

Richter, Cornelia; Keuchel, Stephan

**Article**

## Modelling mode choice in passenger transport with integrated hierarchical information integration

Journal of Choice Modelling

**Provided in Cooperation with:**

Journal of Choice Modelling

*Suggested Citation:* Richter, Cornelia; Keuchel, Stephan (2012) : Modelling mode choice in passenger transport with integrated hierarchical information integration, Journal of Choice Modelling, ISSN 1755-5345, University of Leeds, Institute for Transport Studies, Leeds, Vol. 5, Iss. 1, pp. 1-21

This Version is available at:

<https://hdl.handle.net/10419/66845>

**Standard-Nutzungsbedingungen:**

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

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

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

**Terms of use:**

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

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

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



<http://creativecommons.org/licenses/by-nc/2.0/uk/>

# Modelling Mode Choice in Passenger Transport with Integrated Hierarchical Information Integration

Cornelia Richter<sup>1,\*</sup>

Stephan Keuchel<sup>1,†</sup>

<sup>1</sup>Department of Business Engineering, Westphalian University of Applied Sciences,  
August-Schmidt-Ring 10, D 45665 Recklinghausen, Germany

Received 30 April 2011, revised version received 23 September 2011, accepted 23 December 2011

## Abstract

The Integrated Hierarchical Information Integration (HII-I) approach allows to include a larger number of attributes in choice experiments by summarising similar attributes into constructs. In separate sub-experiments, one construct is described by its attributes while the other constructs are included by summarising construct values. This approach allows for testing of process equality in order to know if the different sub-experiments may be concatenated into an overall model.

In this paper, the HII-I approach is applied to model the mode choice between a regional train, a (hypothetical) regional bus and a car (only available for car users). Test results show that process equality is given when analysing only the data of the bi-modal sub-experiments whereas the assumption of process equality is rejected for data of the tri-modal sub-experiments, where differences in error variances between the sub-experiments are found. This empirical finding suggests that it is possible to construct separate sub-experiments while arriving at a single concatenated model.

*Keywords: integrated choice experiments, hierarchical information integration, public transport, service quality, discrete choice experiments*

---

\* Corresponding author, T: +49-2361 915488, F: + 49-2361 915756, [cornelia.richter@w-hs.de](mailto:cornelia.richter@w-hs.de)  
† T: +49-2361 915428, F: + 49-2361 915756, [stephan.keuchel@w-hs.de](mailto:stephan.keuchel@w-hs.de)



## 1 Introduction

The improvement of service quality is seen as a means to make public transport more attractive in contrast to individual transport modes (Regulation (EC) No. 1371/2007). The list of service quality attributes that potentially influence mode choice in passenger transport is rather long. Transport mode choice is usually analysed with discrete choice analysis using stated or revealed preference data. In stated choice experiments respondents are asked to choose between different transport modes which are described by a combination of different attributes with varying levels. It is therefore possible to analyse transport modes beyond the current alternatives. When there are too many attributes included in stated choice experiments, the risk of high dropout rates and biased results increases.

One solution to handle a larger number of attributes is the use of the Integrated Hierarchical Information Integration (HII-I) approach proposed by Oppewal et al. (1994). This approach improves the Hierarchical Information Integration (HII) approach proposed by Louviere (1984). It is based on the idea that similar attributes are grouped into constructs. In separate sub-experiments, one construct is described by its attributes and the other constructs are included by construct values. Before concatenating the separate sub-experiments into an overall model it is necessary to test for process equality across sub-experiments.

To our knowledge, there are only two applications of the Integrated Hierarchical Information Integration for which process equality was tested. Oppewal et al. (1994) tested the hypothesis of equal taste parameters in an HII-I application on consumer choice of shopping centre which included four constructs. They did not find evidence for equal taste parameters across sub-experiments. Van de Vyvere et al. (1998) tested the hypothesis of equal taste parameters and equal scale parameters in an HII-I application on residential choices which included three constructs. They could not reject the hypothesis of equal taste parameters but scale parameters differed across sub-experiments.

The aim of the present paper is to apply the Integrated Hierarchical Information Integration approach to model mode choice in public transport and to test if the separate sub-experiments can be concatenated into an overall model.

To this end the remainder of this paper is organised as follows: The underlying methodology is introduced in section 2. The research design and the data collection process are outlined in section 3, followed by a description of the sample in section 4. The analysis and the results are presented in section 5. Section 6 concludes and discusses aspects for further analysis.

## 2 Methodology

The original Hierarchical Information Integration (HII) approach has been proposed by Louviere (1984). It allows for the analysis of a larger number of attributes and is based on the idea that similar attributes can be grouped into constructs. If there are  $I$  constructs,  $I$  separate sub-experiments have to be created. Each of these  $I$  sub-experiments is designed as a rating task in which the trade-off between the attributes belonging to only one single construct is measured. Further, the  $I$  constructs whose levels are summarising construct values are traded-off in a bridging experiments. This bridging experiment was traditionally designed as a rating experiment to evaluate a single alternative but can also be designed as a choice experiment with the choice

between alternatives that are described by combinations of summarising construct values (Louviere and Timmermans 1990).

Oppewal et al. (1994) outline several problems and limitations of the original HII approach: The assumed hierarchical structure cannot be tested and several models rather than a single model are estimated. In the sub-experiments, the remaining constructs are omitted and inferences of these constructs are unclear. Further, the difficulty of the bridging task is unclear and sub-experiments cannot be designed as choice experiments. It is not possible to estimate interactions between attributes and constructs because there is no trade-off between attributes of one construct and the other constructs in the experiments. To overcome these problems, they proposed the Integrated Hierarchical Information Integration (HII-I) approach.

