Quantifying travellers' evaluation of waiting time uncertainty in public transport networks

Shelat, Sanmay; Cats, Oded; van Lint, J. W.C.

DOI
10.1016/j.tbs.2021.07.009

Publication date
2021

Document Version
Final published version

Published in
Travel Behaviour and Society

Travel Behaviour and Society 25 (2021) 209–222

Contents lists available at ScienceDirect

Travel Behaviour and Society

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

Quantifying travellers' evaluation of waiting time uncertainty in public transport networks

Sanmay Shelat *, Oded Cats, J.W.C. van Lint

Department of Transport and Planning, Delft University of Technology, Stevinweg 1, 2628CN Delft, the Netherlands

ARTICLE INFO

Keywords:
Public transportation
Waiting time
Decision under uncertainty
Certainty equivalent
Stated preference experiment

ABSTRACT

Although waiting times are inherently uncertain in public transport networks, previous research has primarily studied route choice behaviour under risk. Since understanding such behaviour is important to correctly model network flows and gain better understanding of traveller satisfaction, we propose a method to assess travellers' route choice behaviour under natural ambiguity. Specifically, we devise a realistic route choice situation whereby travellers' attitudes and perceptions towards waiting time uncertainty as well as the effects of situational contexts thereon can be quantified in terms of a certainty equivalent. The choice situation is contextualised in a stated preferences experiment to analyse the premium travellers in the Dutch railways are willing to pay for certainty in waiting time. Results indicate a significantly improved model fit and predictive value when accounting for waiting time uncertainty with travellers, on average, willing to trade-off more than 7 minutes of in-vehicle time to have certainty in their waiting time. Minor non-linear effects of elapsed waiting time and anticipated delays on the value of certainty are also found and heterogeneity analysis indicates that younger travellers tend to seek more certainty. The proposed method provides snapshots of travellers' behaviour under uncertainties in a real-world public transport system and can as such be used to improve transportation models, provide more tailored travel advice, and be used to test the efficacy of different policies.

1. Introduction

Analysing route choice behaviour in public transport networks is important for both supply and demand management. It is an essential input for determining network flows which authorities use to manage service levels and prioritize relevant investments. Furthermore, knowing how such decisions are made, travellers can be nudged into choices that are more optimal for them and the system, and can be suggested options that are likely to result in higher traveller satisfaction. Route choice decisions are largely governed by travellers' attitudes and attributes of the public transport system. Increasingly, route choice models have incorporated service attributes beyond travel time components, including, for instance, graphical distortions of transit network representation (Raveau et al., 2011), transfer station layouts (Guo and Wilson, 2011), and on-board crowding (Yap et al., 2018). This study contributes to this line of research by assessing travellers evaluations of waiting time uncertainty above and beyond nominal values. Given the various sources of stochasticity in public transport networks, its travel time attributes are inherently uncertain; however, as we will show in our literature review, this has not been accounted for properly in existing studies. In order to describe and explain route choice decisions more completely, we develop in this study a route choice model that explicitly accounts for travellers' behaviour under waiting time uncertainties in public transport networks.

First, we clarify what we mean by 'uncertainty'. The Knightian (Knight, 1921) classification of uncertainty is based on whether, for a set of possibly infinite events, objective probabilities exist or not. Decisions under the former regime are said to be made under 'risk' while those under the latter are under 'ambiguity' or 'uncertainty'. Objective probabilities exist either when they are made available to decision-makers (and are trusted by them), there is a consensus amongst decision-makers regarding them, or when they are integrated within the decision problem itself. However, these assumptions are quite stringent and are seldom fulfilled in the real world. Outside of artificial games such as casinos and lotteries, real-world events occur under ambiguity where decisions are based on personal beliefs (Machina and Siniscalchi, 2014). Travellers in public transport networks also do not have access to such objective probabilities for the different attributes involved and

* Corresponding author.
E-mail addresses: s.shelat@tudelft.nl (S. Shelat), o.cats@tudelft.nl (O. Cats), j.w.c.vanlint@tudelft.nl (J.W.C. van Lint).

https://doi.org/10.1016/j.tbs.2021.07.009
Received 5 October 2020; Received in revised form 20 June 2021; Accepted 23 July 2021
Available online 10 August 2021

make their decisions under uncertainty. Even if information is provided on the various aspects of travel time, it is distorted by travellers' beliefs arising from personal characteristics, habits, experiences, and contemporary contextual variables.

Next, we delineate why uncertainty in waiting time is of special interest. Similar to other industries in the service sector, in public transport systems too, waiting times have been found to play a crucial role in consumers' decision-making and satisfaction (Abenoza et al., 2018). While the cost of waiting can usually be objectively calculated in the manufacturing industry, to describe its manifestation in the service sector, Maister (1985) quotes the copywriters of a parcel delivery service: 'waiting is frustrating, demoralizing, agonizing, aggravating, [and] annoying…'. Arguably, these feelings arise from the uncertainty that is often inherently involved with waiting time as well as the context in which it is experienced. In the service industry, apart from the objective magnitude, the perception of waiting time is critical for customer satisfaction (Maister, 1985) and any disparity between objective and subjective expectations of waiting times may lead to sub-optimal decision-making. Therefore, it is vital to analyse travellers' attitudes and perceptions regarding waiting time uncertainty.

The impact of waiting time on route choice behaviour has been typically studied using either expected values or objective probabilities of risk. Both of these approaches fail to account for travellers' beliefs regarding uncertainties associated with waiting time. To that end, the present study proposes a method to assess travellers' route choice behaviour under natural ambiguity without using objective probabilities or assuming specific learning behaviour — important drawbacks in existing studies. Specifically, a realistic route choice situation is proposed whereby travellers' beliefs towards waiting time uncertainty can be quantified in terms of a certainty equivalent. For any gamble, its certainty equivalent is a risk-less value such that the decision maker is indifferent between receiving this risk-less value and playing the gamble. For example, if a decision-maker is indifferent between (a) gambling on a fair coin toss winning $0 on heads and $5 on tails; and (b) winning $2 for sure, then $2 is the certainty equivalent of the gamble offered in (a) for this decision-maker. The certainty equivalent in this case indicates the decision-maker's attitude towards risk — risk-aversion in this case. When the gamble is ambiguous/uncertain, such as wining $0 if a train departs within 1 minute of its scheduled time, $5 otherwise, in addition to attitude towards the uncertainty, the certainty equivalent will also indicate how uncertain an outcome is felt to be by the decision-maker. For instance, if the certainty equivalent for the above was $1, we can infer that the decision-maker feels that the train is more likely to be on time than to be late. The identified choice situation also permits the estimation of the effects context variables have on the certainty equivalent for waiting time. The conditions required for the proposed situation are simple enough that it is fairly common for it to take place structurally (i.e., because of service or network design) in real-world public transport networks; also implying that most travellers will be able to identify with the situation. As a case study, this choice situation is contextualized and used in a stated preferences experiment aimed at understanding route choice behaviour in the Dutch railways.

In the next section, studies on travel behaviour under uncertainty are reviewed, classifying them on the type of uncertainty observed. Section 3 lays out a theoretical framework of choice behaviour under uncertainty. Next, the proposed choice situation is presented in section 4. This is followed by the description of the design and presentation, sample of the stated preferences experiment, and choice analyses in section 5. Section 6 presents and discusses the results of the choice analyses, and finally, the main contributions, outcomes, and limitations are outlined in section 7.

2. Literature review

In this section, we briefly review the large body of literature dedicated to analysing the effect of variability in different aspects of travel time on travellers' decisions. While these studies may fulfil their own objectives, here we analyse drawbacks specifically with respect to observing and analysing behaviour under uncertainty. Decisions have been typically observed under risk, simulated uncertainty, or natural ambiguity. Research approaches — stated preference experiments, laboratory experiments, or analysis of actual trips — have been closely associated with the type of uncertainty under which decision-making has been observed and is accounted for in the analysis.

As discussed above, in the real-world, decisions are made under ambiguity — in the absence of objective probabilities. In contrast, however, travel behaviour under uncertainty is most commonly studied by presenting hypothetical route alternatives with objective distributions of travel times. Furthermore, since such probabilities are usually not available to travellers, conveying objective probabilities is notoriously difficult (Bates et al., 2001; Carrion and Levinson, 2012). This is exclusively the type of uncertainty observed in stated preferences (e.g., Small et al. (1999); Swierstra et al. (2017); Tilahun and Levinson (2010)).

A few laboratory experiments have observed choice in traffic networks under partial uncertainty by offering different levels and accuracies of information to respondents within the context created in a 'travel simulator' (e.g., Ben-Elia et al. (2013); Ben-Elia and Shiftan (2010); Bogers et al. (2005); Bogers et al. (2006); Ramos et al. (2011)). Unlike stated preference questionnaires, respondents do not make single-shot decisions but are required to consider a number of choice situations with or without feedback. These experiments typically focus on analysing learning mechanisms (e.g., Avineri and Prashker (2005; 2006)) and the effects of different uncertainty levels (e.g., Ben-Elia et al. (2008)). In an interesting setup, Kemel and Paraschiv (2013) observe choices under artificial ambiguity using Ellsberg's urns (Ellsberg, 1961). Artificial ambiguity is typically created using an unknown mix of differently coloured balls in an urn. This approach is often used in ambiguity studies to control for likelihood beliefs in a laboratory setting. Since participants do not have any information regarding the proportions of different colours, they cannot form any beliefs about this. Kemel and Paraschiv (2013) as well as a number of authors (as summarized in Baillon et al. (2018)) note that the external validity of such studies could be improved by using natural (real-world) events (for example, stock market prices or actual departure times of public transport vehicles).

Studies observing behaviour under natural ambiguity are sparse and typically use revealed preferences from real-world observations in car traffic. While revealed preferences offer high behavioural validity, unlike stated preferences and laboratory experiments, there is little experimental control. A series of papers observed route choice behaviour in two road-pricing demonstrations in California, involving a free but congested route and a (time-varying) tolled route with low congestion levels (and hence an almost certain travel time), just before the beginning of this millennium (see Brownstone and Small (2005) for an overview). While these fairly unique opportunities offered reasonable choice experiment settings, the studies faced significant issues in data collection and preparation.

