Value of time and reliability for urban pooled on-demand services

Alonso González, María; van Oort, Niels; Cats, Oded; Hoogendoorn-Lanser, Sascha; Hoogendoorn, Serge

DOI
10.1016/j.trc.2020.102621

Publication date
2020

Document Version
Final published version

Published in
Transportation Research. Part C: Emerging Technologies

Transportation Research Part C 115 (2020) 102621

Contents lists available at ScienceDirect

Transportation Research Part C

journal homepage: www.elsevier.com/locate/trc

Value of time and reliability for urban pooled on-demand services

María J. Alonso-González[a,*], Niels van Oort[a], Oded Cats[a], Sascha Hoogendoorn-Lanser[b], Serge Hoogendoorn[a]

[a] Dep. Transport and Planning, Civil Engineering and Geosciences, Delft University of Technology, P.O. Box 5048, 2600 GA Delft, the Netherlands
[b] KiM Netherlands Institute for Transport Policy Analysis, Postbus 20901, 2500 EX The Hague, the Netherlands

ARTICLE INFO

Keywords:
Value of time
Value of reliability
Ridesharing
DRT
Stated preference
Choice modelling

ABSTRACT

The uptake of on-demand services is increasing rapidly all over the world. However, the market share of their pooled version (ridesharing, e.g., UberPOOL or LyftLine) is still low, despite its potential in addressing the mobility challenges that dense urban cities are facing. In this research, we analyse user preferences towards pooled on-demand services regarding their time-reliability-cost trade-offs. We study, via stated preference experiments, the value of time (VOT) and value of reliability (VOR) of the different trip stages (waiting stage, in-vehicle stage, and transfer stage when combined with line-based public transport). We target urban Dutch individuals (N = 1006), and address commuting and leisure trips. Results show in-vehicle VOT for pooled on-demand services to amount to 7.88–10.80 €/h. These values are somewhat higher than known values of traditional public transport. We also find waiting VOT (both before the trip and during the transfer stage) to be lower than values previously reported in literature. In general, we find VOR to be lower than VOT: the reliability ratio (VOR/VOT ratio) for both the waiting stage and the in-vehicle stage being around 0.5. In order to understand different preferences, we also estimate latent class choice models. The analysis shows that the main difference between classes pertains to the overall time-cost and reliability-cost trade-offs (VOT and VOR values) rather than in different valuations of the reliability ratio. In addition to serving as input for demand forecasting models such as macroscopic static assignment and agent-based simulation models, our findings can support service providers in developing their strategy when designing pooled on-demand services.

1. Introduction

Urban transport is changing. Flexible transport services are appearing in cities as transport alternatives to traditional public transport and to privately owned modes. One of these types of services are pooled on-demand services (also known as Demand Responsive Transport (DRT) services). These can be offered by transit operators or by so-called Transport Network Companies (TNCs). UberPool, UberExpressPOOL, Shared Lyft, Ola Share or ViaVan are examples of pooled on-demand services currently in operation offered by TNCs.

They are enabled in large-scale settings thanks to the emergence of new technologies and ubiquitous communication. These services combine the flexibility of taxi services (flexible route and schedule) with the collective nature of public transport (different travel requests are matched together in the same vehicle). Therefore, they can potentially yield a service that offers some of the advantages associated with the private car while attaining some of the supply-side efficiency gains made possible by bundling travel

* Corresponding author.
E-mail address: m.j.alonsogonzalez@tudelft.nl (M.J. Alonso-González).

https://doi.org/10.1016/j.trc.2020.102621
Received 15 November 2019; Received in revised form 10 March 2020; Accepted 11 March 2020

demand.

Previous research has shown that pooled on-demand services can efficiently match current individual trips in urban areas with very little extra travel time for the users (Tachet et al., 2017). Ride pooling also allows for a better performance of on-demand systems, in comparison to individual-only systems (Liang et al., 2020; Vosooghi et al., 2019). Moreover, trip matching can contribute to large decreases in the congestion, pollution and space problems of urban areas, as shown in previous simulation studies (ITF, 2017, 2016). However, these models did not include an underlying behavioural model for describing user preferences in relation to pooled on-demand services; they assigned a fixed demand to these services – either the entire urban travel demand or the existing demand for taxi services. Unreliable behavioural attributes may lead to an under- or over-estimation of the benefits and thus mislead decision making.

In this study, we aim at estimating the time-reliability-cost trade-offs of individuals regarding pooled on-demand services. We study both the value of time (VOT) (i.e., the willingness to pay to reduce travel time) and the value of reliability (VOR) (i.e., the willingness to pay to reduce travel time variability/uncertainty). A good understanding of how individuals value reliability in pooled on-demand services is especially important given a) their flexible nature (with their lack of predefined schedules or routes, in contrast to traditional public transport); and b) their shared nature (with the related inherent reliability challenges, in contrast to the private alternative) (Li et al., 2019). We also analyse the reliability ratio (VOR/VOT).

To this end, we conduct three different stated preference (SP) experiments targeting individuals living in (sub)urban areas of the Netherlands. In them, we analyse the value of time and value of reliability of these services for the different trip stages: waiting time, in-vehicle time, and transfer time when the service feeds traditional public transport. The transfer stage for the intermodal scenario is included since research has found that integration of pooled on-demand services with public transport can complement first/last-mile transit access (Yan et al., 2019), improve overall system efficiency (Luo and Nie, 2019), and significantly enhance mobility (Stiglic et al., 2018). SP experiments provide respondents with hypothetical scenarios in order to analyse individuals' trade-offs. In settings where the studied service has not yet been implemented, this approach allows to estimate behavioural attributes; and in settings where these services are already present, this approach allows to (a) estimate behavioural attributes for individuals who have not opted in, and therefore avoid selection-bias of early-adopters, and (b) estimate preference towards attribute value ranges which are not yet observed and avoid the attribute collinearity often present in real situations.

Previous studies have evaluated the VOT for pooled on-demand services using SP experiments in different settings (e.g., Al-Ayyash et al. (2016) in Lebanon, Frei et al. (2017) in the USA, König et al. (2018) in Germany, and Ryley et al. (2014) in the UK). However, none of them analyse the VOR of these services, even though there is a common understanding that reliability is a fundamental determinant of travel behaviour (Carrion and Levinson, 2012; Jin et al., 2015; Prashker, 1979). One recent study, Bansal et al. (2019), did consider pick-up reliability in their on-demand SP experiment, and found that reliability plays an important role in the likelihood of passengers to choose these services. Their study, however, exclusively considers the average pick-up delay as reliability attribute portrayed to respondents (rather than a list of possible arrival values, which better represents the uncertainty/variability of the pick-up time). Also, the mentioned study does not consider cost as an attribute, and, therefore, the VOR cannot be calculated. As a result, our study contributes to literature by enabling the analysis of both VOT and VOR of pooled on-demand services and doing so for the different trip stages.

Our study also adds to previous research by including a segmentation approach regarding different time-reliability-cost sensitivities among travellers. Previous research has found that preference heterogeneity exists regarding adoption of different new mobility services (Alemi et al., 2019; Alonso-González et al., 2020b; El Zarwi et al., 2017), and pooled on-demand services in particular (Alonso-González et al., 2019). We identify different latent classes with different values of time and values of reliability. This segmentation can then help on-demand providers develop a portfolio of pooled on-demand services, targeting different market segments (as offered in Atasoy et al., 2015).

The remainder of the paper is structured as follows. Section 2 explains the methodology employed in this research; Section 3 presents the results; Section 4 discusses the implications of the findings for the design of pooled on-demand services, and Section 5 draws the final conclusions.

2. Methodology

The methodology section consists of three sub-sections. First, we discuss in Section 2.1 our research approach with regards to how to convey reliability to respondents. Second, Section 2.2 presents the design of our stated preference experiments given the considerations outlined in Section 2.1. Last, Section 2.3 covers the modelling approach for the analysis of the experiments.

2.1. Approach used to convey reliability to respondents

Due to the importance of the VOT and VOR in the assessment of travel demand studies (Carrion and Levinson, 2012), there is a large body of literature that investigates them (e.g., Bates et al., 2001; Brownstone and Small, 2005; Kouwenhoven et al., 2014; Tseng, 2008; Wardman, 2004). We refer the interested reader to Carrion and Levinson (2012), Li et al. (2010) and Wardman et al. (2016) for some recent literature reviews on the VOT and VOR.

In the literature, there is a common understanding of how to portray and analyse individuals' VOT. However, this differs for the analysis of the VOR. There are various conceptualisations of the notion of reliability and this has led to many discrepancies in how to convey reliability to respondents (which in turn, affect the subsequently obtained values). In the following, we highlight the three main distinctive aspects that play a role in how reliability has been operationalised in stated preferences studies in the literature, along with stating our own approach in relation to each of them:

Reliability representation – The first point of discrepancy lies in how to present reliability to respondents in SP experiments. Different representations have been used in literature, ranging from bar diagrams, histograms, circular arrangements, percentages to representing the likelihoods of certain travel times, time variability with a single time component (± X min), and verbal description of five equally probable travel times. Tseng et al. (2009) found the latter to be the most suitable one to represent reliability to respondents and to be appropriate for people with different levels of education. This representation, which was first used in Black and Towriss (1993) and Small et al. (1999), is still considered the state-of-the-art representation for SP preference reliability studies (Carrion and Levinson, 2012) and has been adopted in several recent studies (e.g., Asensio and Matas, 2008; Kouwenhoven et al., 2014; Swierstra et al., 2017). We also adopt it in this study.

Shape of the underlying reliability distribution – Less discussed, though still important, is the shape of the underlying distribution that is considered to obtain the five equally probable travel times. While Small et al. (1999) and Asensio and Matas (2008) used lognormal distributions for percentiles' values (which are the 10th, 30th, 50th, 70th and 90th percentiles of such distributions), Kouwenhoven et al. (2014) (see full report of the study under De Jong et al. (2007)) and Swierstra et al. (2017) consider additional parameters to quantify the increase in time of the points that form the travel time distributions. The first approach (use of a lognormal distribution) has two main advantages, namely: (1) this shape best represents real travel time distributions in field tests of pooled on-demand services (Lu et al., 2017) (as well as best fits in-vehicle travel times for bus (Kieu et al., 2014) and private car (Durán Hormazábal, 2016)); and (2) the ratios between the different percentiles shown follow a certain pattern (due to their common underlying shape), which avoids that the obtained results also depend on different shapes. For consistency reasons, we use the lognormal distribution also for the distribution of waiting times (which is expected to be right skewed as shown in Chen et al. (2017)) and waiting times during the transfer stage.

Two values are used as a starting point for the lognormal distributions: the planned time and the coefficient of variation (Cv)[1]. The μ and σ of the lognormal distribution (from which the 10th, 30th, 50th, 70th and 90th percentile values are calculated) relate to these two design attributes of the non-logarithmised distribution as follows:

$$\mu = \ln\left(\frac{planned\quad time}{\sqrt{1+\frac{Cv}{planned\quad time}}}\right) \tag{1}$$

$$\sigma = \sqrt{\ln\left(1+\frac{Cv}{planned\quad time}\right)} \tag{2}$$

In our study, we do not only consider variability as a source of unreliability, but also the deviation of the real (to be experienced) waiting/in-vehicle time from the expected (announced) waiting/in-vehicle time. For this we include an additional parameter, a systematic lateness that shifts the entire distribution (what we call displacement and is similar to the departure time shift included in Small et al. (1999)). The five equally probable values shown to respondents correspond to the previously rounded percentiles plus the displacement. Any posterior calculation during the modelling stage is performed based on these final values shown to respondents and not on the original lognormal distribution from which the values originate (same as in Noland et al. (1998)).

Reliability conceptual framework – There are two main conceptual frameworks to incorporate reliability in random utility choice models: the mean-variance method (which considers unreliability as the disutility of variability), and the scheduling method (which considers unreliability as the disutility of arriving early or late). It has been shown that the two methods are equivalent under certain assumptions (Fosgerau and Karlström, 2010) and that the mean-variance method is to be preferred on practical grounds, since the scheduling method does not directly yield a valuation of reliability (Bates et al., 2001). Given that our study aims not to test the disutility of arriving late for different individuals but rather the disutility of the variability of the offered service (to help design such services), we consider the mean-variance model as most suitable framework for this research.

