavior might play a role in the answers provided in the questionnaire. By overstating travel times, the respondents might hope to influence policy makers such that they implement 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 actual scheduling behavior. However, for both cases there is no clear reasoning why overstating rather than understating of travel times would result. We therefore do not expect them to be the main driver behind the results obtained in this paper.

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 arising from respondents not only considering in-vehicle time in their answers. Additionally, 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. Moreover, asking for a set of previously experienced travel times rather than averages might help to reduce differences between reported and actual travel times, since this excludes 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.

In this paper, the main focus of the SP- and RP-based choice models was to test whether drivers misperceive travel times to a similar extent as they misreport them. If this was found true, we would conclude that the travel time ratio represents well travel time perceptions. However, since 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 admittedly difficult task of identifying (instrumental) variables that affect the travel time misperception, but not the preferences that drive the scheduling choices.

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 drawn in a wider sense.

## References

- Abrantes, P. A., and 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., and Lindsey, R. (1993). A structural model of peak-period congestion: A traffic bottleneck with elastic demand. *The American Economic Review*, 161–179.
- Arnott, R., De Palma, A., and Lindsey, R. (1994). The welfare effects of congestion tolls with heterogeneous commuters. *Journal of Transport Economics and Policy*, 139–161.
- Baldauf, D., Burgard, E., and Wittmann, M. (2009). Time perception as a workload measure in simulated car driving. *Applied ergonomics*, 40(5), 929–935.
- Block, R. A., Hancock, P. A., and Zakay, D. (2010). How cognitive load affects duration judgments: A meta-analytic review. *Acta psychologica*, 134(3), 330–343.
- Block, R. A., and Zakay, D. (1996). Models of psychological time revisited. In H. Helfrich (Ed.), *Time and mind* (pp. 171–195). Kirkland, WA, USA: Hogrefe & Huber.
- Boltz, M. G., Kupperman, C., and 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., and 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*, 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., and Carrasco, J. (2009). The influence of trip length on marginal time and money values. In R. Kitamura, T. Yoshii, and T. Yamamoto (Eds.), *The expanding sphere of travel behavior research: Selected papers from the 11th International Conference on Travel*. Bingley, UK: Emerald Group.
- Daly, A., and Hess, S. (2011). Simple approaches for random utility modeling with panel data. In *Transportation Research Board 90th Annual Meeting*. Washington, DC.
- Daly, A., Hess, S., and Jong, G. de. (2012). Calculating errors for measures derived from choice modelling estimates. *Transportation Research Part B: Methodological*, 46(2), 333–341.
- Droit-Volet, S., and Meck, W. H. (2007). How emotions colour our perception of time. *Trends in cognitive sciences*, 11(12), 504–513.
- Eisler, A., Eisler, H., and Montgomery, H. (2007). A quantitative model for retrospective subjective duration. *NeuroQuantology*, 2(4).
- Eisler, H. (1976). Experiments on subjective duration 1868-1975: A collection of power function exponents. *Psychological Bulletin*, 83(6), 1154.