While in reality travellers do not have access to objective probabilities, studies using observations of choices under risk face an additional problem that is related to the difficulties in conveying probabilities. Empirical findings suggest that for choices under risk, people do not fully distinguish between different levels of probability (Wakker, 2010, section 7.1) as is assumed in the commonly adopted expected utility regime. In recent years, however, a few studies (see Li and Hensher, 2011; Li and Hensher, 2019); Rasouli and Timmermans (2014) for a review) have used rank-dependent utility (Quiggin, 1982) and cumulative prospect theory (Tversky and Kahneman, 1981) which apply subjective probability weighting that can account for such likelihood insensitivity. However, only a few studies estimate the functional form and parameters of the probability weighting function (Li and Hensher, 2011; Li and Hensher, 2019). Finally, an important issue in studies using

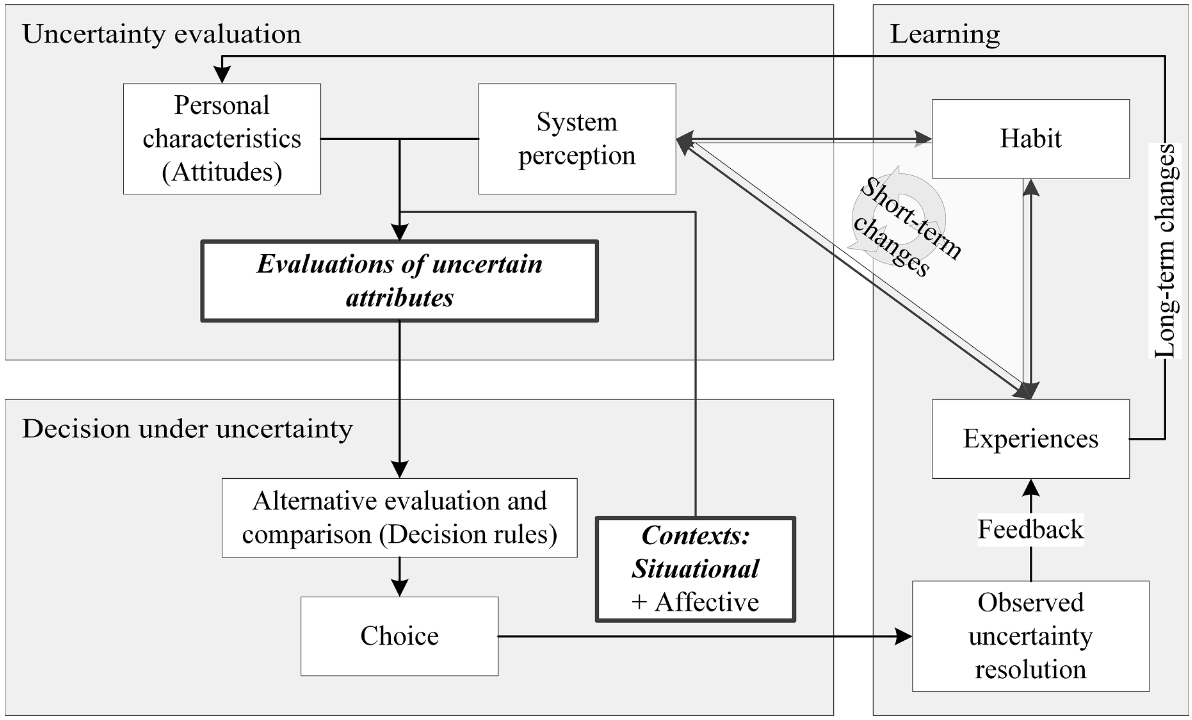

Fig. 1. Theoretical framework of decision-making under uncertainty. Components in bold-italics are the focus of this study.

revealed preferences data is that, although decisions are made under ambiguity, analysis has been commonly carried out using objective probabilities (Carrion and Levinson, 2012) under the assumption that these probabilities are known to the traveller through experience (Ghosh, 2001; Lam and Small, 2001; Small et al., 2005).

An alternative to analysing travellers' attitudes and perceptions regarding uncertainty through choice observations could be to directly ask them about their perceived and expected travel times. The idea is that reported travel time values will incorporate any uncertainties experienced by travellers. This approach has been implemented in a number of studies researching the effect of various aspects of travelling, such as real-time information provision (Dziekan and Vermeulen, 2006; Watkins et al., 2011), on perceived waiting times (see Meng et al. (2018) for a brief overview). This approach is useful to assess a posteriori travel satisfaction. However, Peer et al. (2014) find that reported values do not accurately describe those used for decision making, suggesting that discrepancies between objective and reported values may arise from a number of reasons that do not actually affect travellers' behaviour. Furthermore, even incentivising travellers to report their true beliefs through scoring rules (see Winkler et al. (1996) for an overview) does not seem to reduce discrepancies or improve interpretability (Dixit et al., 2019).

From this review, several drawbacks in existing studies can be identified with respect to analysing route choice behaviour under uncertainty in public transport networks. In most studies, choices observed have been made and/or analysed using objective probability distributions, which are not only missing in the real-world but are also distorted by travellers' prior beliefs that arise from a number of factors such as previous experiences, habits, and contexts, leading to possibly biased outcomes. Studies where choices observed have been made under uncertainty — as in revealed preferences and laboratory studies — were only performed in the context of car traffic networks leaving an important gap for studying behaviour in public transport networks.

To overcome these drawbacks, we identify a choice situation wherein travellers' assessment of waiting time uncertainty in public transport networks can be explicitly quantified directly from observed choices; without external psychometric measurements or collection of reported values. First, however, we present a generic theoretical framework of travel behaviour under uncertainty that outlines the various factors affecting choice and their interactions with the aim of placing the current study in context.

3. Theoretical framework

In order to describe decision-making under uncertainty, we divide the process into three main parts: (i) uncertainty evaluation, (ii) decision-making, and (iii) learning (Fig. 1). Evaluation of uncertainty in attributes is a result of the decision-maker's attitudes (e.g., risk aversion) as well as their perception of the system (e.g., feeling that the system is unreliable). Both attitudes and system perceptions, and therefore the uncertainty evaluation, can be affected by the context, which can be situational or affective. The former affects the environment in which the decision is made while the latter relates to the moods and feelings of the decision-maker at the time. These evaluations are then used to assess and compare alternatives leading to a choice. After making a choice, the resolution of some or all of the uncertainty may be observed by the decision-maker, which feeds back to their experience memory. Previous experiences can lead to longer-lasting changes to their attitudes or shorter-term changes to their system perception. This can take place either over several decision outcomes or after a few extreme ones. Experiences also lead to habit formations and the regularity with which the same choices are made can affect perceptions (e.g., regular cyclists might perceive cycling to be safer than occasional cyclists). Finally, the effect the decision-maker's system perception has on their experienced utility (travel satisfaction) and habits closes this short-term learning loop. Note that we do not propose this framework as a validated scheme but use it to highlight and conceptually place the aspects considered in this study.

In this study, we focus on the evaluation of uncertainty which is dependent on personal characteristics developed over a long period of time and system perceptions that are updated more frequently, as well as the effects contexts (we only study situational contexts and not affective ones) have on them. We assume decisions are made under the random utility maximization paradigm. Furthermore, the focus is on capturing

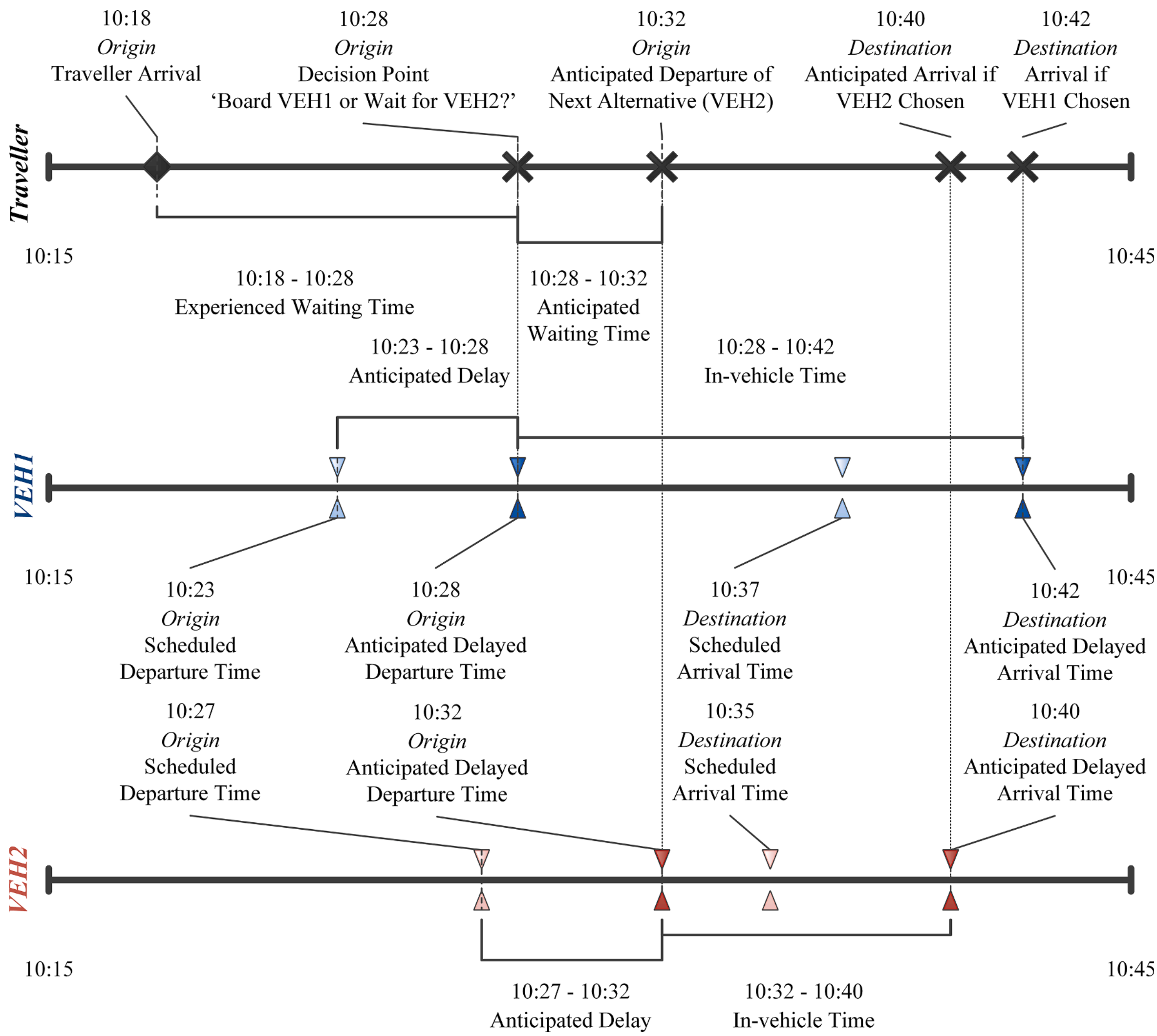

Fig. 2. Choice situation presented in a timeline format.

snapshots of travellers attitudes and perceptions; therefore, we do not study the feedback and learning mechanism involved in uncertainty evaluation.