2.2. Design of the stated preference experiments

Pooled on-demand services can have different characteristics. Ours offers a flexible route and schedule, and it is stop-to-stop (instead of door-to-door) with an average walking time of 1 min to the pick-up point (in line with findings from Zheng et al. (2019)). Fig. 1 shows the explanation presented to respondents in the context of our research. This explanation was slightly modified for individuals without internet connection in their mobile phone (communication via phone call and sms). We branded our service FLEXI, which provided respondents with an easy and intuitive name. Two trip purposes are investigated: commuting trips and leisure trips. Both are framed from home towards the work/leisure activity location. Each individual is assigned one trip purpose exclusively throughout the entire questionnaire (non-working individuals being always assigned to the leisure trip purpose).

We use three different SP experiments to investigate the three trip stages of (un)reliability in pooled on-demand services: waiting stage, in-vehicle stage, and transfer stage for intermodal trips. Figs. 2–4 show an example of a scenario of each of the experiments as an illustration. The large amount of calculations that are necessary to obtain the five equally probable values of the reliability

[1] The coefficient of variation is used instead of the standard deviation as a base attribute for the variability distribution because its standardised value makes it possible to compare the relative degree of variability for distributions with different means.

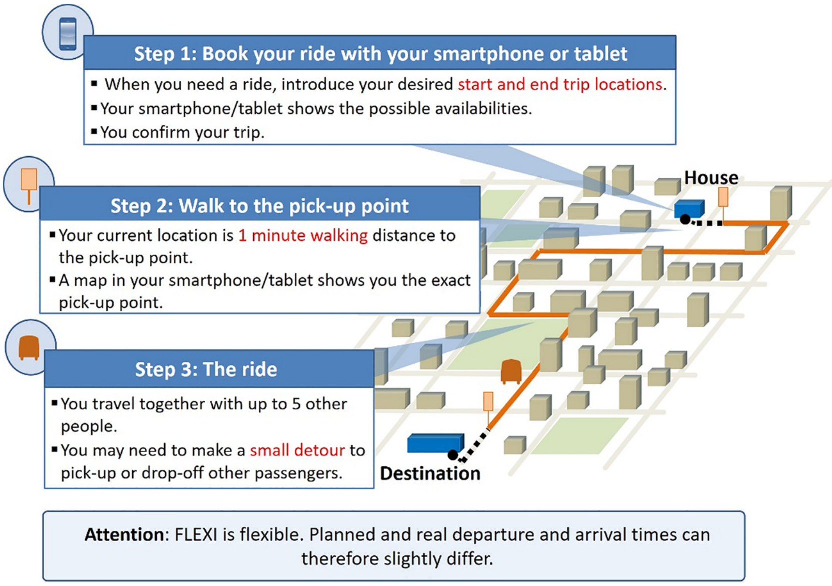

Fig. 1. Description of pooled on-demand services shown to respondents. Layout inspired from Kim et al. (2017).

Fig. 2. Example of a choice task of the waiting time SP experiment.

distributions makes it difficult to have pivoted experiments. To still present respondents with trip times they could relate to, we divide individuals in two segments: those with a reference trip of short distances (< 12 km) and those with a reference trip of medium distances (> 12 km). We present them values accordingly (short and medium version of the experiments), similar to the approach followed in Arentze and Molin (2013).

The experimental design is orthogonal fractional factorial with blocking. Each block consists of four scenarios for each of the two first experiments, and six scenarios for the third experiment. Different Likert-scale attitudinal indicators are placed between the second and third SP experiments as a break between the SP experiments. Each of the attributes has three attribute levels in order to be able to capture non-linearity. We use existing literature to get an indication of which range of values to include in our attribute levels (Chen et al. (2017) and Stiglic et al. (2016) for the time values, and Black and Towriss (1993) and Turnquist and Bowman (1980) for the coefficient of variation). The attribute levels are chosen to be non-equally spaced so as to obtain a larger number of trade-offs from the scenarios and not restrict the power of the design thereof (Small et al., 1999). All attribute levels of each of the experiments can be found in Appendix A.

The two first questionnaires (waiting time and in-vehicle time) are displayed in the classical format of VOT – VOR studies, with

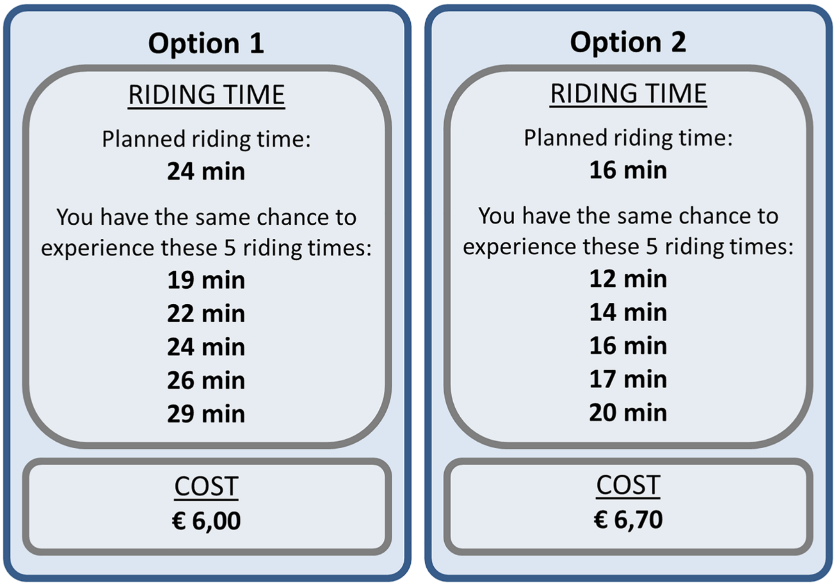

Fig. 3. Example of a choice task of the in-vehicle time SP experiment.

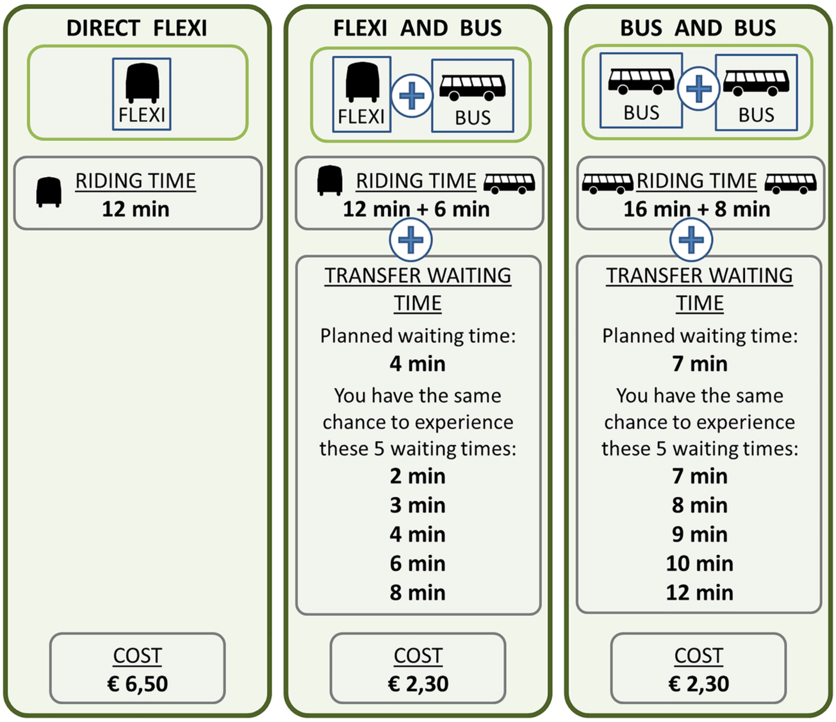

Fig. 4. Example of a choice task of the transfer SP experiment.

the time, reliability distribution and cost components as unique attributes (see Figs. 2 and 3). However, the layout of the transfer SP experiment is more complex (see Fig. 4). It has a larger number of attributes (given the importance of framing during the transfer stage) and it is a labelled experiment. Three alternatives are shown: a direct ride (to allow individuals opt out of transferring), a FLEXI + Bus alternative, and a Bus + Bus alternative. Garcia-Martinez et al. (2018) found that users perceive negatively transfers between different line-based public transport modes. Our design enables examining if a more tailored service (pooled on-demand services) reduces the intermodal transfer burden. Note that the FLEXI + Bus transfer may be perceived differently from the Bus + FLEXI transfer. However, in order to avoid increasing the cognitive workload too much, and given the limited/inexistent experience of respondents with pooled on-demand services, we limit our research to the FLEXI + Bus transfer type.

Reliability experiments rarely have layouts as complex as that of our transfer experiment. An exception is Swierstra et al. (2017), who also showed respondents a transfer SP experiment with the five equally probable values and additional attributes. However, in their research, finding models with significant reliability parameters proved to be difficult. Two explanations can help explain this outcome: 1) individuals are unable to correctly take reliability into account when additional parameters are added to the choice experiment, or 2) reliability attributes are considered negligible in comparison to other attributes in the final decision process (at least regarding the transfer stage). Our experiment can help identify which of the two explanations may play a larger role, given individuals' previous exposure to the other two similar (yet simpler) reliability experiments. To facilitate comparison in the obtained results, we choose to also specify a lognormal underlying reliability distribution for the transfer stage. This ensures consistency with the previous experiments.

2.3. Modelling approach

In this section, we discuss the modelling approach we follow to analyse the data obtained from the SP experiments.

2.3.1. Utility function specifications

We investigate the choices of individuals regarding pooled on-demand services using Logit discrete choice models under the random utility maximisation framework. This framework assumes that, when making choices, individuals try to maximise their utility. We include two final model specifications for each trip purpose of each of the three SP experiments. Our first model is a linear Multinomial Logit (MNL) model, which we consider our base model. For the leisure trip purpose mode, we weight respondents on age and working status, so that the sample mirrors our target population. Our second model is a Mixed Logit (ML) model, estimated with 10,000 Halton draws. We use the software PythonBiogeme (Bierlaire, 2016) to perform the analysis.

The estimated ML models differ from their corresponding MNL models in three main aspects. First, they consider unobserved heterogeneity. We include a panel effect (also known as individual fixed effect), to account for the correlations between the different observations of the same individual; and we test for different nests in the transfer SP experiment, to account for correlations between some of the given alternatives. Second, we include interaction effects in the leisure trip purpose models, in order to account for taste variation between working and non-working individuals. Working status influences individuals' available time and money, both potential determinants of individuals' VOT and VOR. And third, in the ML models, we test for non-linear specifications of the different SP attributes (in particular, quadratic and squared root formulations), in order to test if the disutility associated with the unit increase of a specific attribute is different for different values of the attribute (and for the reliability attributes, to test whether individuals have risk-taking or risk-adverse attitudes instead of the risk-indifferent attitude that the linear parameter would imply (Li et al., 2012)). We compare the different formulations with the likelihood ratio test (p. 164–167 in Ben-Akiva and Lerman (1985)). If including a non-linear specification in an attribute leads to very slight increases of model fit, we keep its linear specification. Slight increases are not considered sufficient to justify the increased complexity in model interpretation.

In the utility function, we include two parameters to model the information originating from the five equally probable values. First, a parameter to account for the variability of the distribution. For this, we test three different specifications: (a) the standard deviation of the shown distribution; (b) the Reliability Buffer Time (RBT) (expressed as the difference between the 90th percentile and the scheduled time, following the results of Swierstra et al. (2017)[2]), and (c) the coefficient of variation (recommended as a measure of service reliability for the passenger in Abkowitz et al. (1978) but largely disregarded in recent studies). And second, a component to account for the inherent delay of the distribution with respect to the planned time. This is expressed in the utility function as the difference between the mean of the new distribution and the expected time announced in the scenario.

