Situational Context and Route Choice in Public Transport: An Exploratory Stated-Choice Study of Travellers' Trade-offs in Germany

Livia Kurbel[1,2,3]

[1]Human Factors, Faculty of Engineering and Technology, Furtwangen University (HFU), Tuttlingen, Germany

[2]In cooperation with the Institute of Transportation Systems, German Aerospace Center (DLR)

[3] Project thesis supervised by Prof. Dr. Stefan Pfeffer (first supervisor, Furtwangen University) and Dr. Esther Johanna Bosch (second supervisor, DLR)

ABSTRACT

Public transport (PT) is central to a sustainable mobility transition, yet attracting travellers requires a detailed understanding of how they choose between alternative routes. Route choice is shaped by attributes such as in-vehicle travel time, the number of transfers, on-board crowding, and walking distance, but most evidence stems from stable, everyday conditions and from countries other than Germany. This study examines how four route attributes: walking time, crowding, travel time, and the number of transfers, are valued, and how their valuation shifts across three situational travel contexts: a relaxed baseline, time pressure, and extreme weather (heat or heavy rain). An exploratory stated-choice experiment was conducted with 99 respondents, each completing 17 choice tasks among three routes, in cooperation with the German Aerospace Center (DLR). Choices were analysed with conditional logit, mixed logit, and context- and sociodemographic-interaction models. Across all contexts, transfers carried by far the strongest penalty, equivalent to roughly 14.7 minutes of travel time per transfer, followed by travel time and walking, whereas crowding was, on average, weak and statistically non-significant. Mixed logit indicated substantial preference heterogeneity, especially for crowding. Attribute valuation was clearly context-dependent: under heat, sensitivity to crowding and walking increased markedly, while sociodemographic moderators were largely non-significant. The findings suggest that situational context, particularly thermal stress, reshapes comfort-related preferences and should inform demand modelling and resilient service design.

Keywords: public transport, route choice, stated-choice experiment, discrete choice modelling, crowding, context effects, travel behaviour

INTRODUCTION

The transport sector is a central field of action for climate policy, because a substantial shift from private cars to public transport (PT) is needed to reduce its greenhouse-gas emissions (Mocanu et al., 2021). Achieving such a shift presupposes a systematic understanding of how passengers decide to use PT, particularly at the level of concrete route choices. Route choice is the process by which travellers select one connection from the set of available alternatives between an origin and a destination. Understanding this process is essential for service planning, passenger-information systems, and demand forecasting in PT networks.

Decades of research have identified a stable set of attributes that govern PT route choice. Travellers consistently prefer routes with shorter in-vehicle travel time (Anderson et al., 2017). The number of transfers is among the strongest deterrents in route evaluation (Raveau et al., 2014). High on-board crowding reduces the perceived utility of a route (Raveau et al., 2011). Access and egress walking is typically valued more negatively than equivalent in-vehicle time (Berggren et al., 2022). Waiting-time uncertainty further shapes route preferences, as travellers are willing to sacrifice travel time for greater certainty (Shelat et al., 2021).

Most route-choice studies, however, estimate a single, context-independent set of attribute weights. Yet contextual and individual factors have been shown to influence how travellers choose their routes (Grison et al., 2017). Disruptions, in particular, weaken the assumptions underlying standard route-choice models. Under network disturbances, passengers often retain their habitual route because they lack timely information about better alternatives (Marra & Corman, 2023). Capacity constraints can lead to denied boarding and cascading network effects that static models do not capture (Cats & Jenelius, 2015). When real-time information is missing, travellers cannot maximise utility and frequently remain in inefficient waiting states (Leng & Corman, 2020). Situational factors such as health risks can likewise shift preferences away from travel time toward other concerns (Chen et al., 2024).

For the German PT context, an empirical research gap remains. Most existing route-choice studies were conducted in other national contexts and transfer only partially to German infrastructural and usage conditions. German studies have analysed the mode-shift potential of PT but have not examined the behavioural factors that guide route choice within the system, nor how these factors combine with the sociodemographic characteristics of travellers (Mocanu et al., 2021). The present project addresses this gap through an exploratory stated-choice experiment conducted in cooperation with the Institute of Transportation Systems of the German Aerospace Center (DLR). It quantifies how four route attributes, walking time, crowding, travel time, and the number of transfers, influence route choice, and how their valuation changes across three situational travel contexts.

The three contexts are deliberately chosen to span everyday and disruptive travel situations. The relaxed and the time-pressure contexts represent everyday states in which travellers are, respectively, unhurried or under schedule pressure. The extreme-weather context represents weather-related disturbances that act as situational stressors on the trip, and it was presented in two variants, heat and heavy rain. Framing adverse weather as a disturbance links the study to research on PT resilience, in which such conditions are known to reshape passenger behaviour (van der Hurk et al., 2015). Examining these contexts within one design therefore allows the resilience of attribute preferences to be assessed, that is, whether the relative importance of walking, crowding, travel time, and transfers remains stable when the travel situation deteriorates.

Three research questions guide the study:

RQ1: How do walking time, crowding, travel time, and the number of transfers influence PT route choice?

RQ2: How does the valuation of these attributes differ across the three situational travel contexts?

RQ3: Are attribute preferences heterogeneous across travellers, and do sociodemographic characteristics moderate them?

Based on the literature, four hypotheses are derived:

H1: All four attributes reduce route utility, with transfers carrying the strongest penalty, followed by travel time and walking (Raveau et al., 2014).

H2: Crowding reduces utility, but its weight is smaller and more context-dependent than that of time and transfers (Raveau et al., 2011).

H3: Situational context moderates attribute valuation, such that time pressure increases the weight of time-related attributes, whereas adverse weather increases the weight of comfort-related attributes (Grison et al., 2017).

H4: Attribute preferences are heterogeneous across individual travellers (Mo et al., 2023).

The remainder of this paper is organised as follows. Section 2 reviews the theoretical background and the determinants of PT route choice. Section 3 describes the stated-choice experiment and the discrete-choice models. Section 4 reports the results, and Section 5 discusses them and draws conclusions.