3.1. *Personal characteristics + system perceptions = uncertainty evaluation*

Theoretically, personal characteristics and (subjective) perceptions of risk are distinguished to study which of these are the driving forces behind behaviour under uncertainty (Weber and Milliman, 1997). Anticipation of regret and attitudes towards risk and uncertainty are amongst the most influential personal characteristics for decisions under uncertainty. These personal characteristics are developed over a long period of time and are not susceptible to frequent changes. They have been quantified in literature in a number of ways from Likert scales to various mathematical formalizations in decision models including expected utility, cumulative prospect theory, and regret theory. Unlike attitudes, subjective perceptions are updated frequently based on habits and experiences (gaps between expectations and outcomes). A number of models (e.g., Bayesian updating, weighted average learning) have been proposed for the learning mechanism through which these three aspects — perceptions, habits, and experiences — interact with one another.

Practically, however, it is difficult to disentangle the effects of personal characteristics and perceptions in observed behaviour. For instance: does a person buy theft insurance because she feels theft is likely to occur or because she is generally risk averse in these matters? In single attribute experiments, outcome valuation and subjective probabilities have been successfully disentangled, for instance using the trade-off method (Wakker, 2010) but it is not obvious how this would be done in multi-attribute decisions such as route choice. When using non-expected utility models for decisions under natural ambiguity, only recently have studies explicitly measured ambiguity aversion whilst controlling for likelihood beliefs (Baillon et al., 2018). Indeed some (Nau, 2001) have argued that the separation of preferences arising from personal characteristics and beliefs is neither possible nor required for decision analysis or economic modelling. Therefore, for this study we consider travellers' uncertainty evaluation as a whole which, in fact, is formed by their personal characteristics and perceptions. We will continue using this term in the remainder of the paper. Note that, we use 'uncertainty evaluation' as an all-encompassing term; referring not only to how likely a decision-maker feels that a particular event will occur but also the impact (or value) thereof.

3.2. *Situational contexts*

Contemporary contexts affect how an attribute (e.g., waiting time) is experienced. For waiting time, Maister (1985) makes a number of propositions that define which contexts make waiting seem longer or shorter than reality; for instance, occupied time feels longer than unoccupied time or that unexplained waits are longer than explained ones. Jones (1996) reviews these propositions in terms of the degree to which service managers can control the related contexts and their impacts on customers. Previous studies in transportation have also explored the differences in value of travel time for different contexts such as free-flow traffic, stop-and-go traffic, and on-ramp delays (Hensher, 2001; Levinson et al., 2004). Ongoing experience is important because it will be taken into account by customers when anticipating the value of uncertain attributes in the upcoming future.

With increasingly prevalent real-time information, seemingly irrelevant information may also affect travellers' evaluation of uncertainty. For instance, delay predictions along the corridor of a traveller or even in other parts of a transportation network might cause increased anxiety and a breakdown of trust in the system, leading to choices that indicate a disproportionately higher degree of pessimism or risk/ambiguity aversion.

As a contextual variable, the amount of waiting time already experienced by the time of decision may have two opposite effects of varying magnitude. On the one hand, greater experienced waiting time translates to increasing stress and frustration (Osuna, 1985), on the other, there may be a sunk-cost effect (Thaler, 1980) wherein having waited for some time is in itself an impetus to wait some more. In an explicit study on the sunk-cost effect for time (rather than money which most authors examine), Soman (2001) finds that because people do not have the ability to account for time in the way they do for money, the effect is not found. However, he does not consider travel time in transportation choices where, often, one time component is traded-off with another in the same trip which could make it easier for people to open and keep mental accounts of time.

4. Choice situation

In this section, we present the choice situation that will be analysed to obtain travellers' evaluations of uncertainty in waiting time and the effects of contexts thereon. Amongst the sequence of choices faced by a traveller, we look at the decision of whether to board a particular vehicle in the following situation.

Consider a traveller who arrives at a public transport stop. From here, either of the next two vehicles can take her to her destination. Both of these vehicles are identical in every way except for their departure and arrival times at the origin and destination stations, respectively. Furthermore, both of these vehicles will take her directly, without any transfers, to the destination station. As is prevalent in many transit systems worldwide, real-time information regarding anticipated departure times and delays is displayed alongside scheduled departure times. Moreover, she is assumed to know the time both vehicles will take to reach her destination station (either from experience or a travel planner). When the first vehicle (VEH1) arrives, she must make a decision, based on the information available to her and her own uncertainty evaluation for the network, whether to board it or to wait for the next one (VEH2). Fig. 2 shows the proposed choice situation in a timeline format.

Although the vehicles are identical, the options available to the traveller (unlike route alternatives in most choice situations) are not unlabelled, that is, they have alternative-specific properties — in fact, the traveller is comparing a *certain* (as in risk-less) option against an ambiguous one. The vehicle that has already arrived has a *certain* waiting time which is almost zero due to the, usually, negligible difference between boarding, doors closing and departure. Although the anticipated waiting time for the next vehicle is displayed (either directly or as anticipated departure time of the next vehicle), it is ambiguous for the traveller since no concrete probabilities regarding its accuracy are supplied. Rather, she will draw from her own evaluation of this natural source of ambiguity and make a decision.

Aside from trading-off the difference in in-vehicle times against the anticipated waiting time, the traveller may assign an alternative-specific value to the *certain* option which represents her uncertainty evaluation for the anticipated waiting time. Thus, from this situation, the certainty equivalent of the ambiguous waiting time for the second (uncertain) option can be obtained by estimating the value assigned, holding other things equal, to the *certain* option. The traveller's uncertainty evaluation, and by extension her value of certainty, may also be affected by situational contexts such as delays in the system and the time she already spent waiting before the decision point. Since travellers are not likely to believe that the actual waiting time will be significantly lower than the displayed prediction, it is reasonable to expect that they do not dislike certainty — they are either indifferent or like certainty. This implies that if travellers, in general, believe the shown anticipated waiting time, the value of certainty would be lower than if there is a general perception of poor reliability.

For the proposed situation to take place, there must be a difference in in-vehicle times between the travel options. Moreover, the schedules or real-time delays must be such that the slower and faster vehicles are the certain and uncertain options, respectively (i.e., the slower vehicle arrives first at the origin). To assess the value of certainty in waiting time, choices between non(-strictly)-dominated alternatives must be observed. Assuming that travellers either like or are indifferent to certainty in waiting time, to ensure that the *certain* alternative does not dominate the uncertain one, the former must arrive at the destination later than the latter taking into account any weighting of travel time components. The uncertain vehicle can arrive at the destination before the certain one (i) if it can overtake the latter along a common path or (ii) if they serve two distinct lines.

The conditions outlined for the proposed situation to arise are not stringent and a number of examples can be found in the real world. Using published timetables of real-world public transport networks, specific examples can be found. For instance, the situation arises in the New York City subway and Mumbai commuter railways because express trains can overtake local ones (e.g., local and express lines 1 and 2 between 96 St and Chamber St in New York City; local and express trains between Borivali and Churchgate in Mumbai) (Indian Railways: Western Railway, 2021; MTA New York City Transit, 2020). Examples of the situation arising due to stops being connected by lines with distinct routes can also be found in the New York City subway as well as in the tram network of The Hague (e.g., lines 2 and 4 between 149 St and Franklin Av; lines 9 and 16 between Loevensteinlaan and Station Hollands Spoor) (HTM, 2021; MTA New York City Transit, 2020). Furthermore, even if the situation does not occur in a particular public transport network, given that the setup is fairly common in other networks, it is likely that travellers can identify with the situation. We emphasise that the proposed situation is a probe that permits the measurement of a relevant factor in travel behaviour, that is, uncertainty evaluation. There is little reason to believe that travellers' evaluation of waiting time uncertainty in this situation would be any different in other situations in the networks.

Choice observations may be analysed under the random utility maximization paradigm. To formulate the utility equations for the two options, we consider the four attributes involved in the choice situation described above: two main variables — in-vehicle times (IVT) and anticipated waiting time (AWT) — and two contextual variables — experienced waiting time (EWT) and anticipated delay (DEL). Furthermore, as the alternatives are labelled, the vehicle that arrives at the origin first (VEH1, the *certain* option) is assigned an alternative-specific constant ($\beta^{\text{certainty}}$) that represents the value of certainty attached to it. Since there are only two alternatives and only differences in utility matter, we set the utility of the second vehicle (VEH2) to zero. The (systematic parts of the) utilities of the two alternatives are then specified as follows:

$$\begin{aligned} V_{\text{VEH1}} &= \beta^{\text{certainty}} + \beta^{\text{IVT}} \cdot (IVT_1 - IVT_2) + \beta^{\text{AWT}} \cdot AWT + \beta^{\text{EWT}} \cdot EWT + \beta^{\text{DEL}} \cdot DEL \\ V_{\text{VEH2}} &= 0 \end{aligned} \tag{1}$$

Table 1
Attribute values used in the choice experiment.

Attribute	Attribute values (in minutes)
Experienced waiting time	0, 5, 10, 15
Anticipated delays in both trains	0, 5, 10, 15
In-vehicle time for the first train	14, 28
In-vehicle time for the second train	4, 8
Anticipated waiting time for the second train	4, 10

5. Case study: Dutch railways

As a case study, we assess the waiting time reliability beliefs of travellers in the Dutch railways by implementing the choice situation presented in the previous section in a stated preferences experiment. In the Netherlands, the railways are used widely, for different trip purposes and over a large range of travel times. Since trains may — and different services such as express (*Intercity*) and non-express (*Sprinter*) indeed do — overtake one another by skipping stations, the choice situation would not seem unrealistic to travellers. Furthermore, as required in the proposed choice situation, throughout all railway platforms in the Netherlands, real-time departure information is displayed in a uniform manner.

When the proposed choice situation is presented as a stated preference questionnaire, it has two important advantages over conventional travel time reliability behaviour stated preference experiments. First, since there are no objective probabilities, they do not have to be conveyed to respondents so that everyone can understand them; thus, circumventing a major issue in such experiments. Second, unlike conventional choice experiments where respondents are known to provide protest answers in such experiments to demonstrate (in an exaggerated manner) their dislike towards delays and irregularities in public transport services (Bates et al., 2001), it is less obvious to survey-takers what is being measured and therefore they are likely to indicate their 'true' preferences. Next, we discuss the experiment design, data collection, and the choice analyses.

5.1. Experiment design

The choice situation consists of the following variables: (i) time already waited or the experienced waiting time; (ii) the anticipated delays of the two trains; (iii) the in-vehicle times of the two trains; and (iv) the anticipated waiting time for the second train. The first variable, experienced waiting time, is a context variable as it holds true irrespective of the alternative chosen. Since, the objective is to understand how they affect the value of certainty (rather than their marginal disutility), the anticipated delays for the two trains are changed together. Thus, the anticipated delay in the two trains can also be considered to be a context variable.