The HII-I approach is based in the following assumptions (Oppewal et al. 1994, van de Vyvere et al. 1998, Molin and Timmermans 2009):

An individual's choice is influenced by a set  $X$  that consists of  $N$  attributes. A subset of  $N_i$  attributes  $X_{in}$  ( $n=1, \dots, N_i$ ) can be grouped into  $I$  constructs<sup>1</sup> that are denoted by  $G_i$  ( $i=1, \dots, I$ ). Typically, each attribute  $X_{in}$  is part of one and only one construct. The number of attributes in each construct does not necessarily need to be the same.

In sub-experiments, not only the attributes of one construct but also summarising values<sup>2</sup> of the remaining other constructs are included. Sub-experiments can be designed as discrete choice experiments (and then also can be called Integrated Choice Experiments) that are consistent with random utility theory. A general relationship between the attributes and the constructs is illustrated in Figure 1. The grey shaded attributes and constructs describe a choice alternative of a sub-experiment in which the alternative is characterised by the attributes  $X_{11}$  to  $X_{1N1}$  of construct  $G_1$  and the summarising values of the constructs  $G_2$  to  $G_I$ .

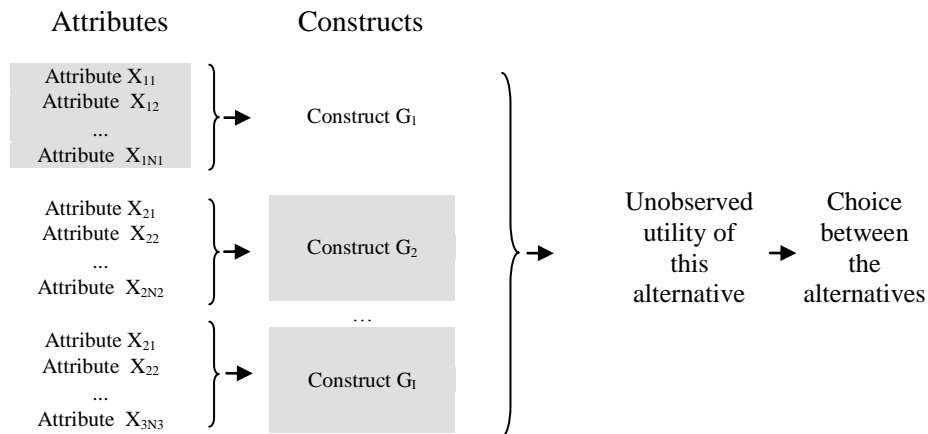


Figure 1 General Relationship between the Attributes and the Constructs

<sup>1</sup> The grouping of attributes into constructs can be based on “logic, theory, or empirical evidence from literature or pre-experimental research” (Oppewal et al. 1994). One possible empirical approach is proposed by Bos et al. (2002 and 2003).

<sup>2</sup> Rating tasks are usually used beforehand in order to familiarise respondents with the attributes, their corresponding constructs, and the construct values.

The systematic component of the utility  $V_{ir}$  of the alternative  $r$  in sub-experiment  $i$  can be described as

$$V_{ir} = X_{ir} \beta_i + G_{rj} \gamma_j, \quad (1)$$

where  $X_{ir}$  is the vector of the attributes,  $\beta_i$  the vector of the taste parameters of the attributes,  $G_{rj}$  the vector of the other constructs ( $j \neq i$ ), and  $\gamma_j$  the vector of the taste parameters of the constructs. Using a multinomial logit (MNL) model, the probability that the alternative  $r$  in sub-experiment  $i$  is chosen from a choice set  $C$  can be expressed as

$$P_{ir} = \frac{e^{\mu_i V_{ir}}}{\sum_{s \in C} e^{\mu_i V_{is}}} = \frac{e^{\mu_i (X_{ir} \beta_i + G_{jr} \gamma_j)}}{\sum_{s \in C} e^{\mu_i (X_{is} \beta_i + G_{js} \gamma_j)}}, \quad (2)$$

where  $\mu_i$  is a scale parameter that is unidentified in a single model and therefore usually set to 1.0.

Each sub-experiment is supposed to describe the alternative with all  $N$  attributes, either directly as an attribute or indirectly as part of a construct. Consequently, for a given profile of attributes, the utility function derived from data of one of the  $I$  sub-experiments should be equal to the utility function derived from data of any other of the  $I$  sub-experiments. This process equality is given when both the scale parameters and the taste parameters of the different sub-experiments are equal. If process equality across sub-experiments exists, the  $I$  separate sub-experiments can be concatenated into an overall model containing all  $N$  attributes  $X$ .

The hypothesis of process equality (equal taste<sup>3</sup> and equal scale parameters) can be written as

$$H: \beta_i = \beta_j = \beta \text{ and } \mu_i = \mu_j = \mu. \quad (3)$$

It can be tested with a two-stage test (Swait and Louviere 1993, van de Vyvere et al. 1998). In the first stage, the hypothesis of equal taste parameters

$$H_1: \beta_i = \beta_j = \beta \quad (4)$$

is tested, while scale parameters may vary across sub-experiments. To this end, data of the  $I$  sub-experiments are concatenated into an artificial nested logit model<sup>4</sup> and taste parameters that are present in several sub-experiments constrained to be equal. Further, one of the scale parameters is arbitrarily set to 1.0, while the other scale parameters are relative. This model is estimated with full information maximum likelihood (FIML). This concatenated FIML model with taste parameters constrained to be equal and scale parameters allowed to vary is tested against the separate MNL models in which taste and scale parameters are confounded. The likelihood ratio test statistic is

$$\lambda_1 = -2 [LL_\mu - \sum LL_i], \quad (5)$$

---

<sup>3</sup> In the following test, the taste parameters  $\beta$  refer to the taste parameters of the attributes and constructs that are tested for equality.

<sup>4</sup> Artificial relates to the fact that respondents are not modelled as choosing between the sub-experiments. Therefore, IV parameters do not have to lie within the 0-1 bound, the necessary condition for random utility maximisation (Hensher and Bradley 1993, Hensher et al. 2005).

where  $LL_{\mu}$  is the log-likelihood value of the concatenated FIML model with scale parameters allowed to vary and  $LL_i$  the log-likelihood value of the  $i^{th}$  sub-experiment.  $\lambda_1$  is asymptotically chi-square distributed with the number of degrees of freedom equal to the sum of parameters in the separate sub-experiments minus the number of parameters in the concatenated model.