THEORETICAL BACKGROUND

This section reviews the theoretical foundations of public-transport route choice, the attributes that determine it, the role of situational context, and the discrete-choice methods on which the study builds.

Route Choice as a Discrete-Choice Problem

PT route choice is commonly modelled within the framework of discrete choice based on random utility maximisation (McFadden, 1974). The framework assumes that a traveller assigns a utility to each alternative route and selects the one with the highest utility (Train, 2009). The utility of a route comprises a systematic component, determined by observable attributes such as travel time or the number of transfers, and a random component representing unobserved influences (Train, 2009). When the random components are assumed to be independently and identically extreme-value distributed, the choice probabilities take the multinomial, or conditional, logit form (McFadden, 1974). Conditional logit estimates attribute coefficients from the characteristics of the alternatives themselves, which makes it well suited to modelling choices defined by route attributes (Train, 2009). The mixed logit model relaxes the restrictive assumptions of conditional logit and represents coefficients as random parameters, thereby capturing preference heterogeneity between travellers (Train, 2009). Stated-choice experiments operationalise this framework by presenting respondents with hypothetical alternatives whose attributes are varied systematically (Louviere et al., 2000).

Determinants of Public Transport Route Choice

Time-related attributes are the most consistently documented determinants of route choice. In a multimodal route-choice study of the Greater Copenhagen Area, in-vehicle travel time was identified as a primary criterion for screening out clearly unattractive alternatives (Anderson et al., 2017). In the same network, a single transfer is equated with roughly twelve to eighteen minutes of in-vehicle time, depending on trip purpose (Anderson et al., 2017). Service frequency and headway additionally shape route choice in mixed schedule- and frequency-based systems (Eltved et al., 2018).

Transfers constitute one of the strongest negative factors in route evaluation. The perceived burden of transferring rises disproportionately beyond the first transfer, which defines a perceived transfer threshold (Jin et al., 2017). The location of a transfer relative to the direct path further affects travellers' route and mode choices (Chia et al., 2020). Amenities at interchange points, such as shops or cafés, can reduce the perceived inconvenience of transferring (Berggren et al., 2022).

On-board crowding is a key comfort attribute. High passenger density substantially reduces the perceived utility of a route (Raveau et al., 2011). Evidence from studies on dense urban metro networks, such as the London Underground, shows that travellers are willing to forgo almost half a minute of in-vehicle time per station to increase their chance of obtaining a seat (Raveau et al., 2014). The risk of being left behind because of crowding is valued as an additional time penalty (Mo et al., 2023). Since the COVID-19 pandemic, passengers appear to value crowding more negatively than before (Yap et al., 2023).

Walking, reliability, and network topology provide further determinants. Access and egress time is weighted around two to three times more negatively than in-vehicle time (Anderson et al., 2017). Travellers are willing to sacrifice more than seven minutes of travel time to obtain certainty about their waiting time (Shelat et al., 2021). Reliability is, moreover, valued differently across travellers, modes, and space (Soza-Parra et al., 2022). Travellers also prefer routes that progress directly toward the destination on the network map (Raveau et al., 2011).

Finally, traveller characteristics moderate route choice. Travellers who habitually use multimodal routes and hold a positive attitude towards PT are the most inclined to choose routes that include transfers (Grison et al., 2017). In Budapest, sociodemographic factors such as gender, income, and car ownership shape transport mode choice (Hamadneh & Jaber, 2023). Commute patterns and individual circumstances jointly shape transit route choice (Eluru et al., 2012).

Situational Context, Disruptions, and Resilience

Beyond stable attributes, the situation in which a trip takes place can change how those attributes are weighed. Contextual factors, including trip purpose and the traveller's state, influence route choice in PT (Grison et al., 2017). This becomes most visible under disruptions, where the assumptions of standard models break down. Passengers frequently keep their habitual route because they lack information about better alternatives during disturbances (Marra & Corman, 2023). Partial capacity degradation increases network vulnerability well before a complete failure occurs (Cats & Jenelius, 2018). Reserve capacity is therefore valuable for the robustness of PT networks (Cats & Jenelius, 2015). The availability of real-time information strongly affects how passengers respond to disruptions (Leng & Corman, 2020). Smart-card analyses show that passengers react to rail disruptions in heterogeneous and often inefficient ways (Mo et al., 2022). During extreme weather, reduced schedules shrink the available choice set so that reachability outweighs time savings (van der Hurk et al., 2015). Delayed information can cause travellers to miss detour opportunities once they have passed a decision point (Cong et al., 2022). Situational health risks can similarly shift preferences away from travel time toward hygiene and comfort (Chen et al., 2024). Together, these findings motivate treating relaxed conditions, time pressure, and extreme weather (heat or heavy rain) as distinct travel contexts that may reshape attribute valuation.

The Present Study

The present study builds on this framework with an exploratory stated-choice experiment. Stated-choice methods allow the analyst to control attribute levels and to present situational contexts that are difficult to observe in revealed-preference data (Louviere et al., 2000). Conditional logit is used as the primary model to quantify the average valuation of the four attributes (McFadden, 1974). Mixed logit is added as a robustness check to detect preference heterogeneity (Train, 2009). Context effects are modelled through interactions between the attributes and the travel contexts. All models are estimated with the Apollo package for choice modelling in R (Hess & Palma, 2019).

METHODS

This section describes the two-step research design, the stated-choice experiment with its attributes and contexts, the sample and procedure, the discrete-choice models, and the checks used to evaluate data quality.

Research Design

The study followed a two-step exploratory design. In a first step, a systematic review of the route-choice literature was conducted to identify the factors that influence route choice in PT. In a second step, the most relevant of these factors were tested in a stated-choice experiment with travellers in Germany. The experiment used a within-subject design in which every respondent completed all choice tasks.