Attribute values are based on the need to adhere to reality and the ability to obtain the required estimates from choice observations. Since there is no clear indication on the direction or magnitude of the effect context variables have on the value of certainty, it is interesting to test them more closely. To this end, four attribute levels are used for each context variable allowing testing for non-linearity. The selected values (Table 1) are quite realistic as delay information in the Dutch railways is indeed shown in five minute intervals while experienced waiting time is often rounded as it is difficult to be more precise when thinking about elapsed time.

The selection of attribute values for in-vehicle times and anticipated waiting times is a little trickier. The values of in-vehicle times and anticipated waiting times must be such that, given the expectations of traveller preferences, alternatives presented must not be dominated for a range of trade-off ratios between anticipated waiting time and in-vehicle time. Commonly, studies have found that waiting time is weighed 1.5–2 times compared to in-vehicle time (e.g., Yap et al., 2018). However, it is also possible that travellers directly compare expected arrival times at the destination, in which case the waiting time and in-vehicle time are weighted equally. Thus, the range of waiting time – in-vehicle time trade-offs considered here is from 1 (arrival time differences) to 2 (higher end amongst most findings). A trial-and-error approach is used to find which attribute values satisfy the set of objectives and constraints described below.

For all three variables — in-vehicle times for the two trains and anticipated waiting time for the second train — only two attribute levels are chosen. This results in 8 ($2 \times 2 \times 2$) possible utility differences for a given waiting to in-vehicle time coefficient ratio. We would like to select attribute values for these variables such that for both the lowest and highest ratios (i.e., 1 & 2), considering the alternatives to be unlabelled (i.e., without an alternate specific constant), amongst the 8 possible utility differences, there are at least: (i) 4 that are in favour of the second train, (ii) 1 that is neutral, and (iii) 1 that is in favour of the first train. The objectives are tilted in favour of the second train because people are expected to be neutral at the least but in general have a preference for certainty and therefore the alternative-specific utility of a certain waiting time is expected to be positive. The latter two objectives are set to prevent respondents from learning that the first train always arrives second at the destination as well as to allow observations to indicate that our expectation regarding the sign of utility of certainty is incorrect. In addition to these objectives, the following constraints are set on the attribute values: (i) the minimum anticipated waiting time is 4 minutes, (ii) the minimum in-vehicle time is 4 minutes, and (iii) the range of all attributes is at least 4 minutes. The first two constraints ensure realism of attribute values. A minimum attribute value range is set because a larger difference in alternative utilities requires fewer observations to estimate parameters. Note that only even values were used in order to reduce the search space. Table 1 shows the attribute values used in the experiment.

With these attributes and values, a simultaneous orthogonal fractional factorial design is found with NGENE. To limit the number of questions per respondent, the design is blocked into two parts. With this specification, a design with a total of 16 choice situations is found with 8 choice situations per respondent.

5.2. Data collection

The choice experiment was included within a larger survey that consisted of four parts, in this order: (0) screening, (1) socio-demographics, (2) choice experiment, and (3) qualitative measurements. The structure, content, and design of an initial draft of the survey were refined based on comments received from a small pilot of about 20 persons. The final version of the survey was offered in Dutch and had an expected completion time of 10 minutes. It was distributed to a predefined sample size of 700 respondents through an online panel, PanelClix. Given that most people in the Netherlands have access to the internet, this method of data collection does not create any obvious biases. The data collection took place in November–December 2018.

5.2.1. Screening and socio-demographics

Respondents were screened out if they used the trains less than once per month on the basis that if respondents do not meet this criterion, they are likely to not have well-formed evaluations of uncertainties in the railways. Regarding trip purpose, the survey aimed to collect about 80% of responses (550 responses) from those who used the railways for commuting either to work or education, and the rest from those with other purposes. The greater focus on commuters and efforts was, again, to ensure that those travelling more frequently are included since this group is more likely to have more well-formed value systems and uncertainty evaluations. Based on previous experience with the online panel, it was known that unemployed persons and those working part-time were slightly over-represented. Therefore, it was agreed, before the beginning of the distribution, that an additional restriction would be

Table 2
Sample characteristics.

Total respondents	703		
	Distribution (%)		
Attribute	Value	Actual	Required
Gender	Female	54.8%	50%
	Male	45.2%	50%
Age	<18	0.1%	0%
	18–24	32.7%	36%
	25–34	24.0%	17%
	35–44	15.4%	13%
	45–54	13.2%	16%
	55–64	10.8%	12%
	>64	3.7%	6%
Trip Purpose: Commuting	Work	53.3%	~80%
	Education	27.9%	
Trip Purpose: Others	Errands	0.7%	~20%
	Recreation	18.1%	
	Others	0.0%	
Trips per Week	0	1.8%	
	1	13.2%	
	2	18.8%	
	3	18.9%	
	4	22.0%	
	5	22.0%	
	6	2.4%	
	7	0.7%	

collected of which 703 met the completion time threshold.[1] While the survey was expected to take about 10 minutes on average, analysis of completion times after the collection of the required sample size revealed an average of about 6 minutes (after removing 12 respondents taking more than 20 minutes) and a median completion time of a little more than 5 minutes. Table 2 shows the distribution of respondent characteristics for the final set of valid responses.

5.2.2. *Choice experiment*

The choice experiment section in the survey begins with an explanation of the choice situation. Next, the respondent first faces a sample question which is not used in the analysis and then the 8 choice situations that will be used for the analysis. Each choice situation is prefaced by the instruction that there were two trains that could take them to their destination from the platform. To evoke the feeling of actually being at a station, respondents are shown information regarding the waiting times and anticipated delays of the two trains (TRN1, TRN2) in a format similar to the signboards found at platforms of the Dutch railways (Fig. 3). Respondents are informed that the images displayed are the state of the signboards at the decision point (as described in section 4). To remind survey-takers of the information shown in different parts of the signboard, an annotated version is also displayed in the example question. Separately from the signboard, information regarding the in-vehicle times and the time already waited is shown as a table and a line of text, respectively. Finally, the respondents are asked to choose

Fig. 3. Information displays at a real station (annotated).

placed in the form of a minimum frequency of travel by commuters, at least twice per week, if too many respondents chose a frequency of once per week or less (enforced after collecting 325 responses).

Desired socio-demographic quotas were obtained from data collected between 2011 and 2015 in a national, one-day, trip diary survey, OViN (*Onderzoek Verplaatsingen in Nederland*) conducted by the Dutch Central Bureau of Statistics (Centraal Bureau voor de Statistiek, 2015). The distribution of age, gender, and household incomes of respondents in that survey who use the railways at least once (during the day of reporting) are used as the desired stratification. It should be noted that these distributions were not weighted by the individual weights given in the survey as the group was reasonably large in itself.

To ensure response validity, those taking less than 4 minutes to complete the survey (40% of the expected time) were eliminated and more responses were added until the predefined target (of 700 responses) was reached. Eventually, a total of 918 responses were

whether they would board TRN1 or wait for TRN2. The order of the 8 situations as well as that of the two options in each situation were scrambled to avoid any biases. Fig. 4 shows a translated screenshot of a question in the choice experiment.

It is likely that respondents' uncertainty evaluations are affected by the time-of-day. Therefore, when not explicitly testing how this belief changes across different time periods in a day, it would be ideal reduce potential bias by not presenting any clock times. However, since the Dutch railways is a schedule-based system, train arrivals are associated with a particular clock-time and travellers are used to seeing this information on the signboards. Therefore, the planned departure time of the first train is fixed at 10:23. This time is somewhat neutral in the sense that it is just outside the morning peak (06:00–09:00) and not too far into the midday off-peak hours. Moreover, respondents may still be able to imagine using this train for different purposes. Finally, a rounded-off time such as 10:00 or 10:15 is intentionally not chosen because it might

[1] Analysis of the removed responses revealed very different behaviour from the rest of the sample confirming our suspicion that they were invalid. Using the original sample of 918 respondents, we also did not find any (non-negligible) systematic effects of completion times on attribute weights.

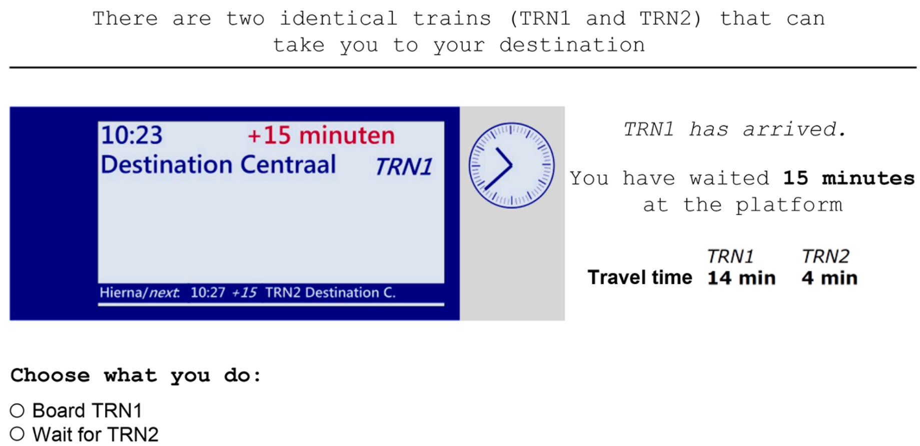

Fig. 4. Screenshot of a question in the choice experiment (translated to English).

Table 3
Overview of attributes included in the choice analyses.

Attributes	Symbol	Explanation	Range
Alternative attribute coefficients			
Certainty constant	$\beta^{certainty}$	–	–
In-vehicle time	β^{IVT}	All time attributes are in minutes	4–28
Anticipated waiting time	β^{AWT}		4–10
Experienced waiting time	β^{EWT}		0–15
Anticipated delays	β^{DEL}		0–15
Personal characteristics			
Socio-demographics			
Age	β^{age}	Ordinal in ascending order: < 18, 18–24, 25–34, 35–44, 45–54, 55–64, >64	1–7
Gender	β^{female}	Categorical (effect coded):female, male	
Net personal income	β^{income}	Ordinal in ascending order: unemployed, €0-11K, €11K-19K, €19K-30K, €30K-60K, >60K	1–6
Trip purpose	$\beta^{commuting}$	Categorical (effect coded): commuting, non-commuting	
Train use frequency	$\beta^{frequency}$	Average number of days train is used in a week	0–7

seem artificial and may induce respondents to act differently than they normally would; for instance, they may become more likely to calculate and focus on the final arrival time as it is easier to do so with round clock-times.

It should be noted that regardless of whether they choose to board the arrived train or wait for the next, respondents are not given any feedback on the outcomes, thus avoiding any learning effects and forcing respondents to continue to depend on evaluations formed in the real-world.