- Flaherty, M. (1999). *A watched pot: How we experience time*. New York, NY, USA: New York University Press.
- Ghosh, A. (2001). *Valuing time and reliability. Commuters' mode choice from a real time congestion pricing experiment*. Phd thesis, University of California at Irvine, 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., and 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., and 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. (2001). Measurement of the valuation of travel time savings. *Journal of Transport Economics and Policy (JTPE)*, 35(1), 71–98.
- Hensher, D. (2010). Hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B: Methodological*, 44(6), 735–752.
- Isacsson, G. (2007). The trade off between time and money: Is there a difference between real and hypothetical choices? *Working Papers*.
- Knockaert, J., Eetema, D., Verhoef, E., Koster, P., Peer, S., and Tseng, Y.-Y. (2012). *Spitsmijden Gouda-Zoetermeer* [Scientific Report].
- Koster, P., and Tseng, Y. (2010). Stated Choice Experimental Designs for Scheduling Models. In *Choice modelling: The state-of-the-art and the state-of-practice: Proceedings from the Inaugural International Choice Modelling Conference* (p. 217).
- Leiser, D., and Stern, E. (1988). Determinants of subjective time estimates in simulated urban driving. *Transportation Research Part A: General*, 22(3), 175–182.
- Li, Y. (2003). Evaluating the urban commute experience: A time perception approach. *Journal of Public Transportation*, 6(4), 41–66.
- Li, Z., Hensher, D., and Rose, J. (2009). Valuation of Travel Time Reliability in an Extended Expected Utility Theory Framework.
- Louviere, J., and Hensher, D. (2001). Combining sources of preference data. In D. Hensher (Ed.), *Travel behaviour research: The leading edge* (pp. 125–144). Oxford: Pergamon.
- Louviere, J., Hensher, D., and Swait, J. (2000). *Stated choice methods: analysis and applications*. Cambridge, UK: Cambridge University Press.
- McCutcheon, A. (1987). *Latent class analysis. Sage University Paper Series on Quantitative Applications in the Social Sciences* (No. 07-064). Newbury Park, CA: Sage Publications.
- McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics*. New York: Academic Press.
- MVA Consultancy. (1987). *The Value of Travel Time Savings: A Report of Research undertaken for the Department of Transport* [prepared by the MVA Consultancy, Institute of Transport Studies, University of Leeds and Transport Studies Unit]. University of Oxford, Newbury, Berkshire, UK: Policy Journals.



- Noland, R., and Small, K. (1995). Travel-time uncertainty, departure time choice, and the cost of the morning commute. *ITS Working Paper Series, UCI-ITS-WP-95-1*.
- O'Farrell, P., and Markham, J. (1974). Commuter perceptions of public transport work journeys. *Environment and Planning A*, 6(1), 79–100.
- Parthasarathi, P., Hochmair, H., and Levinson, D. (2011). Network structure and travel time perception. *Working Paper, Nexus Research Group, University of Minnesota*.
- Peer, S., Knockaert, J., Koster, P., Tseng, Y., and Verhoef, E. (2011). Door-to-door travel times in RP departure time choice models. An approximation method based on GPS data. *Tinbergen Institute Discussion Papers*.
- Recarte, M., and Nunes, L. (1996). Perception of speed in an automobile: Estimation and production. *Journal of Experimental Psychology: Applied*, 2(4), 291.
- Rietveld, P., Zwart, B., Van Wee, B., and 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., and 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., and 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. *The American Economic Review*, 72(3), 467–479.
- Small, K., and Verhoef, E. (2007). *The economics of urban transportation*. London, UK: Routledge.
- Small, K., Winston, C., and Yan, J. (2005). Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica*, 73(4), 1367–1382.
- Swait, J., Louviere, J. J., and 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, UK: Cambridge University Press.
- Tseng, Y., Knockaert, J., and Verhoef, E. (2011). A revealed-preference study of behavioural impacts of real-time traffic information. *Transportation Research Part C: Emerging Technologies*(doi: 10.1016/j.trc.2011.11.006).
- Tseng, Y.-Y., Rietveld, P., and Verhoef, E. (2005). A meta-analysis of valuation of travel time reliability. In *Colloquium vervoersplanologisch speurwerk*. Rotterdam, Netherlands.
- Van de Ven, N., Rijswijk, L. van, and Roy, M. M. (2011). The return trip effect: Why the return trip often seems to take less time. *Psychonomic bulletin & review*, 18(5), 827–832.
- Van Exel, N., and 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. (1969). Congestion theory and transport investment. *The American Economic Review*, 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., and 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., and 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., and Levinson, D. (2005). Variation of the subjective value of travel time on freeways and ramp meters. In *84th Annual Meeting of the Transportation Research Board*. Washington, DC.