Based on the obtained parameters of the final models, the VOT and VOR are calculated. The VOT is the ratio of the marginal utility of time and the marginal utility of money, and the VOR is the ratio of the marginal utility of the reliability attribute and the marginal utility of money. These marginal utilities equal the estimated time and cost parameters when both attributes have a linear specification in the utility function.

2.3.2. Latent class choice models

In order to capture the potential heterogeneity in VOT and VOR of different individuals, we additionally perform a latent class choice model (LCCM) analysis. LCCMs determine the probability of each individual to belong to different classes (which correspond to different market segments). The different classes are identified statistically, rather than using predefined characteristics such as age or income, hence their name (Walker and Ben-Akiva, 2002). To decide on the most suitable number of distinct classes of the different models, we use the BIC (Bayesian Information Criterion) index. The BIC index takes into account both the model fit and the number of parameters in the model.

We adopt the ML model formulations as a starting model for our LCCM analysis, and we constrain all time and cost attribute values to be negative or zero. We weight respondents in the leisure trip purpose models on age and working status to mirror the sample population. In them, taste variation is analysed directly with the different classes. The waiting and in-vehicle stage SP experiments are modelled together in the latent class models, so as to account for the relation between the waiting and the in-vehicle

[2] Note that Swierstra et al. (2017) refer to the RBT as the 80th percentile minus the scheduled time, but they consider the fifth value of their equally probable values as such, and this should correspond to the 90th percentile (middle point between the 80th and the 100th percentile) and not to the 80th percentile.

preferences of the different individuals in the classes specification. Conversely, the transfer stage SP experiment needs to be modelled separately, since it is a labelled experiment. There, given the substantial weight and heterogeneity of the ASC in the utility of the alternatives, we add a random intercept coefficient in the transfer LCCM (modelled in LatentGOLD using a CFactor, see Vermunt and Magidson (2005)). This inclusion leads to final models with lower BIC and fewer classes.

In order to better assign individuals with different characteristics to the diverse classes (improve prediction of the segments), the model can be enriched with active covariates (individual characteristics such as age or income), in what is called the class membership function. However, this comes to the cost of obtaining classes that are also influenced by individual characteristics and not only by the SP attributes. Since our goal is to identify a portfolio of services that can be offered to all users (with different VOT and VOR) rather than to provide unique services to different population subgroups, we opt for not including individual characteristics actively in the membership function. Instead, we examine the characteristics of the individuals that belong to each of the obtained classes (passive covariates). In particular, we look at socioeconomic characteristics (gender, age, education level, working status, working hours, existence of children and urbanisation level), trip characteristics (trip length and trip frequency), and mobility-related characteristics (commuting transport mode, Uber usage, car availability and public transport usage). We perform the latent class analysis using the dedicated latent class software LatentGOLD (version 5.1) (Vermunt and Magidson, 2016).

3. Results

In this section we present the results of our analyses. Section 3.1 describes our data collection and sample. Section 3.2 presents the results of our choice model estimation for the three different SP experiments. Lastly, in Section 3.3, we apply latent class choice models to differentiate market segments.

3.1. Data collection and sample description

We performed an on-line pilot in April 2018 to test our questionnaire. Following the pilot, additional explanations were added to highlight the transition from the waiting stage to the in-vehicle stage SP experiments and hence to improve clarity. The final on-line questionnaire was then distributed in May 2018 (in Dutch). Survey participants were recruited from a panel designed for the longitudinal study of travel behaviour in the Netherlands, the Netherlands Mobility Panel (MobilitieitsPanel Nederland, MPN) (Hoogendoorn-Lanser et al., 2015). We target individuals aged 18 and older, owning a mobile phone and living in (sub)urban areas (known as (very) high urbanised areas according to the Dutch urbanity degree indicator (Centraal Bureau voor de Statistiek (CBS), 1992). Note that we do not restrict our target sample to individuals with specific mobility patterns or socioeconomic characteristics since there are still uncertainties regarding who will adopt pooled on-demand services in urban European contexts. Moreover, preferences of early adopters may differ from those of other potential users.

A total of 1006 individuals were considered valid respondents after data cleaning (93% of the obtained sample). Table 1 shows the socio-economic characteristics of the sample and the average Dutch values for (very) high urbanised areas and for the whole country. Our full sample satisfactorily represents the shares of our target population regarding urbanisation level, gender and age (middle

Table 1
Comparison between the sample and Dutch population for different socio-economic variables. Sources for the population data (Centraal Bureau voor de Statistiek (CBS), 2018a, 2018b, 2018c, 2018d).

Socio-economic variable	Category	Total sample (N = 1006)	Commuting trip purpose sample (N = 308)	Leisure trip purpose sample (N = 698)	Dutch (very) high urbanised areas	Dutch 2018 shares
Gender	Male	48.2%	49.4%	47.7%	48.9%	49.6%
	Female	51.8%	50.6%	52.3%	51.1%	50.4%
Age	18*–39	38.1%	53.2%	31.4%	38.1%	31.8%
	40–64	35.6%	46.5%	30.8%	42.0%	44.0%
	65 and above	26.3%	0.3%	37.8%	19.8%	24.2%
Education	Low	25.2%	16.2%	29.2%		31.5%
	Average	32.5%	35.4%	31.2%		37.8%
	High	42.0%	48.4%	39.3%		29.2%
	Unknown	0.2%	0.0%	0.3%		1.4%
Work status	Working	59.9%	100.0%	42.3%		50.9%
	No working	40.1%	0.0%	57.7%		49.1%
Household	1 person household	49.0%	46.8%	50.0%		38.2%
	> 1 person household	51.0%	53.2%	50.0%		61.8%
Urbanisation level	Very high urbanised (> 2500 inhab./km^2)	46.9%	47.7%	46.6%	48.2%	23.3%
	High urbanised (1500–2500 inhab./km^2)	53.1%	52.3%	53.4%	51.8%	25.1%

* 18–39 for the share sample, but 20–39 for the Dutch population 2018 values.

Table 2
MNL and ML model estimation for the waiting stage SP experiment (p-value: ≤0.01 ***, ≤0.05 **, ≤0.1*).

	Leisure trip purpose		Commuting trip purpose	
	MNL, weighted	ML, taste variation on working status	MNL	ML
	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)
Attribute				
Waiting time	−0.294 (−18.08)***	−0.332 (−6.57)***	−0.369 (−12.80)***	−0.366 (−4.42)***
Squared waiting time	N/A	−0.00605 (−3.34)***	N/A	−0.00824 (−2.91)***
Standard deviation	−0.125 (−4.97)***	−0.0950 (−2.53)***	−0.203 (−5.56)***	−0.319 (−6.57)***
Additional standard deviation parameter (working indiv. only)	N/A	−0.186 (−3.03)***	N/A	N/A
Displacement (i.e. Mean minus scheduled time)	−0.184 (−5.15)***	−0.283 (−6.79)***	−0.0960 (−1.97)**	−0.126 (−2.61)***
Cost	−1.61 (−18.32)***	−2.54 (−15.34)***	−1.51 (−10.61)***	−2.23 (−9.69)***
Sigma panel	N/A	−2.12 (−15.02)***	N/A	−2.05 (−9.65)***
Quality of fit statistics				
Initial log likelihood	−1930.609	−1935.267	−853.957	−853.957
Final log likelihood	−1684.758	−1503.961	−696.247	−622.587
Likelihood ratio test for the initial model	491.702	862.611	315.420	462.741
Rho-square	0.127	0.223	0.185	0.271

aged adults being slightly underrepresented and the elderly population slightly overrepresented). Education level, working status and household composition can only be compared to the average of the Dutch population. Our sample has a higher proportion of highly educated individuals, working respondents and single person households than the national average. We can expect our target population to also differ from the overall Dutch context in these directions. Therefore, we consider the sample to represent our target population adequately.

The large majority of the non-working individuals are retirees (62%), followed by individuals incapacitated to work (11%) and students (10%). Individuals who have a job as a secondary occupation (e.g., students with part-time jobs) are included in the working sample (since they can also relate to commuting trips). The majority of the working individuals (around 70%) were directed to the commuting trip (in order to have enough commuting entries). As a result, non-working individuals are overrepresented in the leisure trip sample. This fact is accounted in our performed models either by weighting the individuals to mirror the target sample or by including interaction effects that capture taste variation between working and non-working individuals (following the market segmentation procedure described in Ben-Akiva and Lerman (1985)).

3.2. Model estimation

Sections 3.2.1 and 3.2.2 present and analyse the choice models. Our results, in line with previous literature, indicate larger VOT for the commuting than for the leisure trip purpose. All our final models include the standard deviation as final variability component given that (a) neither the RBT nor the coefficient of variation perform better than the standard deviation in all models, and (b) the standard deviation allows for an easier comparison to values from other studies.

3.2.1. Results for the waiting and in-vehicle stage SP experiments

Tables 2 and 3 present the results of the waiting stage and the in-vehicle stage SP experiments respectively. The ML models clearly outperform their MNL equivalents (there are significant improvements in the rho-square). In all eight models, all attributes are negative (as expected) and significant at the 0.05 level (the vast majority also at the 0.01 level).

The inclusion of an additional quadratic coefficient for the waiting time attribute of the ML models clearly improves model fit. This means that individuals associate a higher per-minute waiting disutility to longer waiting times. Interaction effects regarding working situation are added to the ML specifications of the leisure purpose when these improve the model fit. Interestingly, we find that in the waiting stage, working individuals are more sensitive than non-working individuals towards variability, while, in the in-vehicle stage, working individuals are more sensitive than non-working individuals towards absolute increases of time (both regarding the expected time and the systematic unexpected delay).

Tables 4 and 5 present the VOT and VOR corresponding to the different models. The in-vehicle VOT values in this study for both commuting (10.80 €/h) and leisure purposes (7.88–9.94 €/h) are well in line with the Dutch car values presented in Kouwenhoven et al. (2014) (9.25 €/h commuting and 7.50 €/h other purposes, referring to the year 2010), and somewhat higher than those

Table 3
MNL and ML model estimation for the in-vehicle stage SP experiment (p-value: ≤0.01 ***, ≤0.05 **, ≤0.1*).

	Leisure trip purpose		Commuting trip purpose	
	MNL, weighted	ML, taste variation on working status	MNL	ML
	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)
Attribute				
In-vehicle time	−0.297 (−20.24)***	−0.420 (−15.55)***	−0.303 (−13.55)***	−0.495 (−11.86)***
Additional in-vehicle time parameter (working individuals only in leisure purpose)	N/A	−0.110 (−4.05)***	N/A	N/A
Standard deviation	−0.105 (−7.10)***	−0.168 (−7.41)***	−0.164 (−7.18)***	−0.264 (−6.71)***
Displacement (i.e. Mean minus scheduled time)	−0.118 (−5.42)***	−0.110 (−3.10)***	−0.120 (−3.57)***	−0.202 (−4.36)***
Additional displacement parameter (working individuals only in leisure trip purpose)	N/A	−0.125 (−2.20)**	N/A	N/A
Cost	−2.04 (−18.29)***	−3.20 (−17.54)***	−1.72 (−10.41)***	−2.75 (−10.16)***
Sigma panel	N/A	−2.03 (−14.03)***	N/A	−2.11 (−9.29)***
Quality of fit statistics				
Initial log likelihood	−1930.609	−1935.267	−853.957	−853.957
Final log likelihood	−1610.213	−1455.763	−681.316	−611.827
Likelihood ratio test for the initial model	640.792	959.009	345.282	484.261
Rho-square	0.166	0.248	0.202	0.284

Table 4
VOT and VOR of the ML specification for the waiting stage SP experiment in €/h.

	Leisure trip purpose	Commuting trip purpose
VOT waiting (5 min)	9.27	12.06
VOT waiting (15 min)	12.13	16.50
VOR waiting standard deviation (non-working individuals)	2.24	N/A
VOR waiting standard deviation (working individuals)	6.64	8.58
VOR waiting displacement	6.69	3.39

Table 5
VOT and VOR of the ML specification for the in-vehicle stage SP experiment in €/h.