5.2.3. Qualitative measurements

Finally, the following factors are measured qualitatively on a Likert scale (with 7 levels): (i) regret anticipation, (ii) perception of reliability, and (iii) engagement level while waiting. The intention is not to include them in the modelling of uncertainty evaluation itself but rather analyse potential relationships between these indicators and stated preferences. The first, anticipation of regret, is considered to be one of the main psychological driving forces of risk aversion which leads to a preference for certainty. A standardized regret scale consisting of five items adopted from Schwartz et al. (2002) is used to measure it. This contains statements such as '*Whenever I make a choice, I try to get information about how the other alternatives turned out*' to which respondents indicate their level of agreement. The second factor assesses the perception of reliability of the network in general and in the presence of delays, and the perceived accuracy of displayed real-time information. This is tested using questions such as '*How reliable do you feel is the train arrival information?*' or '*When you at an NS platform, to what extent is your perception of reliability (for your trip) affected if the next two consecutive trains that you can take to your destination are delayed?*' Finally, as discussed in section 3, context can affect how waiting time is experienced. Occupied time has been consistently shown to reduce perceived waiting time (Jones, 1996; Molin et al., 2020) which could in turn affect beliefs regarding anticipated waiting time; therefore, the level of engagement of respondents at train platforms is measured through the following question: '*Usually, how engaged are you with the activity you perform while waiting at a railway platform?*'

5.3. Choice analysis

The decisions observed in the stated preference experiment are analysed using discrete choice models under the conventional framework of random utility maximization (RUM). Using the utility equations formulated in section 4, first multinomial logit (MNL) models are estimated to demonstrate the effect of accounting for the value of certainty (or the cost of uncertainty) on other choice parameters and to explore non-linear effects of contextual variables. In the RUM paradigm, the utility of an alternative a, U_a, consists of systematic (V_a) and random (ε) components. The systematic component is the product of the vector of taste preferences (β) and the vector of alternative attributes (x_a). Given that the random component in an MNL model is Gumbel distributed, the probability of choosing alternative i from I alternatives is given by the following:

$$U_a = V_a + \varepsilon; \; V_a = \beta \cdot x_a; \; P_{ni} = \frac{e^{V_i}}{\sum_{a=1}^{I} e^{V_a}} \tag{2}$$

Next, heterogeneity in behaviour is assessed using latent class choice models (LCCM) which capture decision-maker heterogeneity through a discrete mixture of choice models. In LCCM, individuals are probabilistically allocated to latent classes each of which have their own choice models. Depending on the objective, different choice models may be used in each class but in this study, the MNL model, based on the utility equations presented in section 4, is used as the underlying behaviour model for each class. To represent this mathematically, consider individual n who belongs to class s (amongst S classes) with probability π_{ns}. Then the probability that this individual selects alternative i is the product-sum of the class membership probabilities and the probability of selecting that alternative for each class (given the vector of taste parameters in that class, β_s):

$$P_{ni} = \sum_{s=1}^{S} \pi_{ns} \cdot P_{ni}(\beta_s) \tag{3}$$

If we assume intra-individual homogeneity in sensitivities, that is, account for panel effects, we essentially say that a particular individual is allocated to each class with the same probability for all choices they make. Thus, the likelihood of observing the sequence of choices $i{:}i_1,\ldots,i_T$ by individual n over T situations is given by the following:

$$L_{ni} = \sum_{s=1}^{S} \pi_{ns} \prod_{t=1}^{T} P_{ni_t}(\beta_s) \tag{4}$$

Apart from accounting for heterogeneity in tastes, an important advantage of LCCM is that individuals' preferences can be explained by using a class membership model to link membership probabilities with individuals' characteristics. The commonly used, logit function is also used here as the class membership model. We use the socio-demographic and qualitative measures collected (see Table 3) as the individual characteristics influencing class membership. For this vector of individual characteristics, z_n, and to-be-estimated, class-specific regression parameters, coefficient vectors, γ_s, and constants, δ_s, the class membership probability is given by:

$$\pi_{ns} = \frac{e^{\delta_s + g(\gamma_s, z_n)}}{\sum_{a=1}^{S} e^{\delta_a + g(\gamma_a, z_n)}} \tag{5}$$

The flexibility of the LCCM means that there are a number of ways to specify the model. The researcher needs to decide the number of classes, the parameters to be included in the choice models in each class, and the parameters to be included in the class membership model. Since there are no prescribed methodologies to arrive at the final model, we define here the sequence of steps taken to obtain our models. First, we include all choice parameters and class membership model constants and find the optimal number of classes. The model fit with different number of classes is assessed using the Bayesian information criterion (BIC) which explicitly penalizes the inclusion of extra parameters. Then for the optimal number of classes, the choice models in each class are finalized by removing highly insignificant ($p > 0.2$) parameters one-by-one. Next, all observable individual characteristics (socio-demographics) are added to the class membership function and the model is finalized by removing those that do not have a significant effect. Table 3 shows an overview of all the attributes used in the choice analysis. The psychometric questions can be found in Appendix A. All choice model estimations are carried out using PythonBiogeme (Bierlaire, 2016).

As noted previously, the collected psychometric indicators were not included in the model itself; instead, the distribution of the unobservable qualitative measures in each class is used for characterising class composition through a posterior analysis of class membership. However, we also estimate a hybrid choice model (HCM) where the class membership model is directly related to the indicators through a measurement model in a framework similar to that employed by Atasoy et al. (2013) and Hurtubia et al. (2014). In this model, the likelihood function given for individual n in the latent class choice model (equation 4) is modified. In addition to the likelihood of observing a particular sequence of choices ($\boldsymbol{i}{:}i_1,\ldots,i_T$), the likelihood of obtaining a particular response pattern ($\boldsymbol{r}{:}r_1,\ldots,r_K$) for the indicators ($\mathbf{K}$) is also included (equation 6 below). The probability of obtaining a particular response ($\pi^{k,r}$) is treated as a constant for each class and is estimated directly as a parameter in the model using the indicator responses. Thus, as Atasoy et al. (2011) note, the measurement model for the psychometric indicators helps identifying the latent classes by using responses to these indicators. Since the HCM accounts for these responses, it might lead to different latent classes or newer insights that do not surface in the posterior analysis of the LCCM.

$$L_{nik} = \sum_{s=1}^{S} \pi_{ns} \left\{ \prod_{t=1}^{T} P_{ni_t}(\beta_s) \right\} \left\{ \prod_{k=1}^{K} \pi_s^{k,r} \right\} \tag{6}$$

A large number of parameters has to be estimated with this approach. If all indicators are used a total of 162 parameters have to be estimated to obtain the indicator response likelihood (162 = 3 classes × 9 indicators × (7–1) levels). Therefore, we reduce the number of indicators by selecting only one each for regret and reliability perception (from a set of 5 and 2, respectively), and the indicators for engagement while waiting and effect of delays. The indicators for regret and reliability perception are selected based on an exploratory factor analysis and overall model fitness. The full results of the HCM can be found in Appendix B. Estimates for the parameters common to the hybrid choice and latent class choice models were found to be fairly similar. Moreover, the parameters estimated for the indicators in the HCM follow the same trends as their corresponding class profiles in the LCCM posterior analysis. Similarities in the two models may be because, in the HCM, the measurement model does not contribute substantially to the identification of the latent classes in comparison to the class membership model or choice models. Therefore, we choose not to use the HCM results because: first, the more complex HCM offers similar interpretation of the heterogeneity in choice behaviour, hence the parsimonious LCCM is considered superior; and second, the extra information obtained in the HCM as class profiles of indicators can also be obtained through the posterior analysis mentioned above.

6. Results

As discussed in section 5.3, we first present results of the multinomial logit models; specifically, the effect of accounting for uncertainty and context variables. Then, heterogeneity in behaviour is presented through distinct behavioural profiles identified by a latent class choice model which also explains membership to these profiles with socio-demographic and other personal factors.

6.1. Multinomial logit models

To analyse the effect of including a certainty parameter, as opposed to conventional route choice models that consider alternatives to be unlabelled, in addition to the labelled MNL model (MNL^L) that uses the equations presented in section 4, an unlabelled version (MNL^U) that does not include $\beta^{\text{certainty}}$ is also estimated (Table 4). The significant and positive alternative-specific constant in the MNL^L model clearly rejects the null hypothesis that there is no effect of uncertainty and shows a preference for certainty. The coefficients for travel time components in both models are also significant and have the expected signs: as the anticipated waiting time increases or the first train is less slow in comparison, the preference for the first train increases. Since the context variable parameters in MNL^L model are small and insignificant ($p > 0.2$), in the model shown in Table 4 they are fixed to zero. The most likely reason for finding these parameters significant in MNL^U but not in MNL^L is that, in the absence of an alternative-specific constant in the former model, these parameters also partially capture respondents' overall preference for certainty. In the MNL^U model, where the contextual variable parameters are significant, the signs of the context variables seem to be reasonable. Regarding delays, one can expect travellers to be increasingly wary of waiting for TRN2 as the delays increase. For experienced waiting time, as discussed in section 3, there is no clear intuition regarding the effect direction since travellers might either experience frustration or take into account sunk costs. Moreover, some people may begin to engage in an activity that distracts them from waiting after some threshold of experienced waiting time. An overall positive effect is found and it may be justified — the more time has elapsed, the more travellers just want to take the train that comes first, all other things being equal (Osuna, 1985).

Table 4
Estimation results of the different multinomial logit models.

Model	MNL^U		MNL^L		$MNL^{L\text{-}nl}$	
# parameters		4		3		6
Initial LL		–3898.260		–3898.260		–3898.260
Final LL		–3366.782		–3331.959		–3321.126
Adjusted ρ^2		0.135		0.145		0.147
BIC		6768.102		6689.822		6694.06
	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
$\beta^{certainty}$	–	–	0.947	0.00	0.94	0.00
β^{IVT}	–0.108	0.00	–0.123	0.00	–0.124	0.00
β^{AWT}	0.132	0.00	0.080	0.00	0.081	0.00
β^{EWT}	0.023	0.00	–	–	–	–
β^{DEL}	0.015	0.00	–	–	–	–
$\beta^{EWT\text{-}0}$	–	–	–	–	–0.206	–
$\beta^{DEL\text{-}0}$	–	–	–	–	0.094	–
$\beta^{EWT\text{-}5}$	–	–	–	–	–	–
$\beta^{DEL\text{-}5}$	–	–	–	–	–0.094	0.03
$\beta^{EWT\text{-}10}$	–	–	–	–	0.091	0.06
$\beta^{DEL\text{-}10}$	–	–	–	–	–	–
$\beta^{EWT\text{-}15}$	–	–	–	–	0.115	0.02
$\beta^{DEL\text{-}15}$	–	–	–	–	–	–

A log-likelihood ratio test between the models shows that, the MNL^L model clearly outperforms its unlabelled counterpart ($p < 0.001$). We cross-validate this improvement using a k-fold procedure with 14 folds such that all observations from one individual are either in the training or testing data set. The cross-validation reveals similar improvements in likelihood of chosen alternatives in the test dataset: −3335.37 versus −3371.20 with the MNL^L and MNL^U models, respectively. An important difference between these models is in the β^{AWT}–β^{IVT} ratio. In the unlabelled model this ratio is 1.22, a value close to results in literature which have commonly found that waiting time weighs higher in travellers minds than in-vehicle time (e.g., Yap et al., 2018). However, once the waiting time uncertainty is accounted for in the MNL^L model, the ratio becomes 0.65 indicating the large role of uncertainty in the travellers' assessment of waiting time. Furthermore, the MNL^L model also shows that travellers are willing to trade-off 7.70 minutes (0.947 ÷ 0.123) of in-vehicle time for certainty in their waiting time.