Systematic Literature Review

The first step served to ground the experimental attributes in the existing evidence base. A systematic review of the peer-reviewed transport-research literature was carried out to identify the factors that consistently influence PT route choice. Sources were selected with attention to their peer-review status, the relevance of the publishing journal, and the recency of the findings. The review confirmed travel time, transfers, crowding, and access and egress walking as the most consistently documented determinants of route choice. In-vehicle travel time was identified as a primary criterion for screening out unattractive route alternatives (Anderson et al., 2017). The number of transfers emerged as one of the strongest deterrents, with a disproportionate increase in burden beyond the first transfer (Jin et al., 2017). On-board crowding was identified as a central comfort attribute that lowers the perceived utility of a route (Raveau et al., 2011). Access and egress walking was found to be valued more negatively than equivalent in-vehicle time (Berggren et al., 2022). On the basis of this review, four attributes, walking time, crowding, travel time, and the number of transfers, were selected for the experiment. These attributes were chosen because they are behaviourally relevant and can be communicated to respondents without technical knowledge.

Choice Tasks and Attributes

In each choice task, respondents selected their preferred route from three alternatives labelled A, B, and C. Each route was described by the four attributes derived from the literature review: walking time, crowding, travel time, and the number of transfers. Walking time was expressed in minutes and varied between 0 and 15 minutes. Crowding was operationalised as a three-level categorical attribute ranging from low to high passenger density. Travel time was expressed in minutes and varied between 10 and 30 minutes. The number of transfers varied between zero and three. Within a single task, two of the four attributes were varied while the remaining two were held constant, so that the pairwise trade-offs between attributes were presented clearly. All three routes were available in every task.

Situational Travel Contexts

Each choice task was embedded in one of three situational travel contexts, presented as a short standardised vignette before the routes. The relaxed context described an unhurried trip home. The time-pressure context described an important work appointment for which arriving late would have negative consequences. The extreme-weather context described travelling under adverse weather and was presented in two variants, one describing hot weather (heat) and one describing heavy rain. In the experimental design, the relaxed and time-pressure contexts each comprised six choice tasks, and the extreme-weather context also comprised six, split evenly between the heat and rain variants. After data preparation, five relaxed tasks entered the analysis, so each respondent contributed 17 tasks. In the analysis, the heat and rain variants were treated as separate conditions, so that the context model reports four sets of coefficients. Tasks from the different contexts were presented in randomised order to control for order and learning effects (Louviere et al., 2000). Because every respondent completed all contexts, the design was fully within-subject.

Sample and Procedure

The experiment was administered as an online survey to travellers in Germany. The final sample comprised 99 respondents who provided valid choices. Before the choice tasks, respondents reported sociodemographic characteristics, including age, gender, place of residence by settlement size, and the frequency of their PT use. These characteristics were later used to test for systematic moderation of attribute valuation. After data preparation, 17 choice tasks per respondent entered the analysis. The resulting data formed a panel of repeated choices nested within respondents.

Statistical Analysis

Route choices were analysed with discrete-choice models estimated by maximum likelihood in the Apollo package for R (Hess & Palma, 2019). The models follow the random utility framework, in which the utility that respondent *i* derives from route *j* in choice task *t* is divided into a systematic component and a random error (McFadden, 1974):

$$U_{ijt} = V_{ijt} + \varepsilon_{ijt} \qquad (1)$$

In the conditional logit model, the systematic utility is a linear function of the four route attributes:

$$V_{ijt} = \mathrm{ASC}_j + \beta_F \cdot F_{ijt} + \beta_C \cdot C_{ijt} + \beta_D \cdot D_{ijt} + \beta_U \cdot U_{ijt} \qquad (2)$$

Here F, C, D, and U denote the walking time, crowding level, travel time, and number of transfers of route *j*, and ASC_j is its alternative-specific constant, with ASC_A fixed to zero as the reference. Assuming independent and identically type-I extreme-value distributed errors, the probability that respondent *i* chooses route *j* takes the multinomial logit form (McFadden, 1974):

$$P_{ijt} = \exp(V_{ijt}) \, / \, \Sigma_k \exp(V_{ikt}) \qquad (3)$$

Because each respondent completed several tasks, the individual likelihood is the product of the choice probabilities across that respondent's 17 tasks. The mixed logit model relaxes the assumption of homogeneous preferences by treating each attribute coefficient as individual-specific and normally distributed, with mean μ and standard deviation σ, and was simulated using 500 Halton draws (Train, 2009). The context model lets each attribute coefficient differ across the contexts through interactions with context indicators, with the heat and rain variants entered separately, illustrated here for travel time:

$$\beta_D = \beta_{D,EN} \cdot \mathrm{EN} + \beta_{D,ST} \cdot \mathrm{ST} + \beta_{D,HI} \cdot \mathrm{HI} + \beta_{D,RG} \cdot \mathrm{RG} \qquad (4)$$

where EN, ST, HI, and RG are dummy variables for the relaxed, time-pressure, heat, and rain conditions; this model additionally includes interactions of travel time with age and PT-use frequency. The sociodemographic model instead expresses each attribute coefficient as a baseline plus interactions with the person characteristics, as shown in Equation (5):

$$\beta_k = \beta_{k,0} + \beta_{k,age} \cdot age + \beta_{k,ori} \cdot ori + \beta_{k,fem} \cdot fem + \beta_{k,PT} \cdot PT \qquad (5)$$

Here $\beta_{k,0}$ is the baseline coefficient of attribute *k* for the average respondent, with *k* indexing walking, crowding, travel time, and transfers, and *age*, *ori*, *fem*, and *PT* denote the z-standardised age, the z-standardised origin, the female indicator, and the z-standardised PT-use frequency. Because the continuous variables are z-standardised, each interaction coefficient gives the change in the valuation of the attribute per one standard deviation of the characteristic. In addition to these population-level models, individual-level coefficients were obtained for each respondent as the posterior (conditional) mean of the random parameters of the mixed logit, given that respondent's own choices (Train, 2009). These individual coefficients underpin the exploratory analyses of sociodemographic tendencies and preference types reported in Sections 4.6 and 4.8. Finally, attribute coefficients were converted into time equivalents using the delta method, which expresses the disutility of walking, crowding, and transfers in minutes of travel time. All reported standard errors are robust standard errors that account for the repeated choices of each respondent, which yields a conservative basis for the significance tests.

In all tables, statistical significance is reported with asterisks: * denotes $p < .05$, ** denotes $p < .01$, and *** denotes $p < .001$. Estimates without an asterisk are not statistically significant (n.s.).