	Leisure trip purpose	Commuting trip purpose
VOT in-vehicle (non-working individuals)	7.88	N/A
VOT in-vehicle (working individuals)	9.94	10.80
VOR in-vehicle standard deviation	3.15	5.76
VOR in-vehicle displacement (non-working individuals)	2.06	N/A
VOR in-vehicle displacement (working individuals)	4.41	4.41

presented in that study for public transport (6.00 €/h commuting and 7.75 €/h other purposes, referring to the year 2010). These differences may be due to user type effects being larger than mode type effects (as found in Wardman (2004)). That is, traditional public transport is usually chosen by individuals who have a lower willingness to pay (Zamparini and Reggiani, 2007), while, in our research, we target the average Dutch population. Moreover, our sample targets Dutch individuals living in urban areas, which have arguably higher values of time than the average Dutch population.

We find the VOT for the waiting stage to be in the range of 9.27–16.50 €/h, depending on trip purpose and total waiting time (increasing with longer waiting times due to the squared component in the utility function). The ratio between waiting and in-vehicle times varies between 1 and 1.5, depending on the waiting time. Traditionally, this ratio has been known to be 2 and above (Wardman, 2004). The waiting VOT has likely decreased thanks to the more widely available and accurate real-time information. Frei et al. (2017) even found the waiting VOT to be lower than the in-vehicle time for pooled on-demand services in the US context. As found in previous research (e.g., Ehreke et al. (2015), Kouwenhoven et al. (2014), Li et al. (2010)), we find waiting and in-vehicle VOT to be somewhat higher – around 30% higher in our research – for the commuting trip purpose than for the leisure trip purpose.

The VOR values regarding the standard deviation (3.15–5.76 €/h) of the in-vehicle time are also in line with those found on Kouwenhoven et al. (2014) for car and public transport in the Dutch context (3.25–4.75 €/h in 2010 terms). In both studies,

Table 6
MNL and ML model estimation for the transfer stage SP experiment (p-value: ≤ 0.01 ***, ≤ 0.05 **, ≤ 0.1*).

	Leisure trip purpose		Commuting trip purpose	
	MNL, weighted	ML, taste variation on working status	MNL	ML
	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)	Parameter (robust *t*-test)
Attribute				
Direct FLEXI – ASC	0	0	0	0
FLEXI-Bus – ASC	0.709(3.03) ***	0.644 (1.45)	1.26(3.55) ***	0.818 (1.25)
Bus-Bus – ASC	0.899(4.06) ***	0.647 (1.46)	1.60(4.45) ***	1.06(1.70) *
Direct FLEXI – In-vehicle time	−0.0957 (−13.59)***	−0.100 (−6.19)***	−0.104 (−9.70)***	−0.160 (−8.19)***
Additional Direct FLEXI – In-vehicle time parameter (working people only)	N/A	−0.0589 (−3.00)***	N/A	N/A
FLEXI-Bus – In-vehicle time leg 1	−0.103 (−11.19)***	N/A	−0.119 (−8.26)***	N/A
FLEXI-Bus – In-vehicle time leg 2	−0.106 (−11.48)***	N/A	−0.113 (−7.98)***	N/A
FLEXI-Bus – Total in-vehicle time	N/A	−0.129 (−10.79)***	N/A	−0.165 (−10.14)***
Additional FLEXI-Bus – Total in-vehicle time parameter (working people only)	N/A	−0.0648 (−3.85)***	N/A	N/A
Bus-Bus – In-vehicle time leg 1	−0.112 (−12.01)***	−0.143 (−11.08)***	−0.125 (−8.73)***	−0.184 (−10.07)***
Bus-Bus – In-vehicle time leg 2	−0.0933 (−10.26)***	−0.106 (−8.20)***	−0.130 (−8.95)***	−0.177 (−9.02)***
Additional Bus-Bus – Total in-vehicle time parameter (working people only)	N/A	−0.0650 (−3.89)***	N/A	N/A
FLEXI-Bus – Exp. waiting time	−0.0983 (−8.14)***	−0.118 (−6.43)***	−0.130 (−6.88)***	−0.176 (−8.35)***
Additional FLEXI-Bus – Exp. waiting time parameter (working people only)	N/A	−0.0503 (−2.01)**	N/A	N/A
Bus-Bus – Exp. waiting time	−0.0817 (−6.73)***	−0.0782 (−4.53)***	−0.128 (−6.74)***	−0.172 (−7.22)***
Additional Bus-Bus – Exp. waiting time parameter (working people only)	N/A	−0.0766 (−2.92)***	N/A	N/A
FLEXI-Bus – Standard deviation waiting time	−0.0587 (−1.69)*	−0.0555 (−1.28)	−0.0473 (−0.89)	−0.0768 (−1.15)
Bus-Bus – Standard deviation waiting time	−0.0626 (−1.87)*	−0.0113 (−0.33)	−0.0666 (−1.28)	−0.0352 (−0.62)
FLEXI-Bus – Displacement waiting time	−0.0612 (−1.46)	−0.0508 (−1.09)	−0.0861 (−1.37)	−0.103 (−1.39)
Bus-Bus – Displacement waiting time	−0.176 (−4.04)***	−0.160 (−3.05)***	−0.229 (−3.29)***	−0.211 (−2.51)***
Direct FLEXI – cost	−0.514 (−16.72)***	−0.996 (−17.15)***	−0.498 (−10.59)***	−0.877 (−10.79)***
FLEXI-Bus – cost	−0.797 (−16.75)***	−1.11 (−17.56)***	−0.868 (−11.69)***	−1.16 (−12.19)***
Bus-Bus – cost	−0.849 (−17.92)***	−1.16 (−19.71)***	−0.841 (−11.66)***	−1.13 (−11.92)***
Transfer error component	N/A	3.21(17.34) ***	N/A	2.73(12.11) ***
Quality of fit statistics				
Initial log likelihood	−4589.914	−4254.988	−2030.236	−1891.774
Final log likelihood	−3888.750	−3205.868	−1670.066	−1414.436
Likelihood ratio test for the init. model	1402.329	2098.241	720.340	954.675
Rho-square for the init. model	0.153	0.247	0.177	0.252

reliability ratios (RR), ratio between the VOR and the VOT, are, on average, around 0.5. This is a reassuring finding for the Dutch context, given that different studies worldwide reported a wide range of RR values, ranging from 0.1 to 2.5 (Carrion and Levinson, 2012). For the waiting stage, the range of VOR found is a bit more spread out, 2.24–8.58 €/h, with RR values ranging between 0.2 and 0.7 depending on the working situation and waiting time.

Table 7
VOT and VOR of the ML specification for the transfer stage SP experiment in €/h.

	Leisure trip purpose	Commuting trip purpose
VOT in-vehicle FLEXI (non-working individuals)	6.02	N/A
VOT in-vehicle FLEXI (working individuals)	9.57	10.95
VOT in-vehicle FLEXI-Bus (non-working individuals)	6.97	N/A
VOT in-vehicle FLEXI-Bus (working individuals)	10.48	8.53
VOT in-vehicle Bus-Bus (leg 1, non-working individuals)	7.40	N/A
VOT in-vehicle Bus-Bus (leg 2, non-working individuals)	5.48	N/A
VOT in-vehicle Bus-Bus (leg 1, working individuals)	10.76	9.77
VOT in-vehicle Bus-Bus (leg 2, working individuals)	8.84	9.40
VOT waiting transfer FLEXI-Bus (non-working individuals)	6.38	N/A
VOT waiting transfer FLEXI-Bus (working individuals)	9.10	9.10
VOT waiting transfer Bus-Bus (non-working individuals)	4.04	N/A
VOT waiting transfer Bus-Bus (working individuals)	8.01	9.13
VOR waiting transfer standard deviation FLEXI-Bus	3.00	3.97
VOR waiting transfer standard deviation Bus-Bus	0.58	1.87
VOR waiting transfer displacement FLEXI-Bus	2.75	5.33
VOR waiting transfer displacement Bus-Bus	8.28	11.20

To measure reliability, other than measuring the traditional variability, expressed with the standard deviation, we modelled the displacement value (calculated as the difference between the expected time and the mean of the presented distribution). Interestingly, the VOR value associated with this systematic delay tends to be similar to the corresponding VOR of the variability of the distribution (ranging between 2.06 €/h and 6.69 €/h), and always lower than the corresponding VOT. This suggests that the effect of a (small) unannounced yet systematic additional time in the utility function is lesser than that of the announced times. Even though this seems counterintuitive and contrary to findings from previous research (unexpected times are more heavily penalised by individuals than expected times (Currie and Wallis, 2008)), this suggests that a small additional unannounced increase in time may not be fully accounted for by individuals.

3.2.2. Results for the transfer stage SP experiment

Results of the MNL and ML models of both trip purposes for the transfer stage SP experiment are shown in Table 6. A nest structure is found between the two transfer alternatives. This is included in the ML model with a common transfer random error component. Also, interaction effects are added in the leisure trip purpose ML model to address the higher sensitivity of working individuals (with respect to non-working individuals) regarding the in-vehicle time and the expected waiting time.

The signs of all time and cost parameters are negative, as expected. The two reliability related parameters for the FLEXI-Bus alternative and the standard deviation parameter for the Bus-Bus one are, however, not significantly different from zero at the 95% confident level. In line with the results of Swierstra et al. (2017), finding significant parameters for the reliability attributes in more complex reliability SP experiments proved to be difficult. Since respondents were familiar with the reliability distribution representation from the waiting and in-vehicle experiments, and reliability values were significant in those, we hypothesise that the main explanation for non-significance is that individuals consider reliability attributes proportionally less important than other attributes in more complex decision processes (at least regarding the transfer stage).

Table 7 shows the VOT and VOR of the different parameters of the ML transfer stage SP experiment. In-vehicle VOT values range between 5.48 €/h and 10.95 €/h, depending on trip purpose and working status. Differences in the range of VOT values among the three alternatives are not pronounced. Contrary to what may be expected, all VOT for the waiting transfer times are lower than their corresponding in-vehicle times. Waiting transfer VOT range between 4.04 €/h and 9.13 €/h. We hypothesise that this finding may be due to the fact that part of the disutility from the transfer waiting stage arises from the uncertainty of the waiting time, which is modelled separately in this study. However, further research is needed to test this hypothesis and empirically underpin it.

As previously mentioned, three of the four reliability parameters are not significant at the 95% level. For the Bus-Bus alternative, the systematic delay (displacement) incurs in a much higher VOR than the variability (standard deviation). However, these values are much similar for the FLEXI-Bus alternative. This finding may suggest that individuals associate a higher disutility towards the systematic (i.e., certain) delay in services with fixed schedules, but that unreliability stemming from variability is less desired for more flexible (and more unknown) transport services.

3.3. Service differentiation for different market segments

Tables 8 and 9 depict the VOT and VOR of the different latent classes (model parameters of the latent class models are included in Appendix B). The latent class models strongly improve model fit of all models (rho-squared vales increase from 0.16 to 0.19 for the one latent class MNL models up to 0.58–0.61 for the latent class models). To get insights on our segment composition, we also describe in Tables 8 and 9 the different segments in terms of their passive covariates pertaining to socioeconomic, trip, and mobility-related characteristics.

The waiting plus in-vehicle stage LCCMs have four different classes for the leisure trip purpose and three different classes for the

Table 8
Results for the latent class estimation of the waiting and in-vehicle stage SP experiments. For the cluster profile identification, we highlight in bold the class with the highest share of each characteristic.

