Available online at www.sciencedirect.com

ScienceDirect

Transportation Research Procedia 97 (2026) 744–753

Transportation Research Procedia

www.elsevier.com/locate/procedia

13th International Conference on Transport Survey Methods

Reliability of Public Transportation Systems with Autonomous First-/Last-Mile Services in a Stated Mode Choice Experiment

Christian Klinkhardt[a,*], Kim Kandler[a], Martin Kagerbauer[a], Peter Vortisch[a]

[a]*Institute for Transport Studies at Karlsruher Institute of Technology, Kaiserstr. 12, 76131 Karlsruhe, Germany*

Abstract

To achieve a modal shift away from private car use, public transportation must become a more attractive alternative, particularly with regard to reliability. While travel time and cost have been extensively studied as key determinants of mode choice, the impact of travel time reliability has received increasing attention only in recent years. However, quantitative surveys addressing this issue remain scarce, and most existing evidence is based on stated choice studies with highly heterogeneous results. One reason for this inconsistency is the complexity of representing reliability, as it can only be described through a distribution. Many forms of representation appear to overwhelm respondents. This study therefore seeks to bridge this gap by conducting a stated choice experiment with an improved visualization of reliability, presenting travel time and five equally likely delay times as bar charts and incorporating realistic choice situations that account for both trip purpose and transported luggage.

The survey, conducted with a net sample of 859 respondents, reveals that travel time reliability is a decisive factor influencing mode choice. Delays in public transportation are particularly disfavored, with the mean delay rated 3.7 times more negatively than travel time itself. This valuation significantly exceeds the ratios applied in existing cost-benefit analyses, such as those used in German transport infrastructure planning. These findings suggest that the perceived importance of reliability is substantially higher than previously assumed. For on-demand services, the study indicates that displaying slightly longer expected travel times with early arrivals is preferable to promising shorter times that result in delays. These insights underscore the need for policymakers and transport planners to prioritize reliability improvements in public transportation planning and investments.

Keywords: travel time reliability; mode choice experiment; public transportation; autonomous first-/last-mile service

1. Introduction

Due to its environmental and climate benefits, public transportation is expected to play an essential role in our transportation system. The long-term objective is to reduce private car traffic by encouraging increased public transport usage. To accomplish the modal shift, public transportation needs to be an appealing alternative to private vehicles within urban and rural areas. In addition to travel time, independence, convenience, and travel costs, travel time

* Corresponding author Christian Klinkhardt Tel.: +49 721 608 47782.
E-mail address: christian.klinkhardt@kit.edu

10.1016/j.trpro.2026.04.212

reliability is a crucial factor influencing mode choice [12, 23]. While the influence of travel time and travel costs on mode choice has been extensively studied, the impact of travel time reliability has only gained focus in recent years [1]. However, there is still a lack of quantitative surveys in particular [1]. This gap needs to be addressed to obtain an improved understanding of the effects of reliability on mode choice and to determine the necessary degree of reliability of public transportation to facilitate the desired modal shift. Therefore, we designed and conducted a new survey with an improved reliability display in a stated mode choice experiment. Considering the results in decision-making processes, e.g., when allocating funding or carrying out cost-benefit analyses may lead to a more demand-orientated public transportation system. Reliability becomes even more critical in public transportation systems with transfers and first-/last-mile services as the uncertainty of the total travel time increases with each transfer. Transport operators assume that such first-/last-mile services will be operated using autonomous shuttles in the future. The survey therefore also focuses on this topic.

In the following sections, the state of the art is presented, focusing on concepts of reliability and its integration into choice modeling. This is followed by the methodology section, detailing the developed concept and the implementation of the conducted survey. Finally, the results of the model estimations are discussed, highlighting the influence of reliability parameters on mode choice.

2. State of the Art

2.1. Concept for Consideration of Reliability of Public Transportation

Several pieces of information are essential for considering reliability in decision-making processes like cost-benefit analyses: (1) the change in reliability due to a specific measure, (2) how travelers' behavior changes in response to changes in reliability, and (3) the monetization of the changes in reliability (reduced travel time and reduced travel time variability) and behavior (e.g., reduced emissions due to fewer vehicle kilometers traveled in private cars) [15]. The first two aspects have received limited research attention so far [16].

While the first aspect falls beyond the scope of this paper, the second aspect of behavior change is the focus of this study. Travel demand models are typical tools for addressing the second aspect. However, most current travel demand models do not account for the reliability of public transportation [22]. It is generally assumed that travel times by public transportation correspond to its schedule. Consequently, public transportation is often presented too positively in terms of travel times, and the models cannot illustrate the effects of improved reliability. Nevertheless, several studies examine the effects of reliability. Most of these focus on route choice behavior in road traffic, although some address route choices in public transportation [26]. Only a few studies address reliability in the context of mode choice. Besides stated choice surveys, some studies have explored passive data collection methods, such as those utilizing smartcard systems [25]. However, the analysis potential of such data in mode choice models is limited since individuals remain anonymous, and alternative modes of transport, sociodemographic characteristics, or trip purposes are not recorded.