Although the context variables seem to have no effect in the MNL^L model, since four levels were included for each variable it is possible to check whether they really do not affect decision-making or they have a non-linear nature which averages out. While this is less likely for delays where we have a clear intuition regarding the effect direction, it may very well be true for experienced waiting time where there is an interplay between the effects of frustration and sunk time costs. The variables are effect coded with the level with 0 minutes as the reference. Effect coding allows us to separate the effect of the reference level from the constant. The variables for 10 and 15 minutes of delay, and for 5 minutes of experienced waiting time have high p-values ($p > 0.2$) and are therefore fixed to zero. The final model is shown in Table 4 as $MNL^{L\text{-}nl}$. The results include the coefficient for the reference level which is computed as the sum of the negatives of all the other coefficients for that attribute. Using the log-likelihood ratio test, this model is found to perform better than the MNL^L model ($p < 0.001$). The signs for anticipated delays are not as expected and it is difficult to hypothesize why a delay of 5 minutes seems to make it more likely that the second train will be chosen. The signs for experienced waiting time, however, can be explained by a combination of frustration/anxiety effects and sunk time/activity engagement effects. The likelihood of choosing TRN1 first increases up to 5 minutes (arguably due to frustration/anxiety), then stabilizes between 5 and 10 minutes (more likely to be engaged in an activity), and then falls again (sunk time/activity engagement).

6.2. Latent class choice models

Using the steps defined in section 5.3 yields a 4-class model as the one with the best trade-off between efficiency and model fit. However, two classes have a membership of less than 10% which means that the choice parameter estimations within these models would likely have high errors. Therefore, we remove one class and estimate a 3-class model which has a comparable model fit, has reasonable class sizes and offers better interpretability. Table 5 shows the final model. To report results, the class with the smallest size is used as the reference for the class membership model (i.e., for the smallest class, $\delta_s = 0, \gamma_s = 0$).

In the largest class (55%), behaviour is similar to the MNL^L model with an additional effect wherein the value of certainty increases slightly with delays. Travellers in this class are willing to trade-off about 5.3 minutes of extra in-vehicle time to remove uncertainty in their waiting time. With each minute of delay, travellers are willing to further accept approximately 4 seconds of additional in-vehicle time. Similar to the MNL^L model, once the value of certainty is accounted for, they weigh anticipated waiting time slightly less than in-vehicle time (0.86:1). Group membership is more likely for younger travellers. Similar to their

Table 5
Estimation results of the 3-class latent class choice model.

Model	LCCM 3-Class					
# parameters	12					
Initial LL	−4159.808					
Final LL	−3063.483					
Adjusted ρ^2	0.261					
BIC	6230.584					
	Class 1		***Class 2***		***Class 3***	
Class Size	54.74%		28.41%		16.84%	
	Class-specific choice models					
	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
$\beta^{certainty}$	1.61	0.00	–	–	0.983	0.01
β^{IVT}	−0.301	0.00	−0.061	0.00	−0.0487	0.01
β^{AWT}	0.258	0.00	–	–	0.126	0.00
β^{EWT}	–	–	–	–	0.0268	0.12
β^{DEL}	0.019	0.18	–	–	–	–
	Class membership models					
	Class 1		***Class 2***		***Class 3 (ref.)***	
	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
$\beta^{intercept}$	2.00	0.00	–	–	0	–
β^{age}	−0.236	0.00	0.134	0.00	–	–

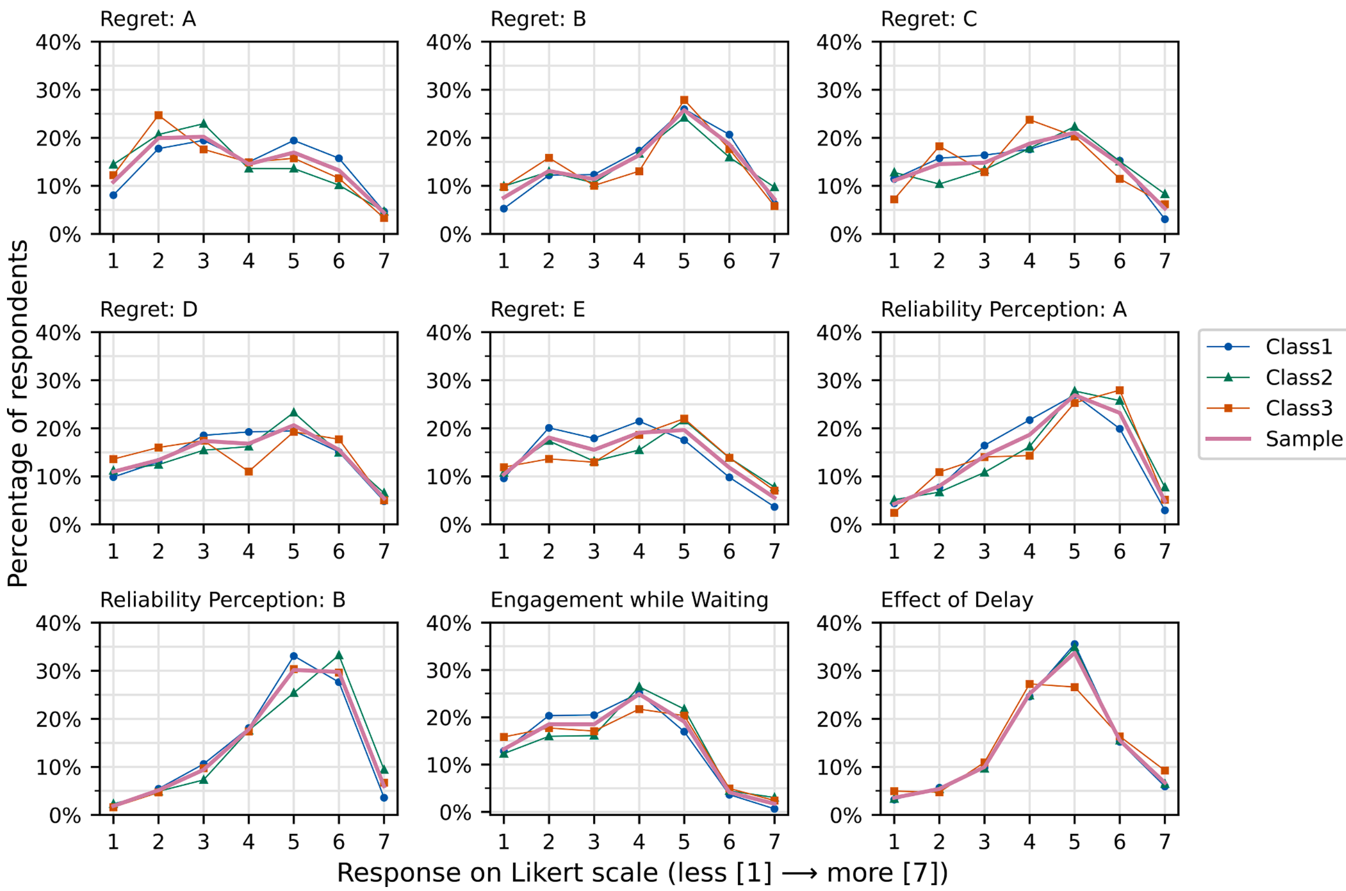

Fig. 5. Class profiles of the psychometric indicators (see Appendix A).

preference for certainty here, younger travellers are also found to be more risk averse by de Palma and Picard (2005) in their departure time choice study.

The second group (28%) shows lexicographic preferences (at least in the range of the attribute levels presented in the stated preferences experiment) for faster trains, making decisions only on the basis of in-vehicle time and, apparently, not caring about other factors. In addition to their inherent preferences, it is possible that those who strongly prefer the faster train, may have translated the offered alternatives into real-life services, where the trains are, in fact, different, and thus chosen one train type over another for reasons not measured in the survey. In the Netherlands, the express trains (*Intercity*) offer additional services such as air-conditioning, Wi-Fi internet, and toilets. Older travellers are more likely to be in this class.

Although, the third group (17%) shows some compensatory behaviour, travellers in this group seem to strongly dislike uncertainty and are willing to accept more than 20 minutes of extra in-vehicle time for certainty in their waiting time. Thus, their preferences are nearly lexicographic in favour of the first train to arrive. Furthermore, frustration seems to play a substantial role for this group: with every minute spent waiting in the past (which should therefore be irrelevant for the decision at hand), there is a willingness to accept an additional 33 seconds of in-vehicle time for certainty in waiting time.

Posterior analysis of the class membership does not reveal substantial differences between classes in terms of distribution of psychometric indicators (Fig. 5). Visual inspection of the trends shows that those showing fully compensatory behaviour (Class 1) have a slightly lower trust in the reliability (indicators Reliability Perception A and B) of the system. Moreover, based on regret indicators C, D, and E, respondents in this group are also a little less regret-averse than the sample is on average.

7. Conclusion

Although decisions in the real world are almost always taken under uncertainty, that is, in the absence of objective probabilities, most existing studies on the effects of waiting time reliability on travel behaviour observe or analyse travel decisions (as if) made using objective probabilities. Capturing travellers' evaluations, which are a result of complex interactions between their perceptions and attitudes, regarding uncertainty in public transport waiting times is difficult. Therefore, this study identifies a realistic route choice situation where such evaluations can be quantified under natural ambiguity without using objective probabilities or assuming specific learning behaviour. In the slow/fast lines experiment proposed, uncertainty evaluations can be quantified as a certainty equivalent or, as shown, an alternative-specific constant under the random utility maximization regime. Studies in behavioural economics and psychology have indicated that contexts are important in decision making. In addition to quantifying the evaluations in general, we are also able to estimate the effect of contextual attributes on them; for instance, the effect of time spent waiting before making a decision based on anticipated time to be waited.