If  $H_1$  is rejected,  $H$  is also rejected. If  $H_1$  is not rejected, the hypothesis of equal scale parameters

$$H_2: \mu_i = \mu_j = \mu \quad (6)$$

is tested in the second stage. To this end, data of the  $I$  sub-experiments are concatenated into a MNL model in which the taste parameters that are present in several sub-experiments are constrained to be equal and scale parameters identical by definition. This concatenated model with taste and scale parameters constrained to be equal is tested against the concatenated model with taste parameters constrained to be equal and unconstrained scale parameters. The likelihood ratio test statistic is

$$\lambda_2 = -2 [LL_p - LL_{\mu}], \quad (7)$$

where  $LL_p$  is the log-likelihood value of the concatenated model with equal scale parameters.  $\lambda_2$  is asymptotically chi-square distributed with the number of degrees of freedom equal to the number of parameters in the concatenated model with scale parameters allowed to vary minus the number of parameters in the concatenated model with equal scale parameters.

If both  $H_1$  and  $H_2$  are not rejected, the hypothesis of equal taste and equal scale parameters is supported. If only  $H_2$  is rejected, the relative scale parameters can be interpreted as measures of differences in error variances between the sub-experiments.

### 3 Research Design and Data Collection

The first step in constructing integrated choice experiments to model the choice between different transport modes concerns the selection of attributes and their corresponding constructs.

Attributes that could potentially influence mode choice were obtained from the literature and interviews with experts of three public transport companies. An empirical pre-study was conducted to cluster the attributes into constructs, following closely the approach proposed by Bos et al. (2002 and 2003). Respondents were asked to group similar attributes and to name these groups. Data were collected in computer assisted personal interviews with over 500 train and car users. Data were analysed on aggregate level using the method of multidimensional scaling: The more often respondents grouped two attributes together, the closer these two attributes were arranged in a multidimensional space and vice versa. As a result, three constructs, namely 'Quality of Connection', 'Comfort', and 'Information', were selected for inclusion in the integrated choice experiments. Additionally, the total travel time and the fare/price were included as attributes in all sub-experiments.<sup>5</sup> The relationship between the attributes and their corresponding constructs is illustrated in Figure 2. The grey shaded attributes and constructs were used in the sub-experiment 'Quality of Connection' for the description of the train alternative. Analogously, the sub-experiments 'Comfort' and 'Information' were created. In contrast to the train and the

---

<sup>5</sup> It was assumed that respondents sum up different time and fare/price related components.

(hypothetical) bus alternative, the car alternative – if available – was described only by the total travel time and the price.

In the next step, the attribute levels were determined by the authors and experts of the three public transport companies. The attributes, that were part of the constructs, were described by two, three, and four levels, respectively (Table 1). The attribute values for the two public transport alternatives were calculated with reference to the respondent's current train alternative and the car alternative with reference to the respondent's current car alternative. Fare and price varied between 90 percent, 95 percent, 100 percent, 105 percent, and 110 percent of the current values and total travel time between +10 minutes, +5 minutes, +0 minutes, -5 minutes, and -10 minutes. The three constructs were described by the three levels '- -', '+ +', and '+ + + +'<sup>6</sup> which represented evaluations on a scale ranging from '- - - -' (very bad) to '+ + + + + ' (very good).

Data were collected on board of regional trains of three public transport companies in Westphalia/Germany using computer assisted personal interviews. Questionnaires were programmed in MS Access allowing for more individual questionnaires in contrast to traditional paper questionnaires. Train users were randomly selected by trained interviewers and asked to participate in an interview during their current trip.<sup>7</sup>

If willing to participate, they were asked to describe their current trip, in particular with respect to the total travel time and the fare. Respondents who indicated to be able to use a car for the current trip (also as a passenger) were further asked to

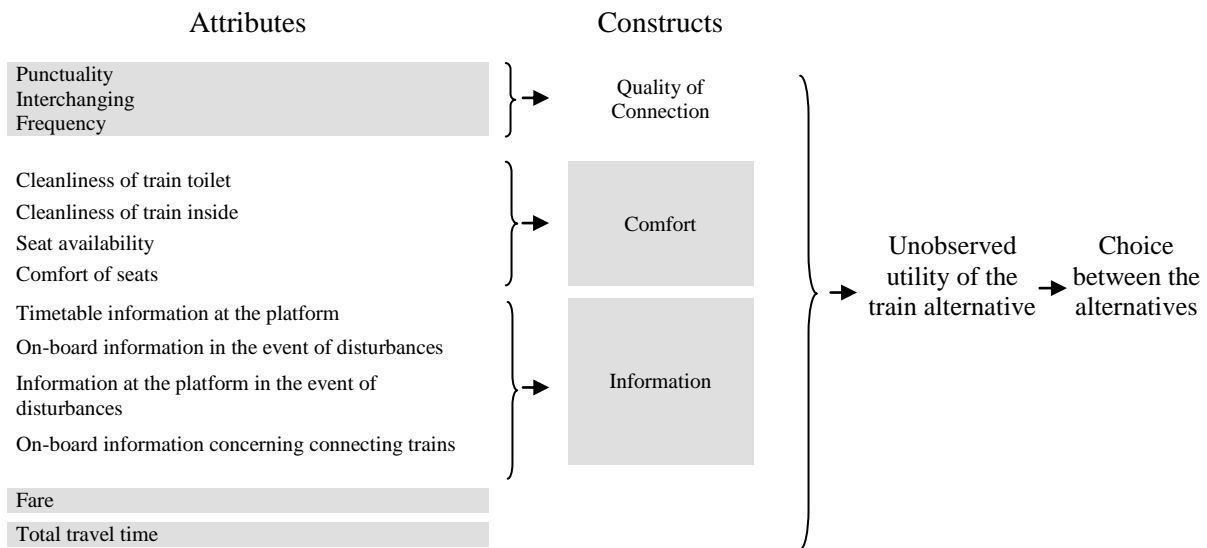


Figure 2: Relationship between the Attributes and the Constructs in the Sub-Experiments for the Train Alternative

<sup>6</sup> A pre-test showed that respondents rated the construct values for their current trip rather good or very good, therefore, very bad construct values were not included in the design in order to be as realistic as possible.