The third aspect, monetization, has been studied before [16]. Research on the monetization of reliability gains focuses on the value of travel time reliability (VOR), which represents the monetary value associated with a change in a unit of reliability [16]. Various theoretical concepts are used to describe this unit of reliability. According to [16], mean-dispersion models (also called mean-variance models) and scheduling models are the most common. Mean-dispersion models consider the average travel time and a measure of deviation from it, while scheduling models focus on the expected arrival time and the time of early or late arrival. Additional models and parameters are discussed in a literature review [18]. For further analyses, estimated parameters can be used for the calculation of the VOR and the Reliability Ratio (RR). To achieve this, a wide range of surveys has been conducted, with stated choice experiments being a commonly used method [4]. There is a summary of VOR results for road transport [15]. In general, most of this research has focused on road transport rather than public transportation [4]. However, a few studies on VOR for public transportation and road traffic are available, for example, in Germany and the Netherlands [7, 21]. Initial studies also examine pooled vehicles [2]. Other researchers emphasize the effects of transfers and use a buffer time metric [6].

Car
travel cost 6€
travel time without traffic 40 min
Equally probable actual travel times [min]
44
45
43
42
48

Train with walking
travel cost 4€
headway 60 min
scheduled travel time 60
Equally probable actual travel times [min]
82
64
61
62
67

Train with shuttle on demand
Travel Cost 6€
headway 10 min
pre-booking time 10 min
scheduled travel time 65
Equally probable actual travel times [min]
75
72
74
65
68

Train with scheduled shuttle
Travel Cost 1€
headway 15 min
scheduled travel time 35
Equally probable actual travel times [min]
41
38
36
44
36

Fig. 1. Typical representation of mode choice with reliability

2.2. Mode Choice Experiment on Reliability

Studies employing the stated choice approach to estimate a VOR aim to present reliability descriptions of alternatives in a way that is most understandable to respondents. The difficulty lies in the fact that travel time reliability cannot be described with a single value but requires the probability distribution of different travel times. This complicates the definition of reliability in revealed preference surveys and its representation in stated choice experiments. Several methods have been employed to provide the necessary information to respondents, including graphs, statistics in tables, and textual descriptions. Among these methods, stated choice experiments performed best when presenting average travel time and five equally likely travel times per alternative in a tabular format, as illustrated in [28]. An example is given in figure 1. This approach has become standard practice [16]. Recent research has further highlighted the importance of uncertainty in the presentation and valuation of travel time reliability in stated choice experiments. A study by [19] demonstrates how uncertainty can be quantified and effectively communicated using Bayesian inference. It also introduces a representation of travel time uncertainty through quantile dotplots, which respondents found both helpful and understandable.

In a large value of time and reliability study for German infrastructure planning, stated choice scenarios with a similar design were used to evaluate the impact of unreliable travel times [7]. The authors found that one minute of average delay in public transportation is perceived as negatively as one minute of travel time. This finding contradicts several other surveys on the subject [3]. The authors suspect that respondents may underestimate the risk of delay occurrence in stated choice scenarios.

However, stated choice experiments with all necessary information do not automatically reflect real-world decision-making situations where individuals choose between train, car, or future autonomous buses in combination with trains. In reality, decisions are often based on simplistic heuristics rather than complete information [24].

In summary, the influence of reliability identified in previous studies varies greatly depending on the method, reliability model, and presentation format in choice experiments. Moreover, passive data can only provide valid results for specific questions since they always capture a limited view of reality and lack sociodemographic information. Thus, there is a need to enhance the realism of stated choice surveys for estimating mode choice models to improve behavior forecasting. This is especially relevant for agents' mode choice when public transportation is considered in detail, such as regional transport with first-/last-mile solutions, for which no large-scale surveys are yet available. This study addresses this gap by developing a particularly realistic representation of reliability.

3. Methodology

3.1. Concept for representation in survey

This paper proposes an improved representation of reliability in a stated choice experiment to enable a more realistic evaluation of the risk of delays by respondents. A second focus of this study is the consideration of autonomous buses as a first/last-mile solution, where reliability plays a critical role in achieving user acceptance. The goal is to estimate parameters for travel time and various reliability metrics in combination with other trip attributes and sociodemographic influences.

The experiment specifically targets mode choice for medium distances (10 km to 50 km). Two primary reasons support this focus: First, according to the survey ”Mobility in Germany 2017” [14], trips within this distance range account for 44 percent of total person-kilometers in Germany. Second, trips involving first/last-mile solutions are generally only practical for distances above a certain threshold, as shorter distances are less attractive due to the additional transfer penalty. In contrast, long-distance travel is typically driven by different purposes and follows distinct decision-making logics.