	Leisure trip purpose					Commuting trip purpose			
	1LC	4LC				1LC	3LC		
Main class characteristic		Time-cost balance. Waiting ~ In-vehicle	Time-cost balance. Waiting > In-vehicle	Time sensitive	Cost sensitive		Time-cost balance	Time sensitive	Cost sensitive
Class size	100%	40%	36%	13%	11%	100%	60%	26%	14%
VOT and VOR values (€/h)									
VOT waiting 5 min	9.75	4.37	19.49	23.65	0.99	12.71	9.86	120.84	2.87
VOT waiting 15 min	11.74	9.18	19.49	32.92	2.98	16.78	15.65	130.14	2.87
VOR sigma waiting	4.93	3.00	10.99	2.08	0.73	9.15	*9.34*	*10.93*	*5.06*
VOR displacement waiting	7.12	5.78	13.17	11.01	0.86	4.15	4.18	57.93	0.00
VOT in-vehicle	8.71	7.72	9.74	58.02	1.85	10.59	9.25	64.42	2.96
VOR sigma in-vehicle	3.08	3.00	2.83	38.19	0.55	5.72	*5.15*	*34.99*	*1.36*
VOR displacement in-vehicle	3.46	3.05	3.39	24.64	0.94	4.19	3.44	29.18	1.37
Cluster profile identification									
Socioeconomic characteristics									
Gender									
Male	49%	49%	48%	**51%**	44%	49%	47%	51%	**58%**
Female	51%	51%	52%	49%	**56%**	51%	**53%**	49%	42%
Age									
18–34	30%	**34%**	29%	23%	24%	41%	**47%**	30%	31%
35–49	21%	**23%**	21%	20%	14%	31%	29%	34%	**36%**
50–64	23%	19%	22%	**33%**	25%	28%	24%	**35%**	33%
65 or above	26%	23%	28%	24%	**36%**	0%	0%	**1%**	0%
Education level									
Low	23%	21%	22%	27%	**34%**	16%	14%	16%	**25%**
Middle	32%	32%	**35%**	25%	34%	35%	35%	**40%**	31%
High	44%	47%	43%	**48%**	32%	48%	**51%**	44%	44%
Working status									
Not working	40%	37%	41%	31%	**60%**				
Working	60%	63%	59%	**69%**	40%				
Working hours									
≤35 h a week						47%	**49%**	44%	44%
> 35 h a week						53%	51%	**56%**	**56%**
Has children aged ≥ 12									
No	88%	86%	**91%**	87%	88%	83%	82%	84%	**88%**
Yes	12%	**14%**	9%	13%	12%	17%	**18%**	16%	12%
Urbanisation level									
Highly urbanised areas	53%	54%	49%	52%	**59%**	52%	**54%**	49%	51%
Very highly urbanised areas	47%	46%	**51%**	48%	41%	48%	46%	**51%**	49%
Trip characteristics									
Trip length									
≤12 km trip	50%	48%	48%	54%	**61%**	49%	48%	49%	**55%**
> 12 km trip	50%	**52%**	**52%**	46%	39%	51%	**52%**	51%	45%
Trip frequency									
1–3 times a month	71%	72%	**73%**	66%	68%				
≥4 times a month	29%	28%	27%	**34%**	32%				
Mobility-related characteristics									
Commuting transport mode									
Car						38%	35%	**49%**	32%
Public transport						19%	18%	**23%**	20%
Active modes (bike or walk)						36%	**40%**	25%	**41%**
Other						6%	**7%**	3%	**7%**
Uber ever used									
No	90%	91%	86%	89%	**97%**	87%	87%	84%	**95%**

(*continued on next page*)

Table 8 (*continued*)

Main class characteristic	Leisure trip purpose					Commuting trip purpose			
	1LC	4LC				1LC	3LC		
		Time-cost balance. Waiting ~ In-vehicle	Time-cost balance. Waiting > In-vehicle	Time sensitive	Cost sensitive		Time-cost balance	Time sensitive	Cost sensitive
Yes	10%	9%	**14%**	11%	3%	13%	13%	**16%**	5%
Car availability									
No car in household	22%	**23%**	22%	20%	22%	26%	29%	17%	**31%**
Yes, but not always available	16%	17%	15%	10%	**20%**	24%	24%	**27%**	17%
Yes, and always available	62%	61%	63%	**70%**	57%	50%	47%	**56%**	51%
Public transport (PT) usage									
No PT used previous week	46%	49%	43%	**52%**	39%	47%	47%	45%	**52%**
PT used previous week	54%	51%	57%	48%	**61%**	53%	53%	**55%**	48%

commuting trip purpose (Table 8). The two largest classes of the leisure trip purpose (40% and 36% of individuals respectively) have the most balanced VOT values. They mainly differ in the waiting versus in-vehicle disutility. While the waiting VOT and VOR values are similar to those of the in-vehicle stage for individuals in the first class, the waiting VOT and VOR values are twice as high as their in-vehicle counterparts for individuals in the second class. In this aspect, the largest class of the commuting trip purpose (60%) can be seen as a middle point of these two balanced leisure trip purpose classes. For both trip purposes, the two smallest classes represent individuals who are either very time sensitive or very cost sensitive. In all classes, the VOR values tend to be somewhat lower than the VOT related values. This suggests a harmonious perception of the reliability ratio (ratio between VOR and VOT) across different market segments. The more balanced classes are characterised by a higher percentage of young individuals. Cost sensitive classes are formed primarily by low educated individuals, and, for the leisure trip purpose, individuals aged 65 years old or older and non-working individuals. For both time sensitive classes, the main common characteristics is the higher car availability and the higher share of individuals aged 50–64.

For the transfer stage LCCMs, we find three and two classes for the leisure and commuting trip purposes respectively (Table 9). As was the case for the waiting and in-vehicle LCCM, classes mainly differ from each other in their overall time/cost sensitivity degree. Within each class, we tend to find similar in-vehicle times for the three alternatives. Also, the in-vehicle VOT are in line with their corresponding transfer waiting VOT. Regarding the VOR values, they tend to be lower than their respective waiting VOT (except for the displacement VOR of the Bus-Bus alternative). Still, most of these values (as was the case for the one class models) are not significant. Regarding class composition, the more cost sensitive classes have larger percentages of non-working individuals and part-time workers (< 35 h a week), as expected. The opposite applies for the more time sensitive classes. As was the case for the waiting plus in-vehicle stage LCCMs, in the transfer LCCMs, the more balanced class (for the leisure trip purpose) has a larger percentage of young individuals.

4. Implications of the VOT and VOR analysis for the design of pooled on-demand services and further reliability considerations

The obtained parameters can be included in transport demand forecasting models such as macroscopic static assignment and agent-based simulation models to assess the possible modal shift towards new pooled on-demand services. Our findings can also help assess the impact of service provision design on users' choices, supporting service providers in developing their design strategies. For example, the ratio between the VOT for the waiting time and in-vehicle time can help service providers select the most suitable strategy to match new users and (re)route their vehicles. Also, the ratio between the VOT for commuting and leisure trips could help them set the price for different times of the day, in order to maximise profit and reduce the need to deploy a larger fleet during peak hour.

The latent class analyses presented in this paper help identify different market segments for on-demand services. Service differentiation may offer the more cost sensitive individuals the option to book a service with higher uncertainty and larger detours for a lower fare. Simultaneously, services with lower uncertainty or shorter waiting times can be offered for the more time sensitive individuals. Offering different services to cater for heterogeneity in preferences can additionally increase patronage and is the current strategy of ride-sourcing companies such as Uber. Service portfolios have also been suggested by Al-Ayyash et al. (2016) and Atasoy et al. (2015).

For the reliability specification, we opted for the mean-variance approach instead of the scheduling approach. Even if these two specifications have been proven to be similar, one underlying difference exists. In our approach, no specific (desired) arrival time is indicated. As a result, individuals are not presented with a situation of lateness. Instead, disutility stems exclusively from the variability/uncertainty of the distribution, which would allow them to choose slack time so as to minimise arriving late (even if this

Table 9
Results for the latent class estimation of the transfer stage SP experiment. For the cluster profile identification, we highlight in bold the class with the highest share of each characteristic.

	Leisure trip purpose				Commuting trip purpose		
Number of classes	1LC	3LC			1LC	2LC	
Main class characteristic		More time sensitive	Time-cost balance	Cost sensitive		More time sensitive	More cost sensitive
Class size	100%	52%	35%	13%	100%	55%	45%
VOT and VOR values (€/h)							
VOT direct flexi	11.18	14.45	3.30	0.00	12.54	16.78	5.49
VOT flexibus1	7.76	18.36	4.42	1.13	8.24	18.92	3.30
VOT flexibus2	8.01	14.62	8.64	0.00	7.84	15.79	4.67
VOT busbus1	7.89	16.59	6.26	2.18	8.94	24.63	5.81
VOT busbus2	6.60	12.95	6.27	1.80	9.30	29.48	5.16
VOT waiting transfer flexibus	7.40	11.21	8.67	0.00	8.99	13.32	7.00
VOT waiting transfer busbus	5.77	14.27	5.26	0.00	9.12	23.62	5.46
VOR sd flexibus	4.41	0.50	6.75	0.00	3.27	17.78	0.00
VOR sd busbus	4.42	0.00	0.00	0.00	4.75	5.80	3.01
VOR disp flexibus	4.61	3.09	3.59	7.43	5.95	11.59	2.11
VOR disp busbus	12.42	18.10	0.58	0.00	16.32	12.29	12.18
ASC monetisation (€/trip)							
ASC_FLEXI/(∂V/∂cost_FLEXI)	1.04	−0.14	−0.25	1.45	1.92	0.22	0.25
ASC_FLEXI-Bus/(∂V/∂cost_FLEXI-Bus)	−0.22	0.41	−0.07	0.48	−0.35	−0.26	0.16
ASC_Bus-Bus/(∂V/∂cost_Bus-Bus)	−0.43	−0.24	0.40	−1.32	−0.77	0.10	−0.41
Cluster profile identification							
Socioeconomic characteristics							
Gender							
Male	49%	**51%**	45%	**51%**	49%	47%	**52%**
Female	51%	49%	**55%**	49%	51%	**53%**	48%
Age							
18–34	30%	27%	**36%**	25%	41%	37%	**45%**
35–49	21%	**24%**	18%	18%	31%	**35%**	27%
50–64	23%	**24%**	22%	18%	28%	28%	28%
65 or above	26%	25%	24%	**39%**	0%	0%	**1%**
Education level							
Low	23%	22%	22%	**32%**	16%	16%	**17%**
Middle	32%	31%	**34%**	31%	35%	**39%**	32%
High	44%	**46%**	44%	37%	48%	46%	**52%**
Working status							
Not working	40%	36%	42%	**52%**			
Working	60%	**64%**	58%	48%			
Working hours							
≤35 h a week					47%	41%	**54%**
> 35 h a week					53%	**59%**	46%
Has children aged ≥ 12							
No	88%	85%	**93%**	86%	83%	82%	**85%**
Yes	12%	**15%**	7%	14%	17%	**18%**	15%
Urbanisation level							
Highly urbanised areas	53%	53%	**56%**	44%	52%	**53%**	51%
Very highly urbanised areas	47%	47%	44%	**56%**	48%	47%	**49%**
Trip characteristics							
Trip length							
≤12 km trip	50%	50%	**51%**	48%	49%	47%	**52%**
> 12 km trip	50%	50%	49%	**52%**	51%	**53%**	48%
Trip frequency							
1–3 times a month	71%	71%	70%	**72%**			
≥4 times a month	29%	29%	**30%**	28%			
Mobility-related characteristics							
Commuting transport mode							
Car					38%	**39%**	36%
Public transport					19%	19%	**20%**
Active modes (bike or walk)					36%	35%	**38%**
Other					6%	**7%**	6%

(*continued on next page*)

Table 9 (*continued*)

	Leisure trip purpose				Commuting trip purpose		
Number of classes	1LC	3LC			1LC	2LC	
Main class characteristic		More time sensitive	Time-cost balance	Cost sensitive		More time sensitive	More cost sensitive
Uber ever used							
No	90%	88%	91%	**92%**	87%	86%	**88%**
Yes	10%	**12%**	9%	8%	13%	**14%**	12%
Car availability							
No car in household	22%	19%	**26%**	24%	26%	24%	**29%**
Yes, but not always available	16%	**16%**	15%	14%	24%	23%	**24%**
Yes, and always available	62%	**65%**	59%	62%	50%	**53%**	47%
Public transport (PT) usage							
No PT used previous week	46%	**49%**	44%	42%	47%	**50%**	43%
PT used previous week	54%	51%	56%	**58%**	53%	50%	**57%**

comes at the cost of arriving too early). This means that, even if we believe that this approach better represents the real VOR of the individual, when presented with an unexpected situation, the VOR may be higher than found in this study to avoid lateness.