Data Quality and Instrument Evaluation

Because the instrument is a stated-choice experiment rather than a psychometric scale, the classical test-theory indices of item difficulty, item discrimination, and internal-consistency reliability (Cronbach's alpha) do not apply. These indices presuppose a set of items that measure a single latent construct and are aggregated into a total score, which is not the case for choice tasks that elicit preferences over experimentally varied attributes (Louviere et al., 2000). The quality of a stated-choice instrument is instead judged from the validity of the experimental design and the consistency of the observed choices. The present study therefore evaluated the design for dominated alternatives and the responses for straight-lining and completeness. The reliability of the estimates is reflected in the robust standard errors, the confidence intervals, and the model fit, while the effect sizes are expressed by the attribute coefficients and their time equivalents.

RESULTS

This section reports the data-quality checks, the sample, the attribute valuations from the conditional and mixed logit models, the context-dependent and sociodemographic effects, and the individual-level preference analysis.

Data Quality

Three checks support the quality of the choice data. First, the experimental design contained no dominated alternatives: in all 17 tasks the three routes represented genuine trade-offs between two attributes, so that no task could be solved by selecting an obviously superior option. Second, the responses showed no straight-lining. No respondent chose the same route position in every task, and only one respondent selected a single position in more than 90% of the tasks, with a mean position concentration of 54% against a chance level of 33%. Third, all 99 respondents completed all 17 choice tasks, so there was no item non-response. The aggregate choice shares differed across positions, with route A chosen in 17%, route B in 38%, and route C in 45% of the tasks, and this residual position effect was absorbed by the alternative-specific constants in the models.

Sample

The final sample comprised 99 respondents. Respondents were on average 34.2 years old (SD = 15.0), and their ages ranged from 18 to 75 years. The gender distribution was balanced, with 50 women, 47 men, and two respondents who did not state their gender. Only four respondents reported a mobility restriction. Table 1 summarises the sample characteristics.

Table 1: Sample characteristics (N = 99).

Characteristic	Value
Sample size (N)	99
Age in years, M (SD)	34.2 (15.0)
Age range	18-75
Female	50 (51%)
Male	47 (47%)
Gender not stated	2 (2%)
Mobility restriction (yes)	4 (4%)

Figure 1 shows the distributions of age, origin, and PT-use frequency. The sample skewed young but covered the full age range. Respondents came from all settlement sizes, from metropolitan areas to rural regions. PT-use frequency was bimodal, with a sizeable group of daily users and a sizeable group of infrequent users. Figure 2 reports the distribution of the life-situation indicators, which were dominated by respondents in education or established in their careers.

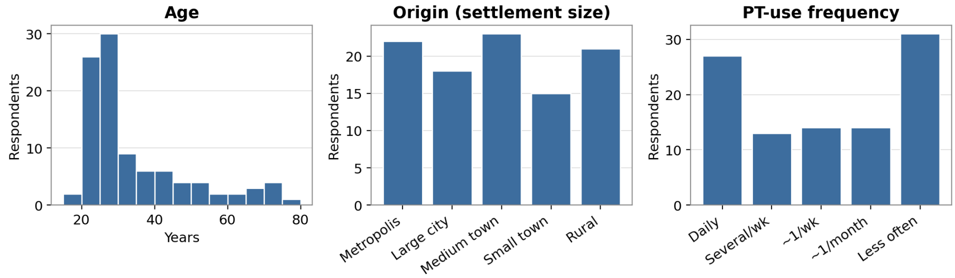

Figure 1: Distributions of respondents' age, origin (settlement size), and public-transport use frequency (N = 99).

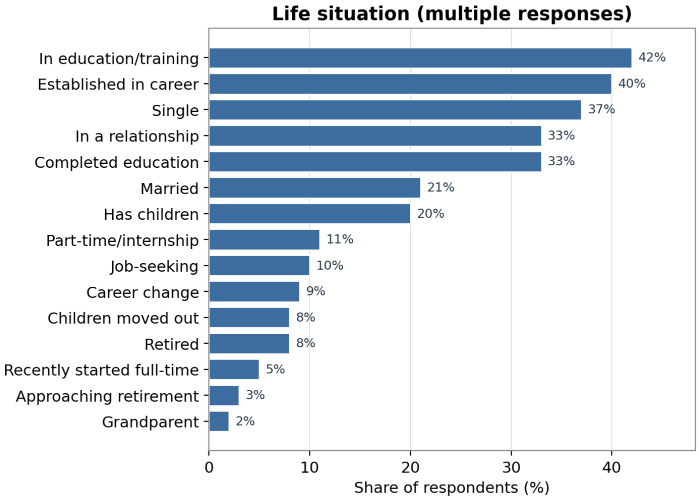

Figure 2: Distribution of respondents' life-situation indicators (multiple responses; N = 99).

Overall Attribute Valuation

Table 2 reports the conditional logit estimates. Three of the four attributes significantly reduced route utility. The number of transfers carried by far the strongest penalty ($b = -1.604$, $p < .001$). Walking time also reduced utility significantly ($b = -0.163$, $p < .001$). Travel time had a smaller but highly significant negative effect ($b = -0.109$, $p < .001$). Crowding was negative but not statistically significant on average ($b = -0.270$, $p = .33$). The alternative-specific constant for route B was significantly negative, which indicates a position effect. Figure 3 visualises these coefficients with their 95% confidence intervals.

Table 2: Conditional logit estimates (robust standard errors).

Parameter	Estimate	Rob. SE	t	Sig.
ASC route B	−0.756	0.166	−4.54	***
ASC route C	−0.094	0.141	−0.67	
Walking	−0.163	0.048	−3.41	***
Crowding	−0.270	0.277	−0.97	
Travel time	−0.109	0.028	−3.88	***
Transfers	−1.604	0.247	−6.50	***

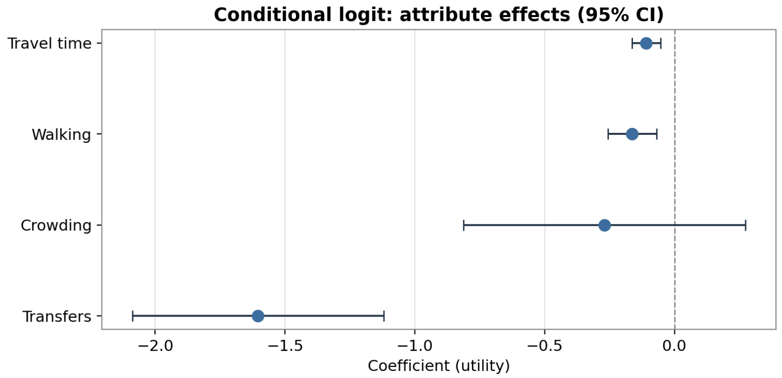

Figure 3: Conditional logit coefficients for the four route attributes with 95% confidence intervals.