Within these medium distances, the private car emerges as the primary realistic alternative to public transportation. Consequently, we present four options for selection: Car, train with walking as the first mile, train with autonomous bus on-demand for the first mile, and train with scheduled autonomous bus for the first mile.

To create realistic stated choice scenarios, respondents are first asked about trips they have already taken for specific purposes within the medium-distance range. These responses are used to ensure that participants are only presented with situations they are familiar with. The typical amount of luggage carried on these trips is also queried. In the experiment, trip purpose and the amount of luggage are explicitly incorporated into the hypothetical choice scenarios, as both factors are expected to have a significant influence on mode choice. The considered trip purposes include work, business travel, education, leisure in the evening, leisure during the day, shopping for daily needs, and shopping for other needs.

To create realistic choice situations, typical travel times, access times, costs, and delay times were defined as attribute levels for distances ranging from 10 km to 50 km. For this purpose, times for specific routes in Baden-Wuerttemberg (Germany) were researched, and travel times were selected using Google Maps and public transportation schedules. The selected attribute levels are presented in table 1.

Table 1. Attribute levels used in the stated choice questions

Attribute	Car	Train with Walking	Train with Shuttle SH	Train with Shuttle OD
travel time without delay [min]	30,40,50	30,40,50,60	30,40,50,60	30,40,50,60
travel cost [€]	2,4,6,8	1,2,4,6	1,2,4,6	1,2,4,6
access foot [min]		6,10,15		
access shuttle [min]			5,8,10	5,8,10
frequency [min]		10,15,30,60	10,15,30,60	10,15,30,60
waiting time [min]				10,20,30,40
delay time [min]	0-17	0-21	0-21	0-21

To optimize this scenario, we created a D-efficient design using the software ngene [5]. This design was chosen to minimize the number of respondents needed while allowing for numerous attribute combinations, ensuring a manageable response burden with only eight situations presented to each respondent. Utilizing a D-efficient design, ngene aims to minimize the D-error, calculated from the determinant of the asymptotic variance-covariance matrix. We experimented with varying numbers of question blocks and determined four blocks to be most suitable for minimizing errors.

To identify the most suitable representation of travel times, concrete examples were discussed in focus groups with both students and non-experts. This included visual formats from [28] as well as newly developed representations based on a timeline concept, where the length of individual bars indicates the duration of each travel time component for a given alternative. This timeline-based approach has already been successfully applied in various surveys and has yielded models with substantial explanatory power [17, 20]. We received extensive qualitative feedback that guided the final choice of representation. Based on this feedback, we decided to use the timeline-based bar chart representation, where travel time reliability is visualized by five delay bars of equal probability (Figure 2). Respondents were instructed to imagine that these delays had actually occurred on the previous five days for the respective alternative. This framing was intended to convey the idea that each of these delays could recur with equal probability in the future. The five actual travel times reflect the mean delay for each alternative as generated by Ngene. Different value distributions were applied: in some cases, all five delays were similar in duration and close to the mean; in other cases, there were substantial outliers with either significantly higher or lower delay values. This design allows for testing whether

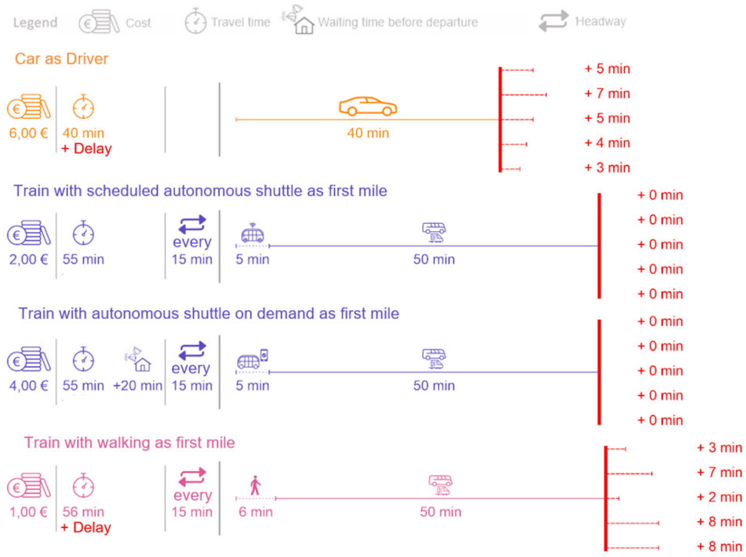

Fig. 2. Representation of the unreliability of the alternatives in the survey (translated from German)

not only the mean but also the distribution and variability of delays influence mode choice behavior. To enhance clarity, respondents were first shown a choice situation with travel times only, without any reliability information. In a second step, the same choice was presented with additional reliability information. To reduce respondent burden, constraints were applied to ensure that not all transportation modes experienced delays simultaneously.