<sup>7</sup> Interviews with car users are ongoing and therefore not described and analysed in this paper.

indicate the total travel time and the estimated price for this alternative. Then, separate rating-experiments were created for each construct in which respondents rated profiles of attribute combinations with random design (Figure 3). The respondents rated three profiles for each of the three constructs on a scale ranging from ‘- - - -’ (very bad) to ‘+ + + +’ (very good). This task aimed to familiarise the respondents with the rating scales of the constructs and their corresponding attributes. Finally, respondents were randomly assigned to one of the three sub-experiments in which they had to choose one of the transport modes proposed in the experiments (Figure 4). They were asked to assume the same context as in their current trip. Respondents without a car alternative had to choose between a regional train and a (hypothetical) regional bus. The car alternative was only available for those who indicated to be able to use a car for the current trip. Random designs were used to create the choice experiments. After five choice situations, the interviewer could ordinarily quit the infinite loop of choice situations. Finally, some demographic questions and questions concerning the usual travel behaviour were posed. A total of 1609 valid interviews were conducted during May and October 2010. This sample included long distance travellers. A sub-sample of 1152 regional travellers is analysed in this paper. Concerning the actual number of HII-I choice situations of this sub-sample, with 90.4 percent a vast majority of the respondents completed five or six choice situations, only 3.0 percent less than five, and 6.7 percent seven or more choice situations. On average, respondents completed

Table 1 Constructs, Attributes, and their Levels

Construct	Attribute	1st attribute level	2nd attribute level	3rd attribute level	4th attribute level
Quality of Connection	Punctuality	On time or up to 3 minutes late	3 to 10 minutes late	10 to 20 minutes late	
	Interchanging	0 interchanges	1 interchange with guaranteed connection	1 interchange without guaranteed connection	
	Frequency	Every 30 minutes	Every 60 minutes	Every 120 minutes	
Comfort	Cleanliness of train/bus toilet	Clean	Dirty	No toilet	
	Cleanliness of train/bus inside	Clean	Floor is sticky, paper lies around on the seats		
	Seat availability Comfort of seats	During the whole trip Comfortable	During half of the trip Not comfortable	No free seat	
Information	Timetable information at the platform/bus stop	Available	Not available		
	On-board information in the event of disturbances	Announcements and display of reason and duration of disturbances	Announcements of reason and duration of disturbances	Display of reason and duration of disturbances	No information
	Information at the platform/bus stop in the event of disturbances	Announcements and display of reason and duration of disturbances	Announcements of reason and duration of disturbances	Display of reason and duration of disturbances	No information
	On-board information concerning connecting trains/buses	Announcements and display	Announcements	Display	No information



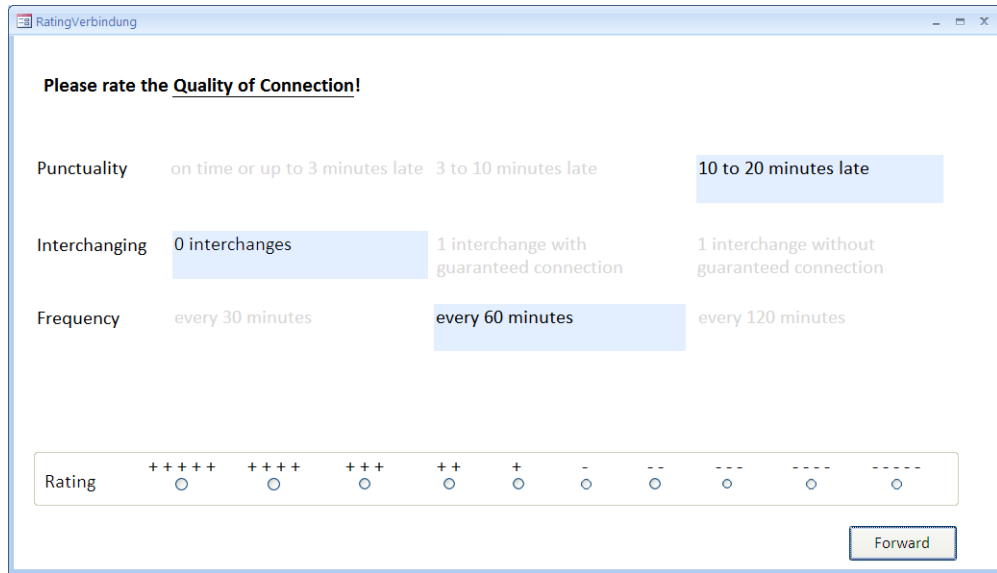


Figure 3 Example Screen of the Rating Experiment 'Quality of Connection'<sup>8</sup>  
(Translated Version)

5.6 HII-I choice situations (standard deviation of 1.2 and a median of 5). Only two respondents completed more than 15 choice situations but only the first 15 situations were included in the analysis.