Through a stated preferences experiment with the identified choice situation, we find a strong preference for certainty in travellers of the

Dutch railways. Accounting for uncertainty explained away some of the waiting time parameter, reducing the waiting to in-vehicle time ratio to less than one. Contextual attributes do not seem to have an effect on average although small, non-linear effects were found for both experienced waiting time and anticipated delays. A latent class choice model indicated three groups of travellers: the biggest group making fully compensatory choices, weighing uncertainty against travel time attributes, and two others showing lexicographic behaviour, choosing the fastest and the first train, respectively. While age seems to affect association with different behavioural profiles, there are only minor differences between the distribution of psychometric indicators in different classes.

The choice situation proposed in this study offers a relatively simple method to obtain snapshots of evaluations of uncertainties in a real-world public transport network. With respect to planning of services, transportation models can benefit from the added accuracy obtained by explicitly quantifying the effects of uncertainty (as indicated by the improved model fit and predictive value). The proposed situation is used to measure uncertainty evaluation and inferences are not limited to this exact situation — for instance, the finding that associated uncertainty has a large role in travellers' assessment of waiting time holds over all decisions of the type 'whether to board or wait' and could be useful for agent-based models (e.g., Cats and Gkioulou (2017)) that commonly simulate this choice. Often when biases are pointed out to decision-makers, they choose to correct their choices to more 'rational' ones (Gilboa, 2009). Journey planner applications may use choice observations in situations similar to the one used in this study to provide feedback highlighting such potential biases (e.g., loss aversion, overweighting of small probabilities) that travellers might want to correct on reconsideration. Moreover, through association of behaviour under uncertainty with introspective psychological measures, such applications can offer targeted actions to specific groups of travellers to bring their evaluations in line with empirical realities. For instance, applications may work on distracting travellers from the boredom of waiting by engaging them in an activity such as reading. Since the experiment also permits measuring the effects of contextual variables, it may be used to analyse situations which exaggerate feelings of uncertainty and take suitable actions for this. On a related note, the certainty equivalent presented here may also be used as an indicator for A/B type tests when transportation authorities wish to introduce new measures aimed at improving feelings regarding uncertainty. For instance, indicating the cause of delays has been proposed to reduced anxiety associated with uncertainty in waiting time (Maister, 1985). The extent to which this measure is effective may be quantified by comparing certainty equivalents obtained for the identified choice situation in the control and treatment groups.

While the choice situation and experiment are carefully designed, there remain some limitations that may affect estimation and interpretation of results. In our experiment, we assume that travellers usually make a conscious choice regarding boarding a train or waiting for the next one. While it is likely that this is a conscious choice, especially for regular travellers who are aware of different lines that can take them to their destination, it is possible that by presenting this choice situation we highlight the uncertainty involved in waiting times thereby making people more averse to it.

As discussed in the theoretical framework for this study, we measure the combined outcome of travellers' perceptions and attitudes on their decisions under uncertainty as subjective beliefs. However, disentangling the effects of these individual determinants on travel behaviour may allow more effective policies and travel advice. In order to analyse attitudes and perceptions separately, we might need to model more complex decision rules and, perhaps, observe different choice situations or sequence of decisions. The challenge will be to do this also directly from observations of real-world trips (i.e., not in an laboratory experimental context), without having to observe risky choices or ask for matching probabilities of uncertain events, both of which require interaction between the researcher and travellers.

Finally, a potential limitation in our study is related to the experiment type itself. We would like to measure the effects of contextual attributes on subjective beliefs but, arguably, it is difficult for respondents to account for such effects separately from their general aversion to uncertainty. For example, respondents may not be able to feel the effect of having waited ten minutes when making a choice in a stated preference questionnaire yet anecdotal evidence would suggest that this variable indeed has an impact on boarding decision. It is possible that this may be why we do not find strong effects of contextual variables in our case study. Using incentivized laboratory experiments, common in behavioural economics, does not help either because, since these are contextual variables, they cannot be incentivized one way or another. Thus, the effects of contextual variables may be best measured in a revealed preferences setting, that is, from observations of real-world trips where travellers actually experience the context. Since the proposed choice situation is realistic, this might be actually possible for many public transport networks.

Apart from future work indicated by the limitations outlined above, other avenues of research may be found in the theoretical framework presented in this study. In our analysis of Dutch railway travellers, we considered the effects of two situational contexts, namely, experienced waiting time and delays on the travellers' corridor. Similarly, other situational contexts such as the effects of delays in other parts of the network, or the differences between trip purposes, such as travelling to and from work may be studied. Furthermore, the effects of affective contexts on subjective beliefs can be investigated to assess the indirect effects of various factors affecting moods, such as station lighting. In decisions under ambiguity, such as route choice in public transport networks, where decision-makers can observe the choices of others, herding effects also become important and may be analysed. Finally, while our method provides a snapshot of subjective beliefs towards waiting time uncertainties in real-world networks, it would be interesting to observe the evolution of such snapshots over time for different individuals in various networks.

CRediT authorship contribution statement

Sanmay Shelat: Conceptualization, Formal analysis, Methodology, Writing - original draft. **Oded Cats:** Supervision, Writing - review & editing. **J.W.C. van Lint:** Supervision, Writing - review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is supported by the My-TRAC (H2020 Grant No. 777640) project.

Appendix A

Psychometric indicators collected in the survey are presented in Table 6.

Table 6
Questions used for psychometric indicators (in English).

Variable Name in Fig. 5	Question (less [1] → more [7])
Regret: A	Once I make a decision, I don't look back. (the response order is reversed)
Regret: B	Whenever I make a choice, I'm curious about what would have happened if I had chosen differently.
Regret: C	Whenever I make a choice, I try to get information about how the other alternatives turned out.
Regret: D	If I make a choice and it turns out well, I still feel like something of a failure if I find out that another choice would have turned out better.
Regret: E	When I think about how I'm doing in life, I often assess opportunities I have passed up.
Reliability Perception: A	How reliable do you feel is the train arrival information?
Reliability Perception: B	How reliable do you feel is the Dutch Railways in general?
Engagement while Waiting	Usually, how engaged are you with the activity you perform while waiting at a railway platform?
Effect of Delay	When you are at an NS platform, to what extent is your perception of reliability (for your trip) affected if the next two consecutive trains that you can take to your destination are delayed?

Appendix B

Results from the hybrid choice model estimation are presented in Table 7. Since the model is not used in the manuscript, we did not refine the model further after the first estimation (e.g., by removing parameters with p-values above our assumed thresholds of insignificance). We note that the common parameters (highlighted in bold) between this model and that without indicators in Table 5 are fairly similar. Furthermore, the indicator parameters generally follow the trends of the respective values in the class profiles obtained in the posterior analysis in the manuscript.

Table 7
Estimation results of the latent class model with indicators. Indicator descriptions can be found in Table 6. The parameters included in the manuscript model are in bold. Note: $\pi^{k,7} = 1-\sum_{r=1}^{r=6}\pi^{k,r}$

Model	LCCM 3-Class (with indicators)
# parameters	90
Initial LL	−12911.755
Final LL	−7902.731
Adjusted ρ^2	0.381
BIC	16582.593

	Class 1		*Class 2*		*Class 3*	
Parameter	Coeff.	p-val	Coeff.	p-val	Coeff.	p-val
Choice parameters						
$\beta^{certainty}$	**1.410**	0.000	0.361	0.160	**0.811**	0.030
β^{IVT}	**−0.284**	0.000	**−0.080**	0.000	**−0.033**	0.260
β^{AWT}	**0.251**	0.000	−0.004	0.860	**0.116**	0.000
β^{EWT}	0.010	0.470	−0.003	0.820	**0.025**	0.220
β^{DEL}	**0.026**	0.230	−0.007	0.700	0.002	0.930
Class membership parameters						
$\beta^{intercept}$	**1.99**	0	–	–	**0**	–
β^{age}	**−0.235**	0.01	**0.15**	0.01	–	–
Indicator probability: Regret B						
$\pi^{Regret\_B,1}$	0.018	0.310	0.136	0.000	0.100	0.080
$\pi^{Regret\_B,2}$	0.114	0.010	0.136	0.010	0.162	0.060
$\pi^{Regret\_B,3}$	0.129	0.000	0.105	0.010	0.094	0.080
$\pi^{Regret\_B,4}$	0.169	0.000	0.177	0.000	0.122	0.070
$\pi^{Regret\_B,5}$	0.288	0.000	0.203	0.000	0.290	0.040
$\pi^{Regret\_B,6}$	0.224	0.000	0.144	0.000	0.178	0.050
$\pi^{Regret\_B,7}$	0.058	–	0.101	–	0.054	–
Indicator probability: Reliability Perception B						
$\pi^{ReliabilityPercep\_B,1}$	0.011	0.120	0.030	0.040	0.013	0.410
$\pi^{ReliabilityPercep\_B,2}$	0.065	0.000	0.036	0.070	0.046	0.240
$\pi^{ReliabilityPercep\_B,3}$	0.127	0.000	0.051	0.060	0.096	0.170
$\pi^{ReliabilityPercep\_B,4}$	0.183	0.000	0.172	0.020	0.175	0.140
$\pi^{ReliabilityPercep\_B,5}$	0.374	1.000	0.208	0.010	0.305	0.130
$\pi^{ReliabilityPercep\_B,6}$	0.236	0.000	0.371	0.000	0.305	0.110
$\pi^{ReliabilityPercep\_B,7}$	0.004	–	0.132	–	0.060	–
Indicator probability: Engagement while Waiting						
$\pi^{WaitEngage,1}$	0.090	0.000	0.170	0.030	0.163	0.150
$\pi^{WaitEngage,2}$	0.214	0.000	0.148	0.040	0.186	0.170
$\pi^{WaitEngage,3}$	0.235	0.000	0.130	0.070	0.168	0.190
$\pi^{WaitEngage,4}$	0.257	1.000	0.258	0.030	0.202	0.180
$\pi^{WaitEngage,5}$	0.173	0.000	0.207	0.040	0.205	0.180
$\pi^{WaitEngage,6}$	0.029	0.020	0.051	0.050	0.054	0.210
$\pi^{WaitEngage,7}$	0.003	–	0.036	–	0.023	–
Indicator probability: Effect of Delay						
$\pi^{DelayEffect,1}$	0.000	1.000	0.074	0.040	0.048	0.190
$\pi^{DelayEffect,2}$	0.049	0.160	0.065	0.070	0.046	0.160
$\pi^{DelayEffect,3}$	0.105	0.090	0.085	0.030	0.116	0.100
$\pi^{DelayEffect,4}$	0.262	0.070	0.223	0.020	0.286	0.070
$\pi^{DelayEffect,5}$	0.382	0.060	0.325	0.030	0.245	0.060
$\pi^{DelayEffect,6}$	0.155	0.060	0.148	0.020	0.170	0.080
$\pi^{DelayEffect,7}$	0.049	–	0.079	–	0.088	–