The graphical presentation of travel times via bars is designed to allow for a more realistic assessment of reliability compared to previously used formats, such as tables or dot plots. In addition, other alternative-specific attributes such as costs, headway, and pre-booking time (for the autonomous on-demand bus) are presented.

3.2. Survey conduction

The survey was conducted as part of the RABus project, which focuses on the integration of autonomous shuttles into public transportation systems. The survey took place from December 2023 to March 2024 and was carried out exclusively online. It was distributed via social media channels and a paid access panel, targeting the population of Baden-Wuerttemberg (Germany), with a particular focus on the regions around the cities of Mannheim and Friedrichshafen, as these areas are central to the project.

Mannheim represents a typical large German city (approximately 316,000 inhabitants) with a well-developed tram and bus network, as well as connections to regional and long-distance rail services. Its surrounding area is categorized as an urban region. Friedrichshafen, on the other hand, is a smaller city in a rural area, with approximately 63,000 inhabitants. It has a bus network and is primarily connected via regional rail services, with only occasional access to long-distance trains. In the surrounding area, regional rail stops are sparse, and public transportation relies mostly on buses. In this region, the integration of autonomous buses into the public transportation network could significantly improve access to railway stations.

The survey link, which was distributed both via social media and through the panel, was opened 2,134 times. Of these clicks, 1,324 respondents started the survey (62.0%), and 1,053 completed the questionnaire, resulting in a completion rate of 79.5%. After pre-filtering based on completion time and screening for “straightlining” (i.e., when a respondent consistently selects the same answer option, such as always choosing the first), the final net sample consisted of 859 respondents. This yields 6,872 choice situations available for analysis. The sample includes 724 respondents recruited via the panel and 141 respondents recruited through social media. For the entire questionnaire, the average completion time was 23 minutes and 16 seconds, while the median completion time was 17 minutes and 21 seconds.

Table 2. Survey sample composition compared to MiD 2017 (weighted results for Baden-Wuerttemberg)

Attribute	Value	Sample	MiD 2017
gender	female	51.4 %	50.5 %
age	< 18	0 %	17.0 %
	18-29	13.0 %	15.1 % (18.2 %)*
	30-49	33.0 %	26.7 % (32.2 %)*
	50-69	45.7 %	26.7 % (32.2 %)*
	70+	8.3 %	14.5 % (17.4 %)*
net income of household [EUR]	< 900	2.7 %	3.3 %
	900-2000	19.9 %	18.3 %
	2000-3000	23.7 %	29.7 %
	3000-4000	20.0 %	19.8 %
	4000-5000	15.0 %	14.0 %
	5000-6000	8.7 %	8.4 %
	6000-7000	4.6 %	4.1 %
	> 7000	5.4 %	2.4 %
location is metropolis	urban region	61.1 %	62.6 %
	rural region	38.9 %	37.4 %
no. of cars in household	0	19.3 %	17.8 %
	1	47.1 %	53.6 %
	2	27.4 %	24.4 %
	3 or more	6.2 %	4.2 %
bike accessibility in household	yes	66.4 %	74.3 %
public transportation pass	yes	30.6 %	15.4 %

* Shares of the age group in the sample and the population when only the adult population is considered

Table 2 presents a summary of the socio-demographic composition and access to various mobility tools of the sample. The information is compared to the results of “Mobilität in Deutschland” (MiD) by [14], a nationwide study on mobility and travel behavior. In this context, only the weighted results for the state of Baden-Wuerttemberg were compared to the key data of our sample because the targeted sample was also the population of this region.

The comparison between the survey sample and the MiD 2017 data for Baden-Wuerttemberg shows some notable differences. Respondents under 18 years old are excluded from the access panel, which explains their absence in the sample. Consequently, the percentage distribution across age groups is considered excluding those under 18. It is noticeable that the 18-29 age group is slightly underrepresented (13.0% in the sample compared to 18.2% in MiD), as well as those over 70 (8.3% compared to 17.4%). In contrast, the age groups between 30 and 69 are overrepresented in the sample.

The full range of income groups is represented in the sample, with slight deviations from the MiD data. Bicycle ownership is underrepresented in the sample (66.4% vs. 74.3% in MiD), while the share of public transportation pass holders is significantly higher (30.6% vs. 15.4%). This overrepresentation of public transportation pass holders can be attributed to the distribution of the survey through social media channels, particularly by transit providers.

In the choice experiment, all alternatives were selected by respondents. Overall, the car was chosen in 49% of all situations (56.65% when available). The train with walking as access accounted for 20%, the train with scheduled shuttle as access for 16%, and the train with on-demand shuttle as access for 15%. After excluding interruption-related outliers, the average response time was 23 seconds per choice situation.