To ease interpretation, the attribute coefficients were converted into time equivalents using the delta method (Table 3). One additional transfer was equivalent to 14.7 minutes of travel time (95% CI [11.1, 18.3]). One additional minute of walking corresponded to 1.5 minutes of travel time. One additional crowding level corresponded to 2.5 minutes of travel time but was not statistically significant.

Table 3: Attribute trade-offs expressed in minutes of travel time.

Factor	Minutes	95% CI	t
One transfer	14.7*	[11.1; 18.3]	7.96
One minute of walking	1.5*	[1.2; 1.8]	10.27
One crowding level	2.5	[−1.4; 6.4]	1.24

Preference Heterogeneity

A mixed logit model with random coefficients tested for preference heterogeneity (Table 4). The mean coefficients confirmed the conditional logit ordering, and crowding now reached significance on average ($\mu = -0.623$, $p < .05$). The standard deviations were significant for crowding, travel time, and transfers, which indicates heterogeneous preferences between respondents.[1] The heterogeneity was by far the largest for crowding ($\sigma = 0.65$, $p < .001$). Figure 4 illustrates this pattern: the crowding distribution is wide and crosses zero, meaning that some travellers were indifferent to or even tolerant of crowding, whereas the travel-time and walking distributions were narrow. The mixed logit fit the data substantially better than the conditional logit (Table 5).

Table 4: Mixed logit estimates: means (μ) and standard deviations (σ) of the random coefficients (robust standard errors).[2]

Parameter	Estimate	Rob. SE	t	Sig.
ASC route B	−0.595	0.190	−3.13	**
ASC route C	0.176	0.163	1.08	
Walking (μ)	−0.241	0.052	−4.67	***
Crowding (μ)	−0.623	0.310	−2.01	*
Travel time (μ)	−0.150	0.031	−4.78	***
Transfers (μ)	−2.054	0.281	−7.31	***
Walking (σ)	0.030	0.025	1.19	
Crowding (σ)	0.650	0.098	6.64	***
Travel time (σ)	0.057	0.009	6.18	***
Transfers (σ)	0.630	0.087	7.23	***

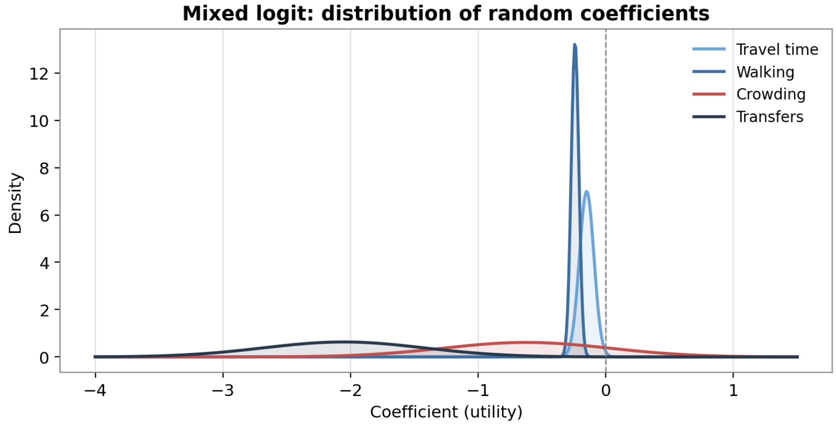

Figure 4: Mixed logit: estimated distribution of the random coefficients across travellers. The wide crowding distribution indicates strong preference heterogeneity.

[1] The sign of an estimated standard deviation is not interpretable; only its magnitude indicates the degree of preference heterogeneity.
[2] *Note.* Significance is based on the robust t-values: *p < .05, **p < .01, ***p < .001. Coefficients without an asterisk are not statistically significant (n.s.).

Table 5: Model comparison (lower AIC and BIC indicate better fit).

Model	k	LL	AIC	BIC	ρ²
Conditional logit	6	−1602.2	3216.4	3249.0	0.133
Mixed logit	10	−1546.2	3112.5	3166.7	0.164
Context model	20	−1500.4	3040.8	3149.3	0.189
Sociodemographic	22	−1590.3	3224.6	3344.0	0.140

Context-Dependent Valuation

The context model estimated separate attribute coefficients for the relaxed, time-pressure, heat, and rain conditions, treating the two weather variants separately, and achieved the best fit of all specifications (Table 5). Figure 5 displays the resulting coefficients. Travel time and transfers were significant deterrents across the relaxed, time-pressure, and heat contexts. The valuation of the comfort attributes, however, depended strongly on the context. Crowding was a significant deterrent only under heat ($b = -1.625$, $p < .001$), where it was about twice as strong as under relaxed conditions. Walking likewise became significant only under heat ($b = -0.238$, $p < .001$). The estimates for the rain variant were unstable and mostly non-significant, which reflects the small number of rain tasks.