König et al. (2018) found that individuals with no ride-pooling experience attach more importance to travel time and fare and less importance to reliability related attributes than those with some ride-pooling experience. As a result, the ratio between the obtained values of reliability and values of times can be expected to somewhat increase with familiarity of these services. Note that service comfort and the sharing experience, not directly addressed in this research, may arguably partially explain this change. Having to share the vehicle with other passengers can be perceived by individuals as a source of inconvenience/discomfort that induces travel impedance. Alonso-González et al. (2020a) and Lavieri and Bhat (2019) estimate the willingness to share rides in pooled on-demand services.

Other than the discomfort associated with sharing per-se, some individuals may consider that the number of co-riders in the vehicle (with the accompanying comfort implication) is related to the level of unreliability (and in-vehicle time duration) of the alternatives shown in our scenarios [3]. In that case, the omission of the number of co-riders in the shown attributes could be linked to a certain omission variable bias, impacting the obtained parameter values. Specifically, this omission bias would have led to a negative bias in the measured related parameters (i.e., to an overestimation of the absolute values of our parameters) due to the (potentially assumed) positive correlation between the number of co-riders and the time and reliability attributes. We believe that the impact of such a potential bias, if any, is very limited and that it does not lead to an overestimation of our obtained parameters for the following two reasons: (a) Alonso-González et al. (2020a) shows that for the Dutch context the disutility associated with the number of co-riders is small, and that the decision choice is driven by the time-cost trade-offs instead (similar results were found in Lavieri and Bhat (2019) for the USA context); and (b) our obtained values and the reliability ratio obtained are in line with those found in a previous study for the Dutch context (Kouwenhoven et al., 2014).

A further aspect related to the flexibility/reliability of the offered on-demand services is uncertainty in the availability. Unguaranteed availability plays a key role in the probability of subscribing to shared mobility alternatives (Kim et al., 2017), and even a low probability of unavailability may be considered unacceptable by users that rely on using on-demand services on a daily basis (Fricker and Gast, 2016). In fact, vehicle unavailability was a decisive reason for individuals of the higher income groups to stop using the pooled on-demand service Kutsuplus (Helsinki, Finland) (Weckström et al., 2017). Additional research is needed to further understand how unavailability influences behaviour if this is a condition users may encounter.

5. Conclusions

We analysed the Value of Time (VOT) and Value of Reliability (VOR) of the different trip stages of pooled on-demand services, namely the waiting stage, the in-vehicle stage and the transfer stage (when combined with traditional public transport). To the best of our knowledge, this is the first study that analyses the time-reliability-cost trade-offs for all trip stages of new flexible transport modes, in particular for pooled on-demand services. This allows for VOT-VOR comparison, both within and between the different trip stages. We have differentiated between commuting and leisure trip purposes, and identified the taste variation between working and non-working individuals for the VOT and VOR values of their leisure trips. Additionally, to further classify the preference heterogeneity among different individuals, we identified different latent market segments.

Our research methodology can be divided in a design phase and an analysis phase. We first designed and executed a series of stated preference (SP) experiments. Our final sample was a representative sample of individuals living in (sub)urban areas in the Netherlands (N = 1006). We then analysed our data using mixed logit and latent class discrete choice models.

[3] Note however that, in reality, no direct relation exists between these attributes. A ride with one extra co-rider may result in higher unreliability and a longer detour time than another with four additional co-riders who are picked-up before us and are dropped-off afterwards.

Results show a higher willingness to pay for (the in-vehicle stage of) pooled on-demand services than known values for traditional public transport: 7.88–10.80 €/h depending on trip purpose and working status. Values of time for the waiting stage (both before the trip and during the transfer stage) are lower than values reported in the literature (around 1–1.5 and around 0.7–1 times in-vehicle VOT, respectively). Two reasons can account for this: first, currently available real time information reduces uncertainty and the related disutility; and second, we separately measure (and model) the waiting uncertainty, which is otherwise masked by the waiting VOT.

Values of reliability in the waiting and in-vehicle stages are found to be lower than their respective values of time, the ratio being around 0.5. This is in line with values found by Kouwenhoven et al. (2014) for car and public transport in the Dutch context. In the transfer stage, most reliability parameters proved insignificant (at the 95% level). The larger amount of attributes in the transfer stage SP experiment may have been the reason to the non-significance of these reliability parameters in our models. Further, the subsequent latent class analysis showed that the main difference between the classes of the different models pertains to the overall price-cost trade-offs rather than in different valuations of reliability in comparison to their corresponding values of time. This suggests a harmonious perception of the reliability ratio (ratio between VOR and VOT) across different market segments.

This research has presented individuals with outcomes of different reliability distributions. In reality, individuals are not directly confronted with this information when taking their transport decisions. Future research may delve into which attributes influence the perception of reliability or on how to better align the reliability characteristics of a service to individuals' perceptions of it. It would also be interesting to investigate if reliability for individual services is perceived differently than for pooled services, given that the latter are presumingly more susceptible to travel time variations. Additional research is also necessary regarding the extent to which VOR can be reduced by providing users with related real-time information. We also recommend future research to further investigate the transfer stage among experienced users, in order to analyse if a 'public transport + pooled on-demand' trip is perceived differently from a 'pooled on-demand + public transport' trip. Finally, future research could analyse whether and how users change their VOT and VOR valuations when gaining experience with pooled on-demand services.

Understanding not only the VOT of pooled on-demand services but also their VOR is of utmost importance, given the premise of their flexibility and the lack of clear reference values such as timetables. Results of this study can be used to forecast modal shift when introducing pooled on-demand services in urban contexts. Additionally, our findings can help in the design of such services by taking users' preferences into consideration.

CRediT authorship contribution statement

María J. Alonso-González: Conceptualization, Methodology, Investigation, Formal analysis, Visualization, Writing - original draft. **Niels van Oort:** Conceptualization, Supervision, Writing - review & editing. **Oded Cats:** Funding acquisition, Conceptualization, Supervision, Writing - review & editing, Project administration. **Sascha Hoogendoorn-Lanser:** Investigation. **Serge Hoogendoorn:** Funding acquisition, Conceptualization, Supervision, Writing - review & editing.

Acknowledgements

We thank the anonymous reviewers for their constructive feedback, which helped improve the quality of the paper. This work was funded by NWO (The Organisation for Scientific Research from the Netherlands), as part of the SCRIPTS (Smart Cities Responsive Intelligent Public Transport Services) research project.

Appendix A. Attribute levels of the stated preference experiments

See Tables A1–A3.

Table A1
Attribute levels of the waiting time SP experiment.

	Short version			Medium version		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Expected waiting time (Option 1 only) [min]	4	7	11	4	7	11
Extra expected waiting time (Option 2 only) [min]	3	5	8	3	5	8
Coefficient of variation (Cv) [–]	0.1	0.6	2	0.1	0.6	2
Displacement [min]	0	1	2	0	1	2
Cost (Option 2 only) [€]	2	4	6	3	5	7
Extra cost (Option 1 only) [€]	0.5	1.0	1.4	0.5	1.0	1.4

Table A2
Attribute levels of the in-vehicle time SP experiment.

	Short version			Medium version		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Expected in-vehicle time (Option 1 only) [min]	12	16	21	25	29	34
Extra expected in-vehicle time (Option 2 only) [min]	3	5	8	4	7	11
Coefficient of variation (Cv) [–]	0.1	0.6	2	0.1	0.6	2
Displacement [min]	0	1	3	0	2	4
Cost (Option 2 only) [€]	2	4	6	3	5	7
Extra cost (Option 1 only) [€]	0.3	0.7	1	0.5	0.9	1.3

Table A3
Attribute levels of the transfer SP experiment.

	Short version			Medium version		
	Level 1	Level 2	Level 3	Level 1	Level 2	Level 3
Total in-vehicle time (transfer alternatives) [min]	14	18	24	24	28	34
Share of the total in-vehicle time for the first vehicle (transfer alternatives) [–]	35%	50%	65%	35%	50%	65%
In-vehicle time (direct alternative) [min]	12	17	22	22	27	32
Expected waiting time transfer [min]	4	7	12	4	7	12
Coefficient of variation transfer (Cv) [–]	0.1	0.6	2	0.1	0.6	2
Displacement transfer [min]	0	1	2	0	1	2
Cost (transfer alternatives) [€]	1.5	2.3	3.3	2.5	3.3	4.3
Cost (direct alternative) [€]	3.5	5	6.5	4.5	6	7.5

Appendix B. Parameter values of the latent class choice models

See Tables B1–B4.

Table B1
Parameters of the leisure trip purpose LCCM for the waiting and in-vehicle stage SP experiments.

Attribute	Class 1 Parameter (z-value)	Class 2 Parameter (z-value)	Class 3 Parameter (z-value)	Class 4 Parameter (z-value)	Wald	p-Value	Mean	Std. Dev.
Waiting time	−0.1544 (−2.10)	−0.3908 (−7.91)	−1.5715 (−3.09)	0.0000 (−)	83.56	5.3e−18	−0.41	0.47
Squared waiting time	−0.0189 (−4.29)	0.0000 (−)	−0.0383 (−0.38)	−0.0055 (−1.43)	20.65	0.00012	−0.01	0.01
Standard deviation waiting	−0.2357 (−3.77)	−0.2204 (−3.99)	−0.1717 (−0.63)	−0.0405 (−0.46)	36.8	2.0e−7	−0.2	0.06
Mean minus scheduled waiting	−0.4539 (−4.32)	−0.2641 (−3.76)	−0.9095 (−1.42)	−0.0474 (−0.31)	40.69	3.1e−8	−0.4	0.23
Cost waiting time	−4.7149 (−7.24)	−1.2029 (−4.71)	−4.9577 (−0.90)	−3.3199 (−4.29)	87.3	4.9e−18	−3.34	1.66
Riding time	−0.5276 (−11.80)	−0.3564 (−11.08)	−2.3758 (−0.97)	−0.5533 (−1.51)	319.93	5.4e−68	−0.71	0.65
Standard deviation in-vehicle	−0.2047 (−5.12)	−0.1036 (−3.60)	−1.5638 (−0.66)	−0.1634 (−0.52)	58.5	6.0e−12	−0.34	0.48
Mean minus scheduled in-vehicle	−0.2085 (−4.04)	−0.1241 (−3.19)	−1.009 (−0.50)	−0.2809 (−0.61)	36.19	2.6e−7	−0.29	0.28
Cost in-vehicle time	−4.1 (−10.73)	−2.1948 (−9.12)	−2.4568 (−0.46)	−17.9007 (−1.97)	236.13	6.3e−50	−4.66	4.64
Model for Classes								
Intercept	0.6468 (−5.36)	0.5238 (−3.83)	−0.477 (−4.54)	−0.6936 (−6.26)	116.5	4.4e−25		

Table B2
Parameters of the commuting purpose LCCM for the waiting and in-vehicle stage SP experiments.

Attribute	Class 1 Parameter (z-value)	Class 2 Parameter (z-value)	Class 3 Parameter (z-value)	Wald	p-Value	Mean	Std.Dev.
Waiting time	−0.2604 (−2.46)	−0.6749 (−1.64)	−0.1585 (−1.68)	13.93	0.003	−0.36	0.19
Squared waiting time	−0.0108 (−2.87)	−0.0027 (−0.17)	0.0000 (−)	8.76	0.013	−0.01	0.00
Standard deviation waiting	−0.3487 (−5.71)	−0.0635 (−0.22)	−0.2793 (−2.22)	41.05	6.4E−09	−0.26	0.12
Mean minus scheduled waiting	−0.1562 (−2.13)	−0.3365 (−1.00)	0.0000 (−)	6.06	0.048	−0.18	0.11
Cost waiting time	−2.241 (−8.08)	−0.3485 (−0.44)	−3.3123 (−4.66)	92.21	7.3E−20	−1.90	0.98
Riding time	−0.4304 (−10.82)	−0.6176 (−2.97)	−0.2435 (−3.01)	157.03	8E−34	−0.45	0.12
Standard deviation in-vehicle	−0.2397 (−6.82)	−0.3354 (−2.33)	−0.1118 (−1.34)	59.61	7.1E−13	−0.25	0.07
Mean minus scheduled in-vehicle	−0.1603 (−3.27)	−0.2797 (−1.50)	−0.1122 (−0.93)	16.55	0.00087	−0.18	0.06
Cost in-vehicle time	−2.7928 (−8.44)	−0.5752 (−0.66)	−4.9279 (−4.43)	94.57	2.3E−20	−2.51	1.35
Model for Classes							
Intercept	0.7743 (−6.65)	−0.064 (−0.38)	−0.7103 (−4.14)	47.58	4.7E−11		

Table B3
Parameters of the leisure trip purpose LCCM for the transfer stage SP experiments.