In the analysis, we further examined how often respondents changed their mode choice between the situation without reliability information and the situation with reliability information. The results indicate a high degree of stability, as a change occurred in only 22% of cases. This is, however, considered plausible.

4. Model Results

To process the results, the open-source package Apollo in the R programming language was used [13]. Using this package, various multinomial logit models were estimated and compared. Since multinomial logit models have been the state-of-the-art for a long time, we refer the reader to [27] for details and an overview of generalized extreme value (GEV) models. Several model structures were tested and evaluated based on their log-likelihood and Bayesian Information Criterion (BIC).

A particular focus was placed on the different formulations for delays and their influence on the utility function. The five equally likely travel times need to be aggregated into indicators that can be incorporated into the model estimation. There are several possible approaches to generating such indicators from individual delays. The mean–dispersion model, the scheduling model, and several quantile-based formulations were tested. Based on both log-likelihood and BIC, the best-performing specification applied the mean–dispersion model (1) for car alternatives, while for public transport alternatives an adapted scheduling model (2) provided the superior fit.

The distinction between the model for the car alternative and the model for the public transportation alternative can be explained by the differing perceptions of travel time in private car use versus public transport. One hypothesis is that certain regular delays in car traffic are perceived as ”normal” and thus more acceptable, whereas significant variability in travel time is viewed negatively. In contrast, regular delays in public transportation appear less tolerable. Instead, the mean delay is judged more critically by users. Additionally, for certain trip purposes another negative component arises when there are significant outliers in delays that are substantially longer than the average travel time.

$$U_{mean\_dispersion\_model} = \ldots + \beta_{tt} \cdot TT_{mean} + \beta_d \cdot D \tag{1}$$

$$U_{scheduling\_model} = \ldots + \beta_{tt\_s} \cdot TT_S + \beta_{tt\_dm} \cdot TT_{DM} + \beta_{tt\_dq\_trip\_purpose} \cdot (TT_{DQ} - TT_{DM}) \tag{2}$$

with:

- [...]: Further attributes and estimates
- β_{tt}: Estimate for mean travel time
- TT_{mean}: Mean travel time
- β_d: Estimate for dispersion
- D: Dispersion (standard deviation)
- β_{tt_s}: Estimate for scheduled travel time
- TT_S: Scheduled travel time
- β_{tt_dm}: Estimate for mean delay
- $\beta_{tt\_dq\_trip\_purpose}$: Estimate for 90th-quantile of delay for specific trip purpose
- TT_{DM}: Mean delay
- TT_{DQ}: 90th-quantile of delay

In the following analysis, various demographic influences and the impact of trip purposes were examined. First, the effects were tested for each alternative specifically, and in the second step, their impact on the delay parameters was evaluated. Various factors influencing the delay parameters were examined, including trip purpose, employment status, luggage, public transport season tickets, and car ownership. Non-significant parameters were subsequently removed in each step. The only exception are the alternative-specific constants, which were retained even if they were not statistically significant. The estimates for the alternative ”Car” were fixed to 0. Additionally, an analysis was conducted to determine the extent to which the Bayesian Information Criterion (BIC) could be reduced by combining parameters with similar values. The results indicate that cost, time, and delay estimates for the three public transportation alternatives should be the same.

The remaining estimates are presented in Table 3.

These results provide valuable insights into how different factors, such as travel time, costs, delays, and sociodemographic characteristics, influence mode choice in the context of travel time reliability. The following sections analyze key findings, with a particular focus on delay parameters, their relative importance compared to travel time, and variations across transportation modes.

4.1. Travel Time

Travel time in a car is perceived more negatively (-0.038) compared to travel time in public transportation alternatives (-0.030). However, access times, particularly walking to the train (-0.108), are rated even more negatively. This aligns with expectations, as sitting comfortably in a vehicle is generally perceived as less burdensome than physically walking to access the service.

Access times for shuttles, both scheduled and on-demand (-0.072), are also rated negatively but less severely compared to walking. This indicates that while shuttles are seen as an improvement for first-/last-mile connections, their additional access times still detract from their overall attractiveness.

Waiting times for on-demand shuttles (-0.012) are perceived as the least negative among all time components. This result likely reflects the comfort associated with waiting at home, where users experience fewer inconveniences compared to waiting at stations or stops. This distinction underscores the potential advantage of on-demand services in improving perceived service quality.