References

Abenoza, R.F., Cats, O., Susilo, Y.O., 2018. How does travel satisfaction sum up? An exploratory analysis in decomposing the door-to-door experience for multimodal trips. Transportation 46 (5), 1615–1642.
Atasoy, B., Glerum, A., Bierlaire, M., 2011. Mode choice with attitudinal latent class: a Swiss case-study. Second Int. Choice Model. Conf.
Atasoy, B., Glerum, A., Bierlaire, M., 2013. Attitudes towards mode choice in Switzerland. *disP-The.* Planning Rev. 49 (2), 101–117.
Avineri, E., Prashker, J.N., 2005. Sensitivity to travel time variability: Travelers' learning perspective. Transp. Res. Part C: Emerg. Technol. 13 (2), 157–183.
Avineri, E., Prashker, J.N., 2006. The Impact of Travel Time Information on Travelers' Learning under Uncertainty. Transportation 33 (4), 393–408.
Baillon, A., Huang, Z., Selim, A., Wakker, P.P., 2018. Measuring ambiguity attitudes for all (Natural) events. Econometrica 86 (5), 1839–1858.
Bates, J., Polak, J., Jones, P., Cook, A., 2001. The valuation of reliability for personal travel. Transp. Res. Part E: Logist. Transp. Rev. 37 (2-3), 191–229.
Ben-Elia, E., Di Pace, R., Bifulco, G.N., Shiftan, Y., 2013. The impact of travel information's accuracy on route-choice. Transp. Res. Part C: Emerg. Technol. 26, 146–159.
Ben-Elia, E., Erev, I., Shiftan, Y., 2008. The combined effect of information and experience on drivers' route-choice behavior. Transportation 35 (2), 165–177.
Ben-Elia, E., Shiftan, Y., 2010. Which road do I take? A learning-based model of route-choice behavior with real-time information. Transp. Res. Part A: Policy Pract. 44 (4), 249–264.
M. Bierlaire. PythonBiogeme: a short introduction, Series on Biogeme. Transport and Mobility Laboratory, School of Architecture, Civil and Environmental Engineering, Ecole Polytechnique Fédérale de Lausanne, Switzerland. 2016.
Bogers, E.A.I., Viti, F., Hoogendoorn, S.P., 2005. Joint modeling of advanced travel information service, habit, and learning impacts on route choice by laboratory simulator experiments. Transp. Res. Record 1926 (1), 189–197.
Bogers, E.A.I., Viti, F., Hoogendoorn, S.P., Van Zuylen, H.J., 2006. Valuation of different types of travel time reliability in route choice:large-scale laboratory experiment. Transp. Res. Record 1985 (1), 162–170.
Brownstone, D., Small, K.A., 2005. Valuing time and reliability: assessing the evidence from road pricing demonstrations. Transp. Res. Part A: Policy Pract. 39 (4), 279–293.
Carrion, C., Levinson, D., 2012. Value of travel time reliability: a review of current evidence. Transp. Res. Part A: Policy Pract. 46 (4), 720–741.
Cats, O., Gkioulou, Z., 2017. Modeling the impacts of public transport reliability and travel information on passengers' waiting-time uncertainty. EURO J. Transp. Logist. 6 (3), 247–270.
Centraal Bureau voor de Statistiek, 2015. Onderzoek Verplaatsingen in Nederland 2014.
de Palma, A., Picard, N., 2005. Route choice decision under travel time uncertainty. Transp. Res. Part A: Policy Pract. 39 (4), 295–324.
Dixit, V., Jian, S., Hassan, A., Robson, E., 2019. Eliciting perceptions of travel time risk and exploring its impact on value of time. Transp. Policy 82, 36–45.

Dziekan, K., Vermeulen, A., 2006. Psychological effects of and design preferences for real-time information displays. J. Public Transp. 9 (1), 1–19.
Ellsberg, D., 1961. Risk, Ambiguity, and the Savage Axioms. Q. J. Econ. 75 (4), 643. https://doi.org/10.2307/1884324.
Ghosh, A., 2001. Valuing Time and Reliability: Commuters' Mode Choice From a Real Time Congestion Pricing Experiment. University of California Irvine.
Gilboa, I., 2009. Theory of Decision Under Uncertainty. Cambridge University Press.
Guo, Z., Wilson, N.H.M., 2011. Assessing the cost of transfer inconvenience in public transport systems: A case study of the London Underground. Transp. Res. Part A: Policy Pract. 45 (2), 91–104.
Hensher, D.A., 2001. Measurement of the valuation of travel time savings. J. Transp. Econ. Policy (JTEP) 35, 71–98.
HTM, Timetable. https://www.htm.nl/en/timetable/. Accessed on: 30 May 2021.
Hurtubia, R., Nguyen, M.H., Glerum, A., Bierlaire, M., 2014. Integrating psychometric indicators in latent class choice models. Transp. Res. Part A: Policy Pract. 64, 135–146.
Indian Railways: Western Railway, Mumbai Suburban Time Table. https://wr.indianrailways.gov.in/view_section.jsp?id=0,6,458. Accessed on: 30 May 2021.
Jones, P., Peppiatt, E., 1996. Managing perceptions of waiting times in service queues. Int. J. Service Ind. Manage. 7 (5), 47–61.
Kemel, E., Paraschiv, C., 2013. Prospect Theory for joint time and money consequences in risk and ambiguity. Transp. Res. Part B: Methodol. 56, 81–95.
Knight, F.H., 1921. Risk, Uncertainty and Profit. Houghton Mifflin Company, Boston, New York.
Lam, T.C., Small, K.A., 2001. The value of time and reliability: measurement from a value pricing experiment. Transp. Res. Part E: Logist. Transp. Rev. 37 (2-3), 231–251.
Levinson, D., Harder, K., Bloomfield, J., Winiarczyk, K., 2004. Weighting waiting: evaluating perception of in-vehicle travel time under moving and stopped conditions. Transp. Res. Record 1898 (1), 61–68.
Li, Z., Hensher, D., 2011. Prospect theoretic contributions in understanding traveller behaviour: a review and some comments. Transp. Rev. 31 (1), 97–115.
Li, Z., Hensher, D., 2019. Understanding risky choice behaviour with travel time variability: a review of recent empirical contributions of alternative behavioural theories. Transp. Lett. 12 (8), 580–590.
Machina, M.J., Siniscalchi, M., 2014. Chapter 13 - Ambiguity and Ambiguity Aversion. In: Machina, M., Viscusi, K. (Eds.), Handbook of the Economics of Risk and Uncertainty. North-Holland, pp. 729–807.
Maister, D., 1985. The psychology of waiting lines. In: Czepiel, J.A., Solomon, M., Surprenant, C.S. (Eds.), The Service Encounter. Lexington Books.
Meng, M., Rau, A., Mahardhika, H., 2018. Public transport travel time perception: effects of socioeconomic characteristics, trip characteristics and facility usage. Transp. Res. Part A: Policy Pract. 114, 24–37.
Molin, E., Adjenughwure, K., de Bruyn, M., Cats, O., Warffemius, P., 2020. Does conducting activities while traveling reduce the value of time? Evidence from a within-subjects choice experiment. Transp. Res. Part A: Policy Pract. 132, 18–29.
MTA New York City Transit, Schedules (Subway Timetable, Effective September 13, 2020). https://new.mta.info/schedules. Accessed on: 30 May 2021.
Nau, R.F., 2001. De finetti was right: probability does not exist. Theor. Decis. 51, 89–124.
Osuna, E.E., 1985. The psychological cost of waiting. J. Math. Psychol. 29 (1), 82–105.
S. Peer, J Knockaert, P. Koster, E.T. Verhoef. Over-reporting versus overreacting: Commuters' perceptions of travel times, Tinbergen Institute Discussion Paper. Tinbergen Institute. 2014.
Quiggin, J., 1982. A theory of anticipated utility. J. Econ. Behav. Organ. 3 (4), 323–343.
Ramos, G.d.M., Daamen, W., Hoogendoorn, S., 2011. Expected utility theory, prospect theory, and regret theory compared for prediction of route choice behavior. Transp. Res. Rec. J. Transp. Res. Board 2230 (1), 19–28.
Rasouli, S., Timmermans, H., 2014. Applications of theories and models of choice and decision-making under conditions of uncertainty in travel behavior research. Travel Behav. Soc. 1 (3), 79–90.
Raveau, S., Muñoz, J.C., de Grange, L., 2011. A topological route choice model for metro. Transp. Res. Part A: Policy Pract. 45 (2), 138–147.
Schwartz, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., Lehman, D.R., 2002. Maximizing versus satisficing: happiness is a matter of choice. J. Pers. Soc. Psychol. 83 (5), 1178–1197.
K.A. Small, R. Noland, X. Chu, D. Lewis. Valuation of travel-time savings and predictibility in congested conditions for highway user-cost estimation. National Cooperative Highway Research Program. 1999.
Small, K.A., Winston, C., Yan, J., 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. Econometrica 73 (4), 1367–1382.
Soman, D., 2001. The mental accounting of sunk time costs: why time is not like money. J. Behav. Dec. Making 14 (3), 169–185.
Swierstra, A.B., van Nes, R., Molin, E.J.E., 2017. Modelling travel time reliability in public transport route choice behaviour. Eur. J. Transp. Infrastruct. Res. 17.
Thaler, R., 1980. Toward a positive theory of consumer choice. J. Econ. Behav. Organ. 1 (1), 39–60.
Tilahun, N.Y., Levinson, D.M., 2010. A moment of time: reliability in route choice using stated preference. J. Intell. Transp. Syst. 14 (3), 179–187.
Tversky, A., Kahneman, D., 1981. The framing of decisions and the psychology of choice. Science 211 (4481), 453–458.
Wakker, P.P., 2010. Prospect Theory: For Risk and Ambiguity. Cambridge University Press.
Watkins, K.E., Ferris, B., Borning, A., Rutherford, G.S., Layton, D., 2011. Where Is My Bus? Impact of mobile real-time information on the perceived and actual wait time of transit riders. Transp. Res. Part A: Policy Pract. 45 (8), 839–848.
Weber, E.U., Milliman, R.A., 1997. Perceived risk attitudes: relating risk perception to risky choice. Manage. Sci. 43 (2), 123–144.
Winkler, R.L., Muñoz, J., Cervera, J.L., Bernardo, J.M., Blattenberger, G., Kadane, J.B., Lindley, D.V., Murphy, A.H., Oliver, R.M., Ríos-Insua, D., 1996. Scoring rules and the evaluation of probabilities. Test 5 (1), 1–60.
M. Yap, O. Cats, B. van Arem. Crowding valuation in urban tram and bus transportation based on smart card data. Transp. A: Transp. Sci. 1-20. 2018.