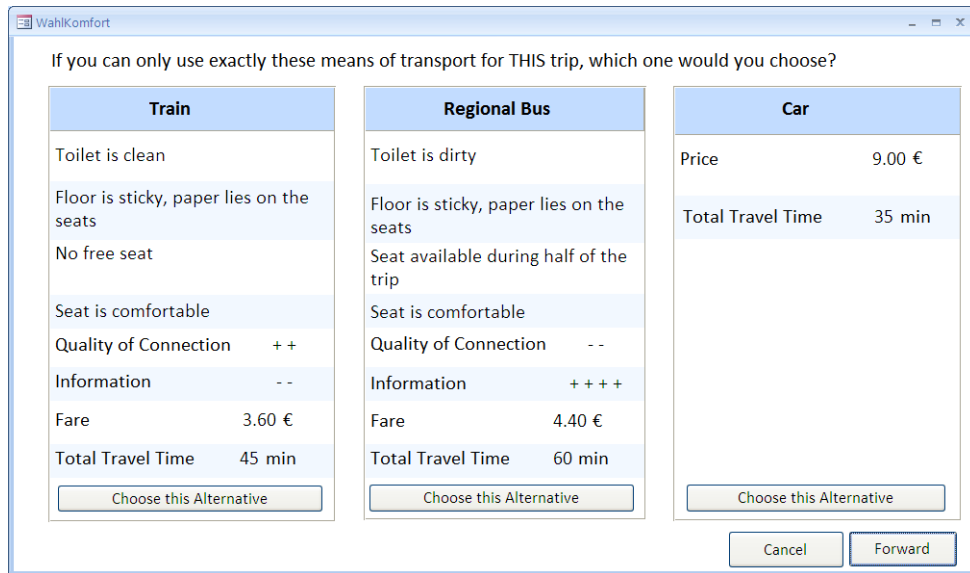


Figure 4 Example Screen of the Sub-Experiment 'Comfort' (Translated Version)

<sup>8</sup> When displaying all attributes and highlighting those attributes that are present in a given situation, respondents might adopt simplifying strategies to rate the situation instead of properly read and trade off the attributes. However, the results of the analysis in section 5.2 did not indicate a systematic ordering effect or a systematic linear relationship. The risk of simplifying strategies might have been lower in this study because personal interviews were used in which interviewers explicitly showed all attributes levels.

## 4 Description of the Sample

In this paper a sub-sample of 1152 respondents was used which was selected with regard to the fare/price (less than or equal to 15€ and 30€ for the train alternative and the car alternative, respectively) and the total travel time (more than 15 minutes and less than or equal to 90 minutes). The characteristics of the respondents of this sample are listed separately for the respondents of the bi-modal and the tri-modal experiments in Table 2, showing that 661 respondents participated in the bi-modal and 491 respondents in the tri-modal experiments; (1) nearly as many men as women participated in the bi-modal experiments and 61 percent of the respondents of the tri-modal experiments were male; (2) most of the respondents of the bi-modal experiments were between 18 and 30 years of age and most of the respondents of the tri-modal experiments were between 31 and 50 years of age; (3) three out of four respondents of the bi-modal experiments had a driving licence and nearly all respondents of the tri-modal experiment<sup>9</sup>; (4) in both experiments three out of four respondents had a season ticket; (5) with 70 and 77 percent, respectively, a large majority of the respondents took the train several times per week; and (6) 30 and 54 percent, respectively, of the respondents took the car several times a week. The average total travel time and the average fare of the train alternative in the bi-modal

Table 2 Characteristics of the Respondents of the Sample

	Bi-modal		Tri-modal	
	Absolute (N=661)	Relative (%)	Absolute (N=491)	Relative (%)
1 Sex				
Male	319	48.3	298	60.7
Female	339	51.3	190	38.7
Missing values	3	0.5	3	0.6
2 Age (years)				
≤ 17	40	6.1	7	1.4
18 - 30	305	46.1	170	34.6
31 - 50	222	33.6	209	42.6
≥ 51	84	12.7	99	20.2
Missing values	10	1.5	6	1.2
3 Driving Licence				
Yes	493	74.6	473	96.3
No	165	25.0	15	3.1
Missing values	3	0.5	3	0.6
4 Season Ticket				
Yes	491	74.3	382	77.8
No	167	25.3	107	21.8
Missing values	3	0.5	2	0.4
5 Frequency of train use				
Several times per week	463	70.0	380	77.4
Several times per month	127	19.2	53	10.8
Rarely	62	9.4	53	10.8
Never	6	0.9	3	0.6
Missing values	3	0.5	2	0.4
6 Frequency of car use				
Several times per week	196	29.7	263	53.6
Several times per month	165	25.0	92	18.7
Rarely	212	32.1	113	23.0
Never	83	12.6	21	4.3
Missing values	5	0.8	2	0.4

<sup>9</sup> The sample of the tri-modal experiments contains data from respondents that used the car as a passenger.

and the tri-modal experiments, respectively, were 55 minutes and 54 minutes, respectively, and 3.63€ and 3.91€, respectively. The average total travel time of the car alternative was 44 minutes and the average price 8.75€. The characteristics of the respondents randomly assigned to the three different sub-experiments were analysed but no major differences were found.

## 5 Analysis and Results<sup>10</sup>

### 5.1 Coding

All categorical attributes were effect coded. To this end, to code an attribute with  $L$  levels,  $L-1$  effect variables were created. An effect variable is set equal to 1 when the corresponding qualitative level is present, -1 when the reference level is present and 0 otherwise. Consequently, the reference level equals the negative sum of the coefficients of effect variables and, in other words, all  $L$  levels sum up to 0. In contrast to the commonly applied dummy coding, effect coding overcomes the problem of correlation of the reference level with the intercept/alternative specific constant in choice experiments (Bech and Gyrd-Hansen 2005, Louviere et al. 2000). The information attributes that had four levels were separated into announcement effects, display effects, and interaction effects.

The construct ratings were coded on a linear scale with ‘- - - -’ being equal to 1 and ‘+ + + +’ equal to 10. Consequently the construct values ‘- -’, ‘+ +’, and ‘+ + + +’ that were used in the sub-experiments were coded as 4, 7, and 9.