Table 3. Model estimates for mode choice experiment (only statistically significant values are presented)

	Car	Train with Walking	Train with Scheduled Shuttle	Train with Shuttle On Demand
Travel Time				
Travel Time [min]	-0.038 (-6.62)***	-0.030 (-10.55)***	-0.030 (-10.55)***	-0.030 (-10.55)***
Travel Time Access [min]		-0.108 (-8.67)***	-0.072 (-5.75)***	-0.072 (-5.75)***
Supply Characteristics				
Waiting Time [min]	-0.012 (-3.63)***			
Headway [min]		-0.007 (-4.32)***		
Cost				
Travel Cost [EUR/trip]	-0.090 (-5.73)***	-0.291 (-13.99)***	-0.291 (-13.99)***	-0.291 (-13.99)***
Delay Parameters				
Delay Mean [min]		-0.111 (-10.74)***	-0.111 (-10.74)***	-0.111 (-10.74)***
Delay SD [min]	-0.0504 (-3.22)**			
Trip Purpose Leisure (Night) Delay Mean [min]		0.104 (2.31)**	0.104 (2.31)**	0.104 (2.31)**
Trip Purpose Education,Leisure (Night) Delay 90th-Quantile [min]		-0.100 (-2.43)**	-0.100 (-2.43)**	-0.100 (-2.43)**
Alternative-specific and Trip-specific Constants				
Constant	0.000 (N/A)	1.5354 (3.71)***	0.4637 (1.09)	0.5975 (1.37)
Male	0.000 (N/A)		0.3693 (2.79)**	0.2632 (1.97)*
Car-ownership	0.000 (N/A)	-1.8286 (-5.92)***	-1.4072 (-4.48)***	-1.1334 (-3.51)***
Public transport-Pass owner	0.000 (N/A)	0.7658 (5.22)***	0.5702 (4.04)**	
Lots of luggage	0.000 (N/A)		-0.3472 (-2.05)*	
Trip Purpose Education	0.000 (N/A)	1.3594 (2.66)**	0.958 (1.85)	0.958 (1.85)
Trip Purpose Shopping (Daily)	0.000 (N/A)	-1.1205 (-5.96)***	-0.9969 (-5.63)***	-0.9806 (-5.28)***
Trip Purpose Shopping (Other)	0.000 (N/A)	-0.3424 (-2.20)*	-0.4614 (-2.88)**	-0.4552 (-2.74)**

$p < 0.001^{***}$; $p < 0.01^{**}$; $p < 0.05^{*}$ (Rob.t-ratio(0) in brackets); $AIC = 13841.74$; $BIC = 14053.63$; $\rho^2 = 0.256$; $Adj.\rho^2 = 0.252$; $LL_{Final} = 6889.869$; $LL_0 = -9257.344$

4.2. Travel Cost

Travel costs for public transportation (-0.287) are perceived as approximately three times more negative compared to car costs (-0.090). This discrepancy often arises because car users tend to underestimate or disregard the true costs of car ownership and operation, perceiving the monetary costs of public transportation as relatively higher. Additionally, there is a higher willingness to pay for travel by private car. This phenomenon highlights a challenge for public transportation providers in communicating the cost-efficiency of their services.

4.3. Delays

Delays for public transportation alternatives are negatively weighted in the utility model. The mean delay for public transportation (-0.111) is rated approximately 3.7 times more negatively than travel time. This significant difference highlights the high value respondents place on reliability in the context of scheduled public transportation. The delay quantile parameter for education and leisure (night) trips (-0.100), while smaller, still indicates that rare but extreme delays affect mode choice especially for these trips. However, the mean delay for leisure in the evening is perceived as less problematic, presumably because there are no appointments at fixed times or follow-up events. Other influences on the delay parameters were not significantly verifiable with the available sample.

When comparing the ratio between the mean delay and travel time in this study to other approaches, significant differences emerge. A study conducted as part of the German federal transport infrastructure planning (BVWP) reports a ratio of 1.0 for travel time and mean delay, with the exception of business trips, where a higher ratio of 1.7 is estimated [8]. The "Standardisierte Bewertung," a benefit-cost analysis framework used for public transportation project funding in Germany, applies a ratio of 2.7 [11]. In Switzerland, a stated choice study identified ratios ranging from 1.5 to 3, depending on the trip purpose and distance [10]. In contrast, the ratio of 3.7 found in this study is substantially higher, suggesting that respondents perceive delays as a significantly greater issue. This discrepancy may be attributed to differences in survey methods and the representation of the delays in the form of bars or a heightened sensitivity to reliability issues among the participants in this experiment.

For cars, the standard deviation of delays is associated with a ratio of 1.1 relative to travel time, which exceeds the value of 0.7 reported in the BVWP. These findings indicate that delays, whether in public transportation or car travel, are assigned more weight in this study compared to other contexts. This variation underscores the critical role of survey design and methodology in shaping respondents' perceptions and preferences.

When comparing the delay parameters between public transport and car travel, delays in public transport are consistently perceived negatively, whereas average delays in car travel are evaluated similarly to regular travel time and therefore entail substantially lower disutility. The findings from this experiment highlight the critical importance of delays in modeling mode choice decisions.

These results highlight the need for policymakers and public transport providers to prioritize reliability improvements as part of their service enhancements.