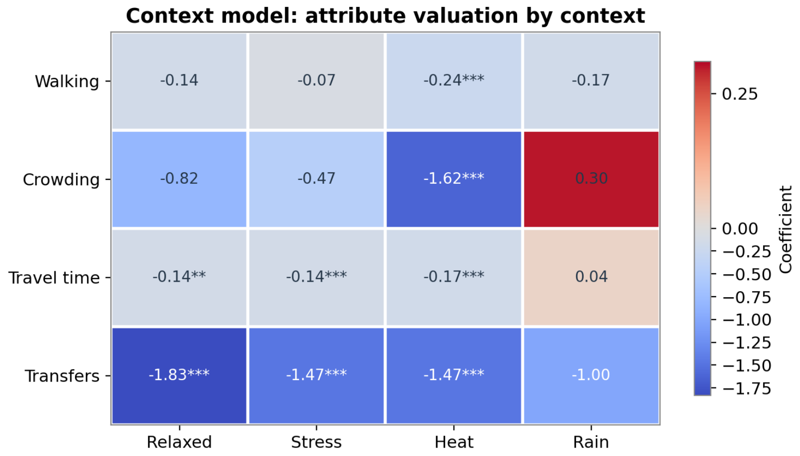

Figure 5: Context model: estimated attribute coefficients by travel context. Asterisks mark significant effects (p < .05, p < .01, **p < .001).

Table 6 expresses the crowding effect as a time equivalent for each context. Under heat, one crowding level was worth 9.6 minutes of travel time. The same crowding level was worth 5.8 minutes under relaxed conditions and a non-significant 3.4 minutes under time pressure. Figure 6 illustrates this amplification of crowding aversion under heat.

Table 6: Disutility of one crowding level expressed in minutes of travel time, by context.

Context	Minutes	95% CI	t
Relaxed	5.8*	[2.6; 9.1]	3.50
Time pressure	3.4	[−0.6; 7.3]	1.66
Heat	9.6*	[7.6; 11.6]	9.44

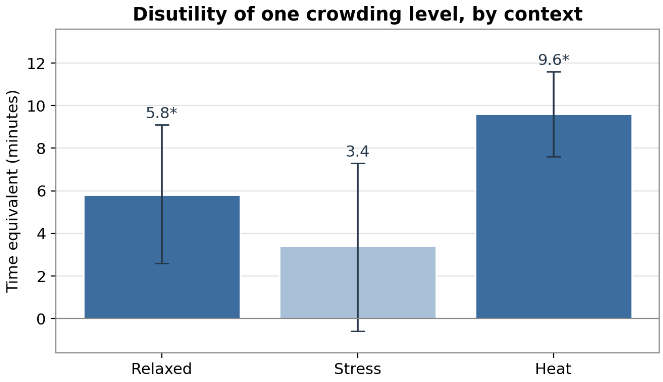

Figure 6: Time-equivalent disutility of one crowding level by context, with 95% confidence intervals. Bars marked with an asterisk are significant.

Sociodemographic Moderation

Finally, a model with interactions between the attributes and the sociodemographic characteristics tested for observed heterogeneity. None of the interactions with age, origin, gender, or PT-use frequency reached significance, and this model had the poorest fit of all specifications (Table 5). The only sociodemographic effect that approached significance appeared in the context model, where older respondents were slightly less sensitive to travel time ($b = 0.015$, $p < .05$). Overall, the measured sociodemographic characteristics explained little of the variation in attribute valuation.

Because such small effects are difficult to detect at this sample size, the non-significant interactions were additionally examined for directional tendencies. For this purpose, individual-level posterior coefficients were computed for each respondent from the mixed logit, following Train (2009), and averaged by sociodemographic group (Table 7). Although no group difference was significant, several weak and directionally plausible tendencies emerged. Respondents from larger cities and frequent PT users valued crowding somewhat less negatively, that is, they were slightly more crowding-tolerant. Female respondents valued transfers somewhat more negatively. Older respondents (aged 40 and above) tended to be slightly less sensitive to travel time, consistent with the borderline effect in the context model, and slightly more averse to crowding, although both age differences were small. All remaining contrasts were essentially flat. These patterns are reported as descriptive tendencies, not as significant effects.

Table 7: Mean individual coefficients by sociodemographic group (descriptive; group differences not significant).

Group	n	Crowding	Transfers	Travel time
Female	50	−0.62	−2.14	−0.153
Male / other	49	−0.64	−1.96	−0.145
Younger (< 40)	73	−0.61	−2.06	−0.154
Older (≥ 40)	26	−0.68	−2.02	−0.135
Larger city	40	−0.59	−2.06	−0.144
Smaller place	59	−0.65	−2.04	−0.152
Frequent PT	40	−0.57	−2.01	−0.151
Infrequent PT	59	−0.67	−2.08	−0.148

Note. Group boundaries were defined as follows. Larger city comprises respondents from a metropolis or large city, whereas smaller place comprises medium towns, small towns, and rural areas. The age groups are split at 40 years. Frequent PT users travel daily or several times per week, and infrequent users travel less often.

Factor Relevance and Effect Sizes

Table 8 links the literature review to the present results. For each of the four attributes, it states why the attribute was selected, the effect size reported for it in the international literature, and the corresponding estimate obtained in this study. The comparison shows that the present results agree in direction and broadly in magnitude with the existing evidence. The transfer penalty of about 14.7 minutes falls within the range of in-vehicle-time equivalents reported for a single transfer (Anderson et al., 2017). Crowding is the only attribute whose average effect was weaker than the literature would suggest, which the context analysis attributes to its dependence on heat.

Table 8: The four route attributes: rationale for inclusion, effect sizes reported in the literature, and the estimate obtained in this study.

Attribute	Rationale for inclusion	Effect reported in the literature	This study
Travel time (Dauer)	Primary criterion for screening route alternatives (Anderson et al., 2017).	Reference time component; other time components are valued relative to it (Anderson et al., 2017).	b = −0.109*** (reference unit)
Transfers (Umstiege)	Burden rises disproportionately beyond the first transfer (Jin et al., 2017).	≈ 12–18 min of in-vehicle time per transfer (Anderson et al., 2017).	≈ 14.7 min per transfer
Walking (Fußweg)	Access and egress are almost always required and valued negatively (Berggren et al., 2022).	≈ 2× in-vehicle time (Anderson et al., 2017).	≈ 1.5 min per walking minute
Crowding	Key comfort attribute that reduces route utility (Raveau et al., 2011).	Standing raises perceived in-vehicle time; valued more negatively since the pandemic (Yap et al., 2023).	≈ 2.5 min per level (n.s.); up to 9.6 min under heat

Preference Typology

The individual-level coefficients introduced in Section 4.6 also permit an exploratory typology of travellers. The heterogeneity was confined to two attributes: the individual coefficients for crowding and transfers varied strongly across respondents (SD about 0.48 and 0.46), whereas those for walking and travel time were almost constant (SD 0.008 and 0.037). To visualise this, respondents were grouped by k-means into two preference types (Figure 7). The number of clusters was examined with the silhouette criterion, which peaked at $k = 2$ but stayed low throughout, at about 0.27 and well below 0.5. The individual coefficients therefore form a continuum rather than sharply separated groups, and the typology is illustrative only. The more sensitive type ($n = 57$; mean crowding coefficient −0.82, transfers −2.30) is on average younger (32 years) and more often female (58%), whereas the more tolerant type ($n = 42$; mean crowding coefficient −0.36, transfers −1.71) is somewhat older (37 years) and more often male.