### 5.2 Rating Experiments

Principally, the rating task was used to familiarise respondents with attributes, constructs, and their levels. Respondents were asked to give overall ratings of the three constructs for different attribute level combinations. The general representation of attributes by constructs was shown by means of linear regression analysis, where the attribute levels were independent variables and the construct rating the dependent variable. Results are listed in Table 3 to Table 5, showing that significant parameters could be estimated for all attribute levels except for the second level of the punctuality attribute which did not differ significantly from the intercept.

### 5.3 Choice Experiments

Depending on car availability, respondents participated either in bi-modal or tri-modal choice experiments. Since car availability might be a reason for differences in the choice process, it was tested if the choice process in the bi-modal and the tri-modal sub-experiments was equal. To that end, three likelihood ratio tests were applied for the three different sub-experiments: for each of the three sub-experiments a joint multinomial logit (MNL) model with bi-modal and tri-modal data was tested against two separate MNL models. The likelihood statistics led to  $(-2[-1553.8 - (-654.3 - 880.1)]) = 38.9 > 19.7$  with  $(11 + 14) - 14 = 11$  degrees of freedom for the sub-experiment ‘Quality of Connection’,  $-2[-1582.7 - (-707.2 - 851.7)] = 47.5 > 19.7$  with

---

<sup>10</sup> Data were analysed using the NLOGIT software package (Version 4.0.1).

Table 3 Regression results for rating task ‘Quality of Connection’  
(N=3522 rating situations)

		Par.	(t-ratio)
Intercept		5.232	(153.24)
Punctuality	<i>On time or up to 3 minutes late</i>	1.109	
	3 to 10 minutes late	0.027	(0.55)
	10 to 20 minutes late	-1.136	(-23.52)
Interchanging	<i>0 interchanges</i>	0.940	
	1 interchange with guaranteed connection	0.226	(4.64)
	1 interchange without guaranteed connection	-1.165	(-24.10)
Frequency	<i>Every 30 minutes</i>	1.623	
	Every 60 minutes	0.179	(3.66)
	Every 120 minutes	-1.801	(-37.29)
R <sup>2</sup>		0.474	

Table 4 Regression results for rating task ‘Comfort’ (N=3525 rating situations)

		Par.	(t-ratio)
Intercept		4.499	(139.37)
Cleanliness of train/bus inside	<i>Clean</i>	1.030	
	Floor is sticky, paper lies around on the seats	-1.030	(-31.92)
Comfort of seats	<i>Comfortable</i>	0.628	
	Not comfortable	-0.628	(-19.43)
Cleanliness of train/bus toilet	<i>Clean</i>	0.925	
	Dirty	-0.379	(-8.33)
	No toilet	-0.546	(-11.93)
Seat availability	<i>During the whole trip</i>	1.557	
	During half of the trip	-0.117	(-2.55)
	No free seat	-1.440	(-31.39)
R <sup>2</sup>		0.489	

Table 5 Regression results for the rating task ‘Information’ (N=3517 rating situations)

		Par.	(t-ratio)
Intercept		5.748	(175.45)
Timetable information at the platform/bus stop	<i>Available</i>	0.862	
	Not available	-0.862	(-26.32)
On-board information in the event of disturbances	Announcement effect	0.649	(19.85)
	Display effect	0.541	(16.52)
	Interaction effect	-0.365	(-11.16)
Information at the platform/bus stop in the event of disturbances	Announcement effect	0.662	(20.21)
	Display effect	0.610	(18.63)
	Interaction effect	-0.362	(-11.05)
On-board information concerning connecting trains/busses	Announcement effect	0.529	(16.14)
	Display effect	0.463	(14.15)
	Interaction effect	-0.322	(-9.85)
R <sup>2</sup>		0.456	

(11 + 14) – 14 = 11 degrees of freedom for the sub-experiment ‘Comfort, and -2[1234.3 – (-610.7 – 608.4)] = 30.2 > 25.0 with (15 + 18) – 18 = 15 degrees of freedom) for the sub-experiment ‘Information’. These differences supported a separate analysis of the bi-modal and tri-modal data.

### 5.3.1 Bi-modal Model

The estimated parameters and t-ratios for the three separate bi-modal MNL models are listed in Table 6. The parameters for the reference levels were also calculated. The number of choice situations, the number of estimated parameters, and the log-likelihood values of the models are indicated at the bottom of the table. All parameters of the sub-experiment ‘Quality of Connection’ were statistically significant, except three parameters which concerned all the second level (of three levels). When the parameters estimated for the second level (of three levels) are not statistically significant, it may be concluded that utility changes linearly with the attribute levels. However, the attribute levels themselves did not/not all have a linear relationship. In the sub-experiment ‘Comfort’, all parameters except two which concerned the second level (of three levels) and the attribute comfort of the seats were significant. Interaction of the comfort of the seats and seat availability was analysed but estimated parameters were not significant. In the sub-experiment ‘Information’, the train constant, fare, time, the two constructs, the timetable information at the platform/bus stop, the display effect concerning information at the platform/bus stop in the event of disturbances, and the interaction effect between the display and the announcements effect of on-board information concerning connecting trains/buses were statistically significant. In the three separate sub-experiments, all significant parameters had the expected signs.

To test for process equality across sub-experiments, a two-stage likelihood ratio test was used. In the first stage, the three separate models were tested against an artificial nested logit model in which the branch levels corresponded to the different sub-experiments. In this concatenated model, the parameters of fare, time, the three constructs and the alternative specific train constant were constrained to be equal. Further, the inclusive value (IV)<sup>11</sup> parameter of the ‘Quality of Connection’ branch was normalised to 1.0, whereas the IV parameters of the other two branches were free to vary. The NL model was estimated with full information maximum likelihood (FIML) estimation under RU2 (normalisation at the upper level). Results are displayed in Table 7. The same parameters as in the separate sub-experiments were significant and signs for significant parameters were as expected. The log-likelihood values of the separate MNL models were -654.25, -707.23, and -610.74, respectively, and the log-likelihood value of the concatenated model was -1977.90. The likelihood ratio test led to  $\lambda_1 = -2[-1977.90 - (-654.25 - 707.23 - 610.74)] = 11.33$ . This value was smaller than 14.07, the critical value of the  $\chi^2$  distribution at 0.05 significance level with  $(11 + 11 + 15) - 30 = 7$  degrees of freedom. Hence, the hypothesis of equal taste parameters could not be rejected.