The substantial variation in delay valuation between studies also underscores the importance of survey design and methodology in capturing user preferences. Since these ratios directly influence the allocation of funding for transport infrastructure, precise determination of reliability impacts is essential. This is consistent with other research findings, such as those by Fosgerau, who emphasize the need for more big data sources to better understand the effects of reliability.

4.4. Alternative-specific and trip-specific constants

Alternative-specific constants reveal additional insights into mode choice behavior. public transportation alternatives generally have positive constants, which shift significantly into the negative range when car ownership is present (-1.829 for train with walking). This confirms that car ownership is one of the strongest determinants of mode choice, as it drastically reduces the likelihood of choosing public transport.

Other factors, such as the trip purposes, also show strong influences. Trips for education are positively associated with train alternatives, suggesting that these users may place greater value on sustainable or efficient travel options. Shopping trips are consistently associated with negative utilities across all public transportation alternatives. Although luggage was additionally analyzed as a separate influence and was found to be significant for the train with the scheduled shuttle, the challenges of carrying items on public transport is also likely to have an influence on these parameters. Public transportation pass ownership significantly improves the utility of train alternatives, but has no observable effect on on-demand shuttles. This suggests that on-demand shuttles are not yet perceived as integral part of traditional public transportation systems by these persons.

5. Conclusion and Recommendations

This study highlights the importance of understanding the impact of travel time reliability on mode choice, addressing the insufficient consideration of this factor in current stated choice experiments. A novel approach is proposed, featuring the visualization of travel and delay times and the inclusion of realistic choice situations tailored to respondents, considering trip purpose and luggage. The survey, conducted with 859 respondents, revealed key insights into the perception of modes: delays in public transportation and private vehicles are perceived differently. For public transportation to become more attractive, reliability in travel time is significantly more important than simply reducing travel times. The study found that the mean delays in public transportation are perceived more negative by a factor of 3.7 compared to travel time itself, emphasizing the importance of reliability in achieving a mode shift. This finding underscores the need to account for reliability in benefit-cost analyses, such as the German framework for infrastructure funding, in which the reliability ratio is currently underestimated.

Moreover, when planning on-demand services, it is more effective to present slightly longer travel times to the customer and ensure early arrivals rather than advertising fast travel times followed by delays. Extreme travel time outliers are especially detrimental to user experience in the context of education trips or leisure trips at night, making their prevention a critical priority in these contexts. In practice, this means that for first-/last-mile solutions, it is better to include sufficient buffers to avoid missed connections, which would otherwise lead to disproportionately long delays.

Acknowledgements

This work was carried out within the project RABus (Living Lab for an Automated Bus Service Operating as Part of Public Transportation in Rural and Urban Areas, https://www.projekt-rabus.de), which is funded by the Baden-Wuerttemberg Ministry of Transport, Germany.

References

[1] Alkubati, M.A., Khalifa, N.A., Al-barakani, H.A., 2023. An overview of public transport reliability studies using a bibliometric analysis. Ain Shams Engineering Journal 14, 101908. URL: https://doi.org/10.1016/j.asej.2022.101908.

[2] Alonso-González, M.J., van Oort, N., Cats, O., Hoogendoorn-Lanser, S., Hoogendoorn, S., 2020. Value of time and reliability for urban pooled on-demand services. Transportation Research Part C: Emerging Technologies 115, 102621. URL: https://doi.org/10.1016/j.trc.2020.102621.

[3] Bocksch, R., 2022. Preis und unzuverlässigkeit größtes manko bei zugreisen [digital image]. URL: https://de.statista.com/infografik/28344/einstellung-zu-zugreisen/. last accessed on 03 February 2024.

[4] Carrion, C., Levinson, D., 2012. Value of travel time reliability: A review of current evidence. Transportation Research Part A: Policy and Practice 46, 720–741. URL: https://doi.org/10.1016/j.tra.2012.01.003.

[5] ChoiceMetrics, 2012. Ngene 1.1.1 user manual & reference guide [computer software]. ChoiceMetrics, Australia.

[6] Dixit, M., Brands, T., van Oort, N., Cats, O., Hoogendoorn, S., 2019. Passenger travel time reliability for multimodal public transport journeys. Transportation Research Record: Journal of the Transportation Research Board 2673, 149–160. URL: https://doi.org/10.1177/0361198118825459.

[7] Dubernet, I., Axhausen, K.W., 2020a. The german value of time and value of reliability study: the survey work. Transportation 47, 1477–1513. URL: https://doi.org/10.1007/s11116-019-10052-4.

[8] Dubernet, I., Axhausen, K.W., 2020b. The german value of time and value of reliability study: the survey work. Transportation 47, 1477–1513.

[9] Fosgerau, M., 2016. The Valuation of Travel Time Variability. volume no.2016/04 of *International Transport Forum Discussion Papers*. OECD Publishing, Paris.