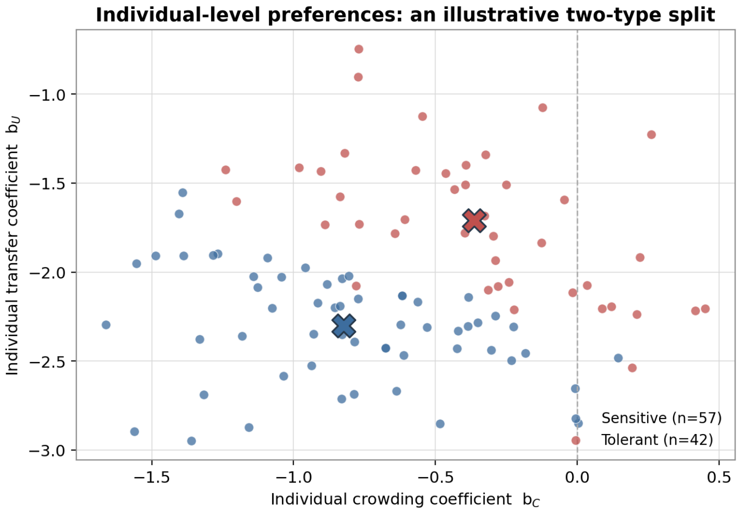

Figure 7: Individual-level preference coefficients for crowding and transfers, coloured by an illustrative two-type k-means split. The overlap of the two groups indicates a continuum rather than distinct types.

Taken together, three independent approaches, the interaction model, the directional tendencies, and the clustering, converge on the same conclusion: the preference heterogeneity is real but is not explained by the measured sociodemographic characteristics and therefore remains largely unobserved.

DISCUSSION AND CONCLUSION

This section interprets the findings in light of the hypotheses and the literature, discusses the limitations and practical implications, and draws overall conclusions.

Interpretation of the Findings

The results largely support the first hypothesis. Transfers carried by far the strongest penalty, followed by walking and travel time, which confirms the expected ordering. The dominance of transfers is consistent with the disproportionate burden that travellers attach to interchanging (Jin et al., 2017).

The transfer penalty of about 14.7 minutes also falls within the range reported for the Copenhagen network (Anderson et al., 2017). Crowding, however, was not a significant deterrent on average, so the first hypothesis holds only partially.

The second hypothesis is supported. Crowding was weaker than the time and transfer attributes and depended strongly on the context. Its average effect became significant only in the mixed logit model, which is consistent with the heterogeneous valuation of crowding reported elsewhere (Raveau et al., 2014). The finding that crowding is increasingly aversive also matches recent post-pandemic evidence (Yap et al., 2023).

The third hypothesis is only partially supported. Adverse weather clearly reshaped the valuation of comfort attributes, as both crowding and walking became significant deterrents under heat. This pattern is plausible because thermal discomfort makes a crowded vehicle and an exposed walking stretch more aversive. The result aligns with the view that contextual factors shape route choice in PT (Grison et al., 2017). The expected effect of time pressure, by contrast, did not materialise, as the valuation of time-related attributes was similar under relaxed and time-pressure conditions. The rain variant produced unstable estimates and did not allow firm conclusions.

The fourth hypothesis is supported. The significant standard deviations show that attribute valuations differed systematically between respondents. The heterogeneity was largest for crowding, where some travellers were tolerant and others strongly averse. This dispersion is consistent with the latent classes of crowding sensitivity identified in smart-card data (Mo et al., 2023). The sociodemographic characteristics measured here explained little of this heterogeneity, which suggests that the relevant differences are largely unobserved.

Limitations

Several limitations qualify these findings. The sample was a convenience sample of 99 respondents and skewed towards younger travellers, which limits its representativeness for the German population. The choices were stated rather than revealed, so the results may be affected by the hypothetical nature of stated-choice tasks (Louviere et al., 2000). The heat and rain variants were each based on only three tasks, which made their estimates unstable, especially for rain. The study did not include a manipulation check, so it cannot be verified that respondents internalised the intended situational framing. Crowding was presented as a three-level categorical attribute but treated as linear in the models, which is a simplification. Finally, the normal distribution assumed for the random coefficients can imply counterintuitive signs for some respondents, and a sign-constrained distribution might be more appropriate for attributes with an unambiguous direction (Train, 2009).

Practical Implications

The findings have several practical implications for PT planning and passenger information. Because transfers dominate route disutility, reducing the number of transfers and improving interchange quality should be a priority for operators. Well-equipped interchange points can reduce the perceived burden of transferring (Berggren et al., 2022). Because crowding becomes far more aversive under heat, capacity and in-vehicle climate should be prioritised on hot days as part of resilient service provision. Reserve capacity is, more generally, valuable for the robustness of PT networks (Cats & Jenelius, 2015). From a Human Factors perspective, passenger-information systems could adapt their route recommendations to the situation, for example by steering travellers towards less crowded routes during heat.

Conclusion and Outlook

This study examined how travellers in Germany value four route attributes and how their valuation changes across three situational contexts. Transfers consistently dominated route choice, travel time and walking were significant deterrents, and crowding mattered mainly under specific conditions. The central contribution is the evidence that situational context, and heat in particular, reshapes comfort-related preferences, which supports treating such contexts explicitly in demand models. Future work should replicate the study with a larger and more representative German sample. A balanced design with more heat and rain tasks would yield more stable context estimates. Adding a manipulation check and validating the results against revealed-preference data would further strengthen the conclusions. Taken together, the study provides a first behavioural account of context-dependent route choice in the German PT system and a basis for more resilient, user-centred service design.