Table 6 Separate Bi-modal Models

	Quality of Connection		Comfort		Information	
	Par.	(t-ratio)	Par.	(t-ratio)	Par.	(t-ratio)
Train constant	0.497	(7.20)	0.643	(9.58)	0.664	(9.11)

<sup>11</sup> The IV parameter from any level of a model is defined as the ratio of the scale parameter at that level to the scale parameter of the level directly below. The closer an IV parameter is to 1.0, the smaller the difference in variance between two levels of the NL tree (Hensher et al. 2005).

Fare		-0.674	(-4.64)	-0.829	(-4.78)	-0.518	(-3.26)
Time		-0.055	(-7.54)	-0.073	(-10.58)	-0.074	(-10.10)
Quality of Connection				0.177	(7.53)	0.242	(9.27)
Comfort		0.187	(7.41)			0.156	(6.31)
Information		0.086	(3.79)	0.071	(3.19)		
Punctuality	<i>On time or up to 3 minutes late</i>	0.321					
	3 to 10 minutes late	0.072	(1.04)				
	10 to 20 minutes late	-0.394	(-5.70)				
Interchanging	<i>0 interchanges</i>	0.403					
	1 interchange with guaranteed connection	-0.117	(-1.68)				
	1 interchange without guaranteed connection	-0.286	(-4.16)				
Frequency	<i>Every 30 minutes</i>	0.683					
	Every 60 minutes	0.018	(0.27)				
	Every 120 minutes	-0.702	(-9.64)				
Cleanliness of train/bus inside	<i>Clean</i>			0.128			
	Floor is sticky, paper lies around on the seats			-0.128	(-2.78)		
Cleanliness of train/bus toilet	<i>Clean</i>			0.139			
	Dirty			0.029	(0.44)		
	No toilet			-0.168	(-2.59)		
Seat availability	<i>During the whole trip</i>			0.635			
	During half of the trip			0.086	(1.34)		
	No free seat			-0.721	(-10.32)		
Comfort of seats	<i>Comfortable</i>			0.024			
	Not comfortable			-0.024	(-0.51)		
Timetable information at the platform/bus stop	<i>Available</i>						
	Not available					-0.119	(-2.44)
On-board information in the event of disturbances	Announcement effect					0.023	(0.45)
	Display effect					0.071	(1.44)
	Interaction effect					0.001	(0.02)
Information at the platform/bus stop in the event of disturbances	Announcement effect					0.050	(1.01)
	Display effect					0.137	(2.74)
	Interaction effect					-0.055	(-1.08)
On-board information concerning connecting trains/busses	Announcement effect					0.020	(0.41)
	Display effect					0.034	(0.68)
	Interaction effect					-0.108	(-2.23)
Number of observation		1213		1351		1121	
Number of parameters		11		11		15	
Log-likelihood value		-654.25		-707.23		-610.74	

Table 7 Concatenated Bi-modal Models (All Bi-modal Sub-Experiments)

		Concatenated model (scale parameters allowed to vary)		Concatenated model (scale parameters constrained to be equal)	
		Par.	( <i>t</i> -ratio)	Par.	( <i>t</i> -ratio)
Train constant		0.538	(10.47)	0.598	(14.95)
Fare		-0.600	(-6.84)	-0.671	(-7.39)
Time		-0.060	(-10.79)	-0.067	(-16.39)
Quality of Connection		0.176	(8.39)	0.202	(11.84)
Comfort		0.154	(8.26)	0.173	(9.89)
Information		0.072	(4.69)	0.079	(4.99)
Punctuality	<i>On time or up to 3 minutes late</i>	0.316		0.326	
	3 to 10 minutes late	0.072	(1.05)	0.073	(1.03)
	10 to 20 minutes late	-0.388	(-5.72)	-0.399	(-5.72)
Interchanging	<i>0 interchanges</i>	0.398		0.409	
	1 interchange with guaranteed connection	-0.110	(-1.58)	-0.110	(-1.57)
	1 interchange without guaranteed connection	-0.288	(-4.21)	-0.299	(-4.29)
Frequency	<i>Every 30 minutes</i>	0.674		0.700	
	Every 60 minutes	0.026	(0.38)	0.023	(0.34)
	Every 120 minutes	-0.699	(-10.09)	-0.724	(-9.97)
Cleanliness of train/bus inside	<i>Clean</i>	0.112		0.128	
	Floor is sticky, paper lies around on the seats	-0.112	(-2.69)	-0.128	(-2.79)
Cleanliness of train/bus toilet	<i>Clean</i>	0.121		0.138	
	Dirty	0.022	(0.39)	0.026	(0.40)
	No toilet	-0.143	(-2.55)	-0.164	(-2.55)
Seat availability	<i>During the whole trip</i>	0.552		0.634	
	During half of the trip	0.074	(1.31)	0.083	(1.30)
	No free seat	-0.626	(-7.25)	-0.717	(-10.44)
Comfort of seats	<i>Comfortable</i>	0.018		0.020	
	Not comfortable	-0.018	(-0.45)	-0.020	(-0.44)
Timetable information at the platform/bus stop	<i>Available</i>	0.096		0.114	
	Not available	-0.096	(-2.35)	-0.114	(-2.38)
On-board information in the event of disturbances	Announcement effect	0.021	(0.49)	0.025	(0.50)
	Display effect	0.058	(1.40)	0.067	(1.37)
	Interaction effect	0.003	(0.06)	0.003	(0.05)
Information at the platform/bus stop in the event of disturbances	Announcement effect	0.045	(1.07)	0.053	(1.09)
	Display effect	0.115	(2.71)	0.134	(2.74)
	Interaction effect	-0.043	(-1.00)	-0.051	(-1.02)
On-board information concerning connecting trains/busses	Announcement effect	0.015	(0.35)	0.019	(0.38)
	Display effect	0.031	(0.77)	0.034	(0.70)
	Interaction effect	-0.091	(-2.17)	-0.108	(-2.26)
IV parameter (normalised at RU2)					
Quality of Connection		1.0	(fixed)		
Comfort		0.867	(-1.40) <sup>1</sup>		
Information		0.833	(-1.86) <sup>1</sup>		
Number of observations		3685		3685	
Number of parameters		30		28	
Log-likelihood value		-1977.90		-1979.46	

<sup>1</sup> *t*-statistics are calculated for the hypothesis that the IV parameter equals 1.0.