Attribute	Class 1 Parameter (z-value)	Class 2 Parameter (z-value)	Class 3 Parameter (z-value)	Wald	p-Value	Mean	Std.Dev.
Direct FLEXI – ASC	0.1023 (0.23)	0.8821 (0.72)	−1.1708 (−1.33)	14.31	0.026	0.21	0.64
FLEXI-BUS – ASC	−0.3178 (−1.05)	0.1865 (0.28)	−0.6660 (−1.08)	–	–	−0.19	0.30
Bus-Bus – ASC	0.2155 (0.63)	−1.0685 (−1.33)	1.8368 (3.01)	–	–	−0.02	0.93
Direct FLEXI – In-vehicle time	−0.1793 (−9.40)	−0.1930 (−4.65)	0.0000 (−)	126.26	3.8e−28	−0.16	0.06
Direct FLEXI – cost	−0.7446 (−8.94)	−3.5108 (−5.04)	−0.8078 (−3.87)	116.88	3.6e−25	−1.72	1.31
FLEXI-Bus – In-vehicle time leg 1	−0.2385 (−10.59)	−0.1907 (−4.67)	−0.0263 (−0.57)	163.55	3.1e−35	−0.19	0.07
FLEXI-Bus – In-vehicle time leg 2	−0.1899 (−8.18)	−0.3730 (−8.16)	0.0000 (−)	122.69	2.3e−27	−0.23	0.12
FLEXI-Bus – Expected waiting time	−0.1456 (−5.40)	−0.3743 (−6.46)	0.0000 (−)	62.41	2.8e−14	−0.21	0.13
FLEXI-Bus – standard deviation waiting time	−0.0065 (−0.09)	−0.2916 (−1.87)	0.0000 (−)	3.91	0.14	−0.11	0.14
FLEXI-Bus – displacement waiting time	−0.0402 (−0.50)	−0.1550 (−0.99)	−0.1730 (−0.76)	2.46	0.48	−0.10	0.06
FLEXI-Bus – cost	−0.7796 (−7.24)	−2.5910 (−7.37)	−1.3978 (−4.48)	120.50	6.0e−26	−1.49	0.83
Bus-Bus – In-vehicle time leg 1	−0.2494 (−9.27)	−0.2776 (−6.40)	−0.0507 (−1.24)	166.88	6.0e−36	−0.23	0.07
Bus-Bus – In-vehicle time leg 2	−0.1947 (−7.06)	−0.2784 (−7.28)	−0.0418 (−1.12)	130.45	4.3e−28	−0.20	0.07
Bus-Bus – Expected waiting time	−0.2145 (−7.27)	−0.2336 (−5.38)	0.0000 (−)	94.97	2.4e−21	−0.19	0.08
Bus-Bus – standard deviation waiting time	0.0000 (−)	0.0000 (−)	0.0000 (−)	0.00	–	0.00	0.00
Bus-Bus – displacement waiting time	−0.2720 (−2.47)	−0.0257 (−0.16)	0.0000 (−)	6.68	0.04	−0.15	0.13
Bus-Bus – cost	−0.9018 (−7.18)	−2.6628 (−7.95)	−1.3961 (−6.53)	149.98	2.7e−32	−1.58	0.81

(*continued on next page*)

Table B3 (*continued*)

Attribute	Class 1 Parameter (z-value)	Class 2 Parameter (z-value)	Class 3 Parameter (z-value)	Wald	p-Value	Mean	Std.Dev.
CFactor1 Direct FLEXI	1.7804 (14.21)	1.7804 (14.21)	1.7804 (14.21)	230.86	7.4e−51	1.78	0.00
CFactor1 FLEXI-Bus	−0.8908 (−12.44)	−0.8908 (−12.44)	−0.8908 (−12.44)	–	–	−0.89	0.00
CFactor1 Bus-Bus	−0.8895 (−8.59)	−0.8895 (−8.59)	−0.8895 (−8.59)	–	–	−0.89	0.00
Model for Classes							
Intercept	0.5880 (4.99)	0.1924 (1.39)	−0.7804 (−5.34)	36.43	1.2e−8	–	–

Table B4
Parameters of the commuting trip purpose LCCM for the transfer stage SP experiments.

Attribute	Class 1 Parameter (z-value)	Class 2 Parameter (z-value)	Wald	p-Value	Mean	Std. Dev.
Direct FLEXI – ASC	−0.1408 (−0.26)	−0.4465 (−0.50)	2.63	0.62	−0.28	0.15
FLEXI-BUS – ASC	0.1970 (0.36)	−0.3247 (−0.50)	–	–	−0.04	0.26
Bus-Bus – ASC	−0.0562 (−0.11)	0.7712 (1.45)	–	–	0.32	0.41
Direct FLEXI – In-vehicle time	−0.1825 (−7.61)	−0.1618 (−3.39)	93.21	5.8e−21	−0.17	0.01
Direct FLEXI – cost	−0.6524 (−6.45)	−1.7699 (−4.52)	94.54	3.0e−21	−1.16	0.56
FLEXI-Bus – In-vehicle time leg 1	−0.2353 (−7.38)	−0.1101 (−3.12)	76.31	2.7e−17	−0.18	0.06
FLEXI-Bus – In-vehicle time leg 2	−0.1964 (−6.47)	−0.1559 (−4.49)	77.49	1.5e−17	−0.18	0.02
FLEXI-Bus – Expected waiting time	−0.1657 (4.62)	−0.2339 (−6.03)	64.45	1.0e−14	−0.20	0.03
FLEXI-Bus – standard deviation waiting time	−0.2211 (−1.96)	0.0000 (−)	3.83	0.05	−0.12	0.11
FLEXI-Bus – displacement waiting time	−0.1441 (−1.00)	−0.0706 (−0.43)	1.66	0.44	−0.11	0.04
FLEXI-Bus – cost	−0.7463 (−5.01)	−2.0047 (−9.14)	113.51	2.2e−25	−1.32	0.63
Bus-Bus – In-vehicle time leg 1	−0.2238 (−6.55)	−0.1803 (−4.97)	94.83	2.6e−21	−0.20	0.02
Bus-Bus – In-vehicle time leg 2	−0.2679 (−7.53)	−0.1599 (−4.88)	88.95	4.8e−20	−0.22	0.05
Bus-Bus – Expected waiting time	−0.2146 (−4.82)	−0.1693 (−4.00)	51.45	6.7e−12	−0.19	0.02
Bus-Bus – standard deviation waiting time	−0.0527 (−0.43)	−0.0933 (−0.88)	1.16	0.56	−0.07	0.02
Bus-Bus – displacement waiting time	−0.1117 (−0.74)	−0.3777 (−2.59)	8.35	0.01	−0.23	0.13
Bus-Bus – cost	−0.5452 (−3.40)	−1.8609 (−9.31)	98.88	3.4e−22	−1.14	0.66
CFactor1 Direct FLEXI	1.3290 (9.29)	1.3290 (9.29)	94.74	2.7e−21	1.33	0.00
CFactor1 FLEXI-Bus	−0.9833 (−8.20)	−0.9833 (−8.20)	–	–	−0.98	0.00
CFactor1 Bus-Bus	−0.3457 (−3.06)	−0.3457 (−3.06)	–	–	−0.35	0.00
Model for Classes						
Intercept	0.0907 (0.88)	−0.0907 (−0.88)	0.78	0.38	–	–

References

Abkowitz, M., Slavin, H., Waksman, R., Englisher, L.S., Wilson, N.H., 1978. Transit Servie Reliability (No. DOT-TSC-UMTA-78-18). United States.

Al-Ayyash, Z., Abou-Zeid, M., Kaysi, I., 2016. Modeling the demand for a shared-ride taxi service: An application to an organization-based context. Transp. Policy 48, 169–182. https://doi.org/10.1016/j.tranpol.2016.02.013.

Alemi, F., Circella, G., Mokhtarian, P., Handy, S., 2019. What drives the use of ridehailing in California? Ordered probit models of the usage frequency of Uber and Lyft. Transp. Res. Part C Emerg. Technol. 102, 233–248. https://doi.org/10.1016/j.trc.2018.12.016.

Alonso-González, M.J., Cats, O., Van Oort, N., Hoogendoorn-Lanser, S., Hoogendoorn, S., 2019. Willingness to share rides in on-demand services for different market segments. In: Presented at Thredbo 16 Conference, Singapore, 2019.

Alonso-Gonzalez, M.J., Cats, O., van Oort, N., Hoogendoorn-Lanser, S., Hoogendoorn, S., 2020. On the influence of cost and time on the willingness to share a ride: a scenario analysis. In: 99th Transportation Research Board (TRB) Annual Meeting, Washington DC.

Alonso-González, M.J., Hoogendoorn-Lanser, S., van Oort, N., Cats, O., Hoogendoorn, S., 2020b. Drivers and barriers in adopting Mobility as a Service (MaaS) – A latent class cluster analysis of attitudes. Transp. Res. Part A Policy Pract. 132, 378–401. https://doi.org/10.1016/j.tra.2019.11.022.

Arentze, T.A., Molin, E.J.E., 2013. Travelers' preferences in multimodal networks: Design and results of a comprehensive series of choice experiments. Transp. Res. Part A Policy Pract. 58, 15–28. https://doi.org/10.1016/j.tra.2013.10.005.

Asensio, J., Matas, A., 2008. Commuters' valuation of travel time variability. Transp. Res. Part E Logist. Transp. Rev. 44, 1074–1085. https://doi.org/10.1016/j.tre.2007.12.002.

Atasoy, B., Ikeda, T., Song, X., Ben-Akiva, M.E., 2015. The concept and impact analysis of a flexible mobility on demand system. Transp. Res. Part C Emerg. Technol. 56, 373–392. https://doi.org/10.1016/j.trc.2015.04.009.

Bansal, P., Liu, Y., Daziano, R., Samaranayake, S., 2019. Can Mobility-on-Demand services do better after discerning reliability preferences of riders? arXiv preprint arXiv:1904.07987.

Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. Transp. Res. Part E Logist. Transp. Rev. 37, 191–229. https://doi.org/10.1016/S1366-5545(00)00011-9.

Ben-Akiva, M., Lerman, S.R., 1985. Discrete Choice Analysis: Theory and Application to Travel Demand, vol. 9 MIT Press.

Bierlaire, M., 2016. PythonBiogeme : a short introduction. Rep. TRANSP-OR 160706, Ser. Biogeme. Transp. Mobil. Lab. Sch. Archit. Civ. Environ. Eng. Ec. Polytech. Fédérale Lausanne, Switzerland.

Black, I., Towriss, J., 1993. Demand Effects of Travel Time Reliability. Centre for Transport Studies, Cranfield Institute of Technology.

Brownstone, D., Small, K.A., 2005. Valuing time and reliability: Assessing the evidence from road pricing demonstrations. Transp. Res. Part A Policy Pract. 39, 279–293. https://doi.org/10.1016/j.tra.2004.11.001.

Carrion, C., Levinson, D., 2012. Value of travel time reliability: A review of current evidence. Transp. Res. Part A Policy Pract. 46, 720–741. https://doi.org/10.1016/j.tra.2012.01.003.