REFERENCES

1. Anderson, M. K., Nielsen, O. A., & Prato, C. G. (2017). Multimodal route choice models of public transport passengers in the Greater Copenhagen Area. *EURO Journal on Transportation and Logistics, 6*(3), 221–245.
2. Berggren, U., Kjær-Rasmussen, T., Thorhauge, M., Svensson, H., & Brundell-Freij, K. (2022). Public transport path choice estimation based on trip data from dedicated smartphone app survey. *Transportmetrica A: Transport Science, 18*(3), 1813–1846.
3. Cats, O., & Jenelius, E. (2015). Planning for the unexpected: The value of reserve capacity for public transport network robustness. *Transportation Research Part A: Policy and Practice, 81*, 47–61.
4. Cats, O., & Jenelius, E. (2018). Beyond a complete failure: The impact of partial capacity degradation on public transport network vulnerability. *Transportmetrica B: Transport Dynamics, 6*(2), 77–96.
5. Chen, T., Fu, X., Hensher, D. A., Li, Z.-C., & Sze, N. N. (2024). Effects of proactive and reactive health control measures on public transport preferences of passengers – A stated preference study during the COVID-19 pandemic. *Transport Policy, 146*, 175–192.
6. Chia, J., Lee, J. B., & Han, H. (2020). How does the location of transfer affect travellers and their choice of travel mode? — A smart spatial analysis approach. *Sensors, 20*(16), 4418.
7. Cong, C., Li, X., Yang, S., Zhang, Q., Lu, L., & Shi, Y. (2022). Impact estimation of unplanned urban rail disruptions on public transport passengers: A multi-agent based simulation approach. *International Journal of Environmental Research and Public Health, 19*(15), 9052.
8. Eltved, M., Nielsen, O. A., & Rasmussen, T. K. (2018). The influence of frequency on route choice in mixed schedule- and frequency-based public transport systems – The case of the Greater Copenhagen Area. In *Proceedings of the Conference on Advanced Systems in Public Transport and TransitData (CASPT)*.
9. Eluru, N., Chakour, V., & El-Geneidy, A. M. (2012). Travel mode choice and transit route choice behavior in Montreal: Insights from McGill University members commute patterns. *Public Transport, 4*(2), 129–149.
10. Grison, E., Burkhardt, J.-M., & Gyselinck, V. (2017). How do users choose their routes in public transport? The effect of individual profile and contextual factors. *Transportation Research Part F: Traffic Psychology and Behaviour, 51*, 24–37.
11. Hamadneh, J., & Jaber, A. (2023). Modeling of intra-city transport choice behaviour in Budapest, Hungary. *Journal of Urban Mobility, 3*, 100049.
12. 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.
13. Jin, F., Yao, E., Zhang, Y., & Liu, S. (2017). Metro passengers' route choice model and its application considering perceived transfer threshold. *PLOS ONE, 12*(9), e0185349.

14. Leng, N., & Corman, F. (2020). The role of information availability to passengers in public transport disruptions: An agent-based simulation approach. *Transportation Research Part A: Policy and Practice, 133*, 214–236.
15. Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). *Stated choice methods: Analysis and applications*. Cambridge University Press.
16. Marra, A. D., & Corman, F. (2023). How different network disturbances affect route choice of public transport passengers: A descriptive study based on tracking. *Expert Systems with Applications, 213*, 119083.
17. McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), *Frontiers in econometrics* (pp. 105–142). Academic Press.
18. Mo, B., Koutsopoulos, H. N., & Zhao, J. (2022). Inferring passenger responses to urban rail disruptions using smart card data: A probabilistic framework. *Transportation Research Part E: Logistics and Transportation Review, 159*, 102628.
19. Mo, B., Ma, Z., Koutsopoulos, H. N., & Zhao, J. (2023). Passenger path choice estimation using smart card data: A latent class approach with panel effects across days. *arXiv*. https://arxiv.org/abs/2301.03808
20. Mocanu, T., Joshi, J., & Winkler, C. (2021). A data-driven analysis of the potential of public transport for German commuters using accessibility indicators. *European Transport Research Review, 13*, Article 54.
21. Raveau, S., Muñoz, J. C., & de Grange, L. (2011). A topological route choice model for metro. *Transportation Research Part A: Policy and Practice, 45*(2), 138–147.
22. Raveau, S., Guo, Z., Muñoz, J. C., & Wilson, N. H. M. (2014). A behavioural comparison of route choice on metro networks: Time, transfers, crowding, topology and socio-demographics. *Transportation Research Part A: Policy and Practice, 66*, 185–195.
23. Shelat, S., Cats, O., & van Lint, J. W. C. (2021). Quantifying travellers' evaluation of waiting time uncertainty in public transport networks. *Travel Behaviour and Society, 25*, 209–222.
24. Soza-Parra, J., Raveau, S., & Muñoz, J. C. (2022). Public transport reliability across preferences, modes, and space. *Transportation, 49*(2), 621–640.
25. Train, K. E. (2009). *Discrete choice methods with simulation* (2nd ed.). Cambridge University Press.
26. van der Hurk, E., Kroon, L., Maróti, G., & Vervest, P. (2015). Deduction of passengers' route choices from smart card data. *IEEE Transactions on Intelligent Transportation Systems, 16*(1), 430–440.
27. Yap, M., Wong, H., & Cats, O. (2023). Public transport crowding valuation in a post-pandemic era. *Transportation, 52*, 287–306.

DECLARATION ON THE USE OF AI TOOLS

During the preparation of this paper, the author used AI-based assistants, specifically Anthropic's Claude and Claude Code, as support tools. They were used to support the literature research, for language editing and formulation (including translation and phrasing) and for debugging the R code used in the analysis. The study design, the data collection, the statistical models, and the results are the author's own work. All AI-assisted content was critically reviewed, verified against the original sources, and edited by the author, who takes full responsibility for the content of this paper.