[10] Fröhlich, P., Juni 2014. Einfluss der Verlässlichkeit der Verkehrssysteme auf das Verkehrsverhalten: Forschungsprojekt SVI 2010/003. volume 1472. Schweizerischer Verband der Strassen- und Verkehrsfachleute (VSS), Zürich.

[11] GmbH, I.C., GmbH, V.V.I.S., 2022. Entwicklung und Fortschreibung der Berechnungsverfahren zur Bewertung von Verkehrswegeinvestitionen des schienengebundenen ÖPNV. Version 2016+ ed., Intraplan Consult GmbH, München, Stuttgart.

[12] Held, M., 1982. Verkehrsmittelwahl der Verbraucher: Beitrag einer kognitiven Motivationstheorie zur Erklärung der Nutzung alternativer Verkehrsmittel. Wirtschaftspsychologische Schriften der Universitäten München und Augsburg: Vol. 8, Duncker und Humbolt.

[13] Hess, S., Palma, D., 2019. Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. Journal of Choice Modelling 32, 100170. URL: https://doi.org/10.1016/j.jocm.2019.100170.

[14] infas, infas 360 GmbH, I.R.G.D.Z.f.L.u.R.D., 2018. Mobilität in deutschland 2017 - tabellenband. URL: https://www.mobilitaet-in-deutschland.de/.

[15] Jong, G.C.d., Bliemer, M.C., 2015. On including travel time reliability of road traffic in appraisal. Transportation Research Part A: Policy and Practice 73, 80–95. URL: https://doi.org/10.1016/j.tra.2015.01.006.

[16] Jong, G.d., Kouwenhoven, M., 2020. Chapter two - value of travel time and travel time reliability. Advances in Transport Policy and Planning: Standard Transport Appraisal Methods 6, 43–74. URL: https://doi.org/10.1016/bs.atpp.2020.07.010.

[17] Kagerbauer, M., Kostorz, N., Wilkes, G., Dandl, F., Engelhardt, R., Glöckl, U., et al., 2021. Ridepooling in der modellierung des gesamtverkehrs - methodenbericht zur moia begleitforschung.

[18] Kathuria, A., Parida, M., Sekhar, C.R., 2020. A review of service reliability measures for public transportation systems. International Journal of Intelligent Transportation Systems Research 18, 243–255. URL: https://doi.org/10.1007/s13177-019-00195-0.

[19] Kehlbacher, A., Stark, K., Gebhardt, L., 2024. Accounting for uncertainty in the presentation and valuation of travel time in sp experiments, in: hEART 2024 Symposium of the European Association for Research in Transportation. URL: https://elib.dlr.de/206067/.

[20] Klinkhardt, C., Kandler, K., Kostorz, N., Heilig, M., Kagerbauer, M., Vortisch, P., 2023. Integrating autonomous buses as door-to-door and first-/last-mile service into public transport: Findings from a stated choice experiment. Transportation Research Record URL: https://doi.org/10.1177/03611981231175900.

[21] Kouwenhoven, M., Jong, G.C.d., Koster, P., van den Berg, V.A., Verhoef, E.T., Bates, J., Warffemius, P.M., 2014. New values of time and reliability in passenger transport in the netherlands. Research in Transportation Economics 47, 37–49. URL: https://doi.org/10.1016/j.retrec.2014.09.017.

[22] van Oort, N., 2014. Incorporating service reliability in public transport design and performance requirements: International survey results and recommendations. Research in Transportation Economics 48, 92–100. URL: https://doi.org/10.1016/j.retrec.2014.09.036.

[23] Redman, L., Friman, M., Gärling, T., Hartig, T., 2013. Quality attributes of public transport that attract car users: A research review. Transport Policy 25, 119–127. URL: https://doi.org/10.1016/j.tranpol.2012.11.005.

[24] Seebauer, S., Lederer-Hutsteiner, T., Heier, A., Hinterreiter, R., Berger, M., 2010. Info-effect. zielgruppenspezifische wirkungen von multimodalen verkehrsinformationen auf individuelles verkehrsverhalten. URL: https://repositum.tuwien.at/handle/20.500.12708/39950.

[25] Soza-Parra, J., Raveau, S., Muñoz, J.C., 2022. Public transport reliability across preferences, modes, and space. Transportation 49, 621–640. URL: https://doi.org/10.1007/s11116-021-10188-2.

[26] Swierstra, A.B., van Nes, R., Molin, E.J., 2017. Modelling travel time reliability in public transport route choice behaviour. European Journal of Transport and Infrastructure Research 17. URL: https://doi.org/10.18757/EJTIR.2017.17.2.3194.

[27] Train, K.E., 2009. Discrete choice methods with simulation. Cambridge University Press.

[28] 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. Journal of Choice Modelling 2, 8–28. URL: https://doi.org/10.1016/S1755-5345(13)70002-2.