Centraal Bureau voor de Statistiek (CBS), 2018a. Bevolking; kerncijfers (in Dutch) [WWW Document]. URL https://opendata.cbs.nl/statline/#/CBS/nl/dataset/37296ned/table?ts=1541413888808 (accessed 11.5.18).

Centraal Bureau voor de Statistiek (CBS), 2018b. Bevolking; hoogstbehaald onderwijsniveau en onderwijsrichting (in Dutch) [WWW Document]. URL https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82816NED/table?ts=1541415076597 (accessed 11.5.18).

Centraal Bureau voor de Statistiek (CBS), 2018c. Arbeidsdeelname; kerncijfers (in Dutch) [WWW Document]. URL https://opendata.cbs.nl/statline/#/CBS/nl/dataset/82309NED/table?ts=1541426506205 (accessed 11.5.18).

Centraal Bureau voor de Statistiek (CBS), 2018d. Regionale kerncijfers Nederland; Bevolking; Stedelijkheidsklasse (in Dutch) [WWW Document]. URL https://opendata.cbs.nl/statline/#/CBS/nl/dataset/70072ned/table?ts=1541423939562 (accessed 11.5.18).

Centraal Bureau voor de Statistiek (CBS), 1992. Sociale kerncijfers - Nieuwe maatstaf van stedelijkheid. Mens en Maatsch (in Dutch).

(Michael) Chen, X., Zahiri, M., Zhang, S., 2017. Understanding ridesplitting behavior of on-demand ride services: An ensemble learning approach. Transp. Res. Part C Emerg. Technol. 76, 51–70. https://doi.org/10.1016/j.trc.2016.12.018.

Currie, G., Wallis, I., 2008. Effective ways to grow urban bus markets - a synthesis of evidence. J. Transp. Geogr. 16, 419–429. https://doi.org/10.1016/j.jtrangeo.2008.04.007.

De Jong, G., Tseng, Y., Kouwenhoven, M., Verhoef, E., Bates, J., 2007. The Value of Travel Time and Travel Time Reliability. The Netherlands Ministry of Transport, Public Works and Water Management.

Durán Hormazábal, E.A., 2016. Caracterización de la variabilidad del tiempo de viaje en la ciudad de Santiago. Universidad de Chile.

Ehreke, I., Hess, S., Weis, C., Axhausen, K.W., 2015. Reliability in the German value of time study. Transp. Res. Rec. J. Transp. Res. Board 2495, 14–22. https://doi.org/10.3141/2495-02.

El Zarwi, F., Vij, A., Walker, J.L., 2017. A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services. Transp. Res. Part C Emerg. Technol. 79, 207–223. https://doi.org/10.1016/j.trc.2017.03.004.

Fosgerau, M., Karlström, A., 2010. The value of reliability. Transp. Res. Part B Methodol. 44, 38–49. https://doi.org/10.1016/j.trb.2009.05.002.

Frei, C., Hyland, M., Mahmassani, H.S., 2017. Flexing service schedules: Assessing the potential for demand-adaptive hybrid transit via a stated preference approach. Transp. Res. Part C Emerg. Technol. 76, 71–89. https://doi.org/10.1016/j.trc.2016.12.017.

Fricker, C., Gast, N., 2016. Incentives and redistribution in homogeneous bike-sharing systems with stations of finite capacity. EURO J. Transp. Logist. 5, 261–291. https://doi.org/10.1007/s13676-014-0053-5.

Garcia-Martinez, A., Cascajo, R., Jara-Diaz, S.R., Chowdhury, S., Monzon, A., 2018. Transfer penalties in multimodal public transport networks. Transp. Res. Part A Policy Pract. 114, 52–66. https://doi.org/10.1016/j.tra.2018.01.016.

Hoogendoorn-Lanser, S., Schaap, N.T.W., Olde Kalter, M.J., 2015. The netherlands mobility panel: An innovative design approach for web-based longitudinal travel data collection. Transp. Res. Procedia 11, 311–329. https://doi.org/10.1016/j.trpro.2015.12.027.

ITF, 2017. Shared Mobility: Simulations for Helsinki, International Transport Forum Policy Papers. OECD Publishing, Paris https://doi.org/10.1787/3d340a2a-en.

ITF, 2016. Shared Mobility: Innovation for Liveable Cities, International Transport Forum Policy Papers. OECD Publishing, Paris https://doi.org/10.1787/5jlwvz8bd4mx-en.

Jin, X., Hossan, S., Asgari, H., 2015. Investigating the Value of Time and Value of Reliability for Managed Lanes Final Report (No. BDV29-977-12). Florida Department of Transportation, Miami.

Kieu, L.M., Bhaskar, A., Chung, E., 2014. Establishing definitions and modeling public transport travel time variability. In: Transportation Research Board 93th Annual Meeting.

Kim, J., Rasouli, S., Timmermans, H., 2017. Satisfaction and uncertainty in car-sharing decisions: An integration of hybrid choice and random regret-based models. Transp. Res. Part A Policy Pract. 95, 13–33. https://doi.org/10.1016/j.tra.2016.11.005.

König, A., Bonus, T., Grippenkoven, J., 2018. Analyzing urban residents' appraisal of ridepooling service attributes with conjoint analysis. Sustainability 10, 3711. https://doi.org/10.3390/su10103711.

Kouwenhoven, M., de Jong, G.C., Koster, P., van den Berg, V.A.C., Verhoef, E.T., Bates, J., Warffemius, P.M.J., 2014. New values of time and reliability in passenger transport in The Netherlands. Res. Transp. Econ. 47, 37–49. https://doi.org/10.1016/j.retrec.2014.09.017.

Lavieri, P.S., Bhat, C.R., 2019. Modeling individuals' willingness to share trips with strangers in an autonomous vehicle future. Transp. Res. Part A Policy Pract. 124, 242–261. https://doi.org/10.1016/j.tra.2019.03.009.

Li, W., Pu, Z., Li, Y., (Jeff) Ban, X., 2019. Characterization of ridesplitting based on observed data: a case study of Chengdu, China. Transp. Res. Part C Emerg. Technol. 100, 330–353. https://doi.org/10.1016/j.trc.2019.01.030.

Li, Z., Hensher, D.A., Rose, J.M., 2010. Willingness to pay for travel time reliability in passenger transport: A review and some new empirical evidence. Transp. Res. Part E Logist. Transp. Rev. 46, 384–403. https://doi.org/10.1016/j.tre.2009.12.005.

Li, Z., Tirachini, A., Hensher, D.A., 2012. Embedding risk attitudes in a scheduling model: application to the study of commuting departure time. Transp. Sci. 46, 170–188. https://doi.org/10.1287/trsc.1110.0389.

Liang, X., Correia, G.H., de, A., An, K., van Arem, B., 2020. Automated taxis' dial-a-ride problem with ride-sharing considering congestion-based dynamic travel times. Transp. Res. Part C Emerg. Technol. 112, 260–281. https://doi.org/10.1016/j.trc.2020.01.024.

Lu, W., Quadrifoglio, L., Petrelli, M., 2017. Reliability analysis of centralized versus decentralized zoning strategies for paratransit services. Transp. Res. Procedia 25, 4100–4113. https://doi.org/10.1016/j.trpro.2017.05.340.

Luo, S., (Marco) Nie, Y., 2019. Impact of ride-pooling on the nature of transit network design. Transp. Res. Part B Methodol. 129, 175–192. https://doi.org/10.1016/j.trb.2019.09.007.

Noland, R.B., Small, K.A., Koskenoja, P.M., Chu, X., 1998. Simulating travel reliability. Reg. Sci. Urban Econ. 28, 535–564.

Prashker, J.N., 1979. Direct analysis of the perceived importance of attributes of reliability of travel modes in urban travel. Transportation (Amst). 8, 329–346. https://doi.org/10.1007/BF00167987.

Ryley, T.J., Stanley, P.A., Enoch, M.P., Zanni, A.M., Quddus, M.A., 2014. Investigating the contribution of Demand Responsive Transport to a sustainable local public transport system. Res. Transp. Econ. 48, 364–372. https://doi.org/10.1016/j.retrec.2014.09.064.

Small, K.A., Noland, R., Chu, X., Lewis, D., 1999. Valuation of Travel Time Savings and Predictability in Congested Conditions for Highway User-Cost Estimation (Report 431). Transportation Research Board.

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., 2018. Enhancing urban mobility: Integrating ride-sharing and public transit. Comput. Oper. Res. 90, 12–21. https://doi.org/10.1016/j.cor.2017.08.016.

Stiglic, M., Agatz, N., Savelsbergh, M., Gradisar, M., 2016. Making dynamic ride-sharing work: The impact of driver and rider flexibility. Transp. Res. Part E Logist. Transp. Rev. 91, 190–207. https://doi.org/10.1016/j.tre.2016.04.010.

Swierstra, A.B., van Nes, R., Molin, E.J.E., 2017. Modelling travel time reliability in public transport. Eur. J. Transp. Infrastruct. Res. 263–278. https://doi.org/10.18757/ejtir.2017.17.2.3194.

Tachet, R., Sagarra, O., Santi, P., Resta, G., Szell, M., Strogatz, S.H., Ratti, C., 2017. Scaling law of urban ride sharing. Sci. Rep. 7, 42868. https://doi.org/10.1038/srep42868.

Tseng, Y.-Y., Verhoef, E., de Jong, G., Kouwenhoven, M., van der Hoorn, T., 2009. A pilot study into the perception of unreliability of travel times using in-depth interviews. J. Choice Model. 2, 8–28. https://doi.org/10.1016/S1755-5345(13)70002-2.

Tseng, Y., 2008. Valuation of Travel Time Reliability in Passenger Transport. Vrije Universiteit Amsterdam, The Netherlands.

Turnquist, M.A., Bowman, L.A., 1980. The effects of network structure on reliability of transit service. Transp. Res. Part B Methodol. 14, 79–86. https://doi.org/10.1016/0191-2615(80)90034-X.

Vermunt, J.K., Magidson, J., 2016. Technical Guide for Latent GOLD 5. 1: Basic, Advanced, and Syntax. Statistical Innovations Inc.

Vermunt, J.K., Magidson, J., 2005. Latent Gold Choice 4.0 User's Guide. Statistical Innovations Inc.

Vosooghi, R., Puchinger, J., Jankovic, M., Vouillon, A., 2019. Shared autonomous vehicle simulation and service design. Transp. Res. Part C Emerg. Technol. 107, 15–33. https://doi.org/10.1016/j.trc.2019.08.006.

Walker, J., Ben-Akiva, M., 2002. Generalized random utility model. Math. Soc. Sci. 43, 303–343. https://doi.org/10.1016/S0165-4896(02)00023-9.

Wardman, M., 2004. Public transport values of time. Transp. Policy 11, 363–377. https://doi.org/10.1016/j.tranpol.2004.05.001.

Wardman, M., Chintakayala, V.P.K., de Jong, G., 2016. Values of travel time in Europe: Review and meta-analysis. Transp. Res. Part A Policy Pract. 94, 93–111. https://doi.org/10.1016/j.tra.2016.08.019.

Weckström, C., Ullah, W., Mladenović, M.N., Nelson, J.D., Givoni, M., 2017. User perspectives on emerging mobility services: ex post analysis of kutsuplus pilot. ICoMaaS 2017 (1st International Conference on Mobility as a Service).

Yan, X., Levine, J., Zhao, X., 2019. Integrating ridesourcing services with public transit: An evaluation of traveler responses combining revealed and stated preference data. Transp. Res. Part C Emerg. Technol. 105, 683–696. https://doi.org/10.1016/j.trc.2018.07.029.

Zamparini, L., Reggiani, A., 2007. Meta-analysis and the value of travel time savings: A transatlantic perspective in passenger transport. Networks Spat. Econ. 7, 377–396. https://doi.org/10.1007/s11067-007-9028-5.

Zheng, Y., Li, W., Qiu, F., Wei, H., 2019. The benefits of introducing meeting points into flex-route transit services. Transp. Res. Part C Emerg. Technol. 106, 98–112. https://doi.org/10.1016/j.trc.2019.07.012.