In the second stage, the hypothesis of equal scale parameters was tested with a likelihood ratio test of the concatenated model against a concatenated model with equal scale parameters. The results of the concatenated model are displayed in Table 7, showing the same results concerning significance levels and signs as the other models displayed in Table 6. The log-likelihood value was -1979.46. The likelihood ratio test led to  $\lambda_2 = -2[-1979.46 - (-1977.90)] = 3.14$ . This value was smaller than 5.99, the critical value of the  $\chi^2$  distribution at 0.05 significance level with  $30 - 28 = 2$  degrees of freedom. The hypothesis of equal scale parameters could not be rejected. Therefore, process equality across the sub-experiments was assumed.<sup>12</sup>

### 5.3.2 Tri-modal Model

The estimated parameters and *t*-ratios of the three separate tri-modal MNL models are displayed in Table 8. Again, the parameters for the reference levels, the number of choice situations, the number of estimated parameters, and the log-likelihood values of the models are indicated. In the sub-experiment 'Quality of Connection', all parameters except two second-level (of three levels) parameters were significant. In the sub-experiment 'Comfort', all parameters except the third level (of three levels) of the cleanliness of the toilet, the second level (of three levels) of the seat availability, and the comfort of the seat were significant. Again, the interaction between the comfort of the seats and the seat availability was analysed but not significant. In the sub-experiment 'Information', only the two alternative specific constants, the fare and price, the two time parameters, the two construct parameters and the timetable information were found significant. In the three sub-experiments, all significant parameters had the expected signs.

Again, process equality across sub-experiments was tested. In the first stage, the three separate tri-modal MNL models were tested against an artificial nested logit model, in which the parameters of price and fare, the two time parameters, the three constructs and the two alternative specific constant parameters were constrained to be equal. The IV parameter of the 'Quality of Connection' branch was normalised to 1.0 and the IV parameters of the other two branches free to vary. Results of this concatenated model are displayed in Table 9. The same parameters as in the separate sub-experiments with the exception of the timetable information parameter were significant and, again, signs for the significant parameters were as expected. The log-likelihood values of the three separate MNL models were -880.09, -851.69, and -608.43, respectively, and the log-likelihood value of the concatenated model was -2346.41. With  $\lambda_1 = -2[-2346.41 - (-880.09 - 851.69 - 608.43)] = 12.39$  being smaller than 22.36, the critical value of the  $\chi^2$  distribution at 0.05 significance level with  $(14 + 14 + 18) - 33 = 13$  degrees of freedom, the hypothesis of equal taste parameters could not be rejected.

---

<sup>12</sup> In this paper only the test of process equality of all the three sub-experiments at the same time is reported. However, process equality across each pair of sub-experiments was tested and only a difference in the scales between the construct 'Quality of Connection' and 'Comfort' was found.



Table 8 Separate Tri-modal Models

		Quality of Connection		Comfort		Information	
		Par.	(t-ratio)	Par.	(t-ratio)	Par.	(t-ratio)
Train constant		0.840	(9.67)	0.709	(8.02)	1.047	(10.57)
Car constant		1.429	(4.08)	1.474	(3.99)	1.518	(3.38)
Fare public transport		-0.182	(-6.98)	-0.177	(-6.15)	-0.244	(-7.05)
Time public transport		-0.033	(-7.85)	-0.037	(-8.49)	-0.057	(-9.59)
Price car		-0.137	(-6.02)	-0.147	(-6.42)	-0.140	(-4.97)
Time car		-0.036	(-6.48)	-0.034	(-6.18)	-0.061	(-7.46)
Quality of Connection				0.134	(5.01)	0.185	(6.05)
Comfort		0.098	(3.82)			0.116	(3.82)
Information		0.076	(2.94)	0.072	(2.81)		
Punctuality	<i>On time or up to 3 minutes late</i>	0.263					
	3 to 10 minutes late	0.121	(1.65)				
	10 to 20 minutes late	-0.384	(-5.03)				
Interchanging	<i>0 interchanges</i>	0.357					
	1 interchange with guaranteed connection	-0.027	(-0.36)				
	1 interchange without guaranteed connection	-0.330	(-4.28)				
Frequency	<i>Every 30 minutes</i>	0.621					
	Every 60 minutes	0.158	(2.15)				
	Every 120 minutes	-0.779	(-9.61)				
Cleanliness of train/bus inside	<i>Clean</i>			0.283			
	Floor is sticky, paper lies around on the seats			-0.283	(-5.19)		
Cleanliness of train/bus toilet	<i>Clean</i>			0.311			
	Dirty			-0.171	(-2.26)		
	No toilet			-0.140	(-1.80)		
Seat availability	<i>During the whole trip</i>			0.617			
	During half of the trip			-0.043	(-0.57)		
	No free seat			-0.574	(-7.22)		
Comfort of seats	<i>Comfortable</i>			0.046			
	Not comfortable			-0.046	(-0.86)		
Timetable information at the platform/bus stop	<i>Available</i>					0.122	
	Not available					-0.122	(-1.98)
On-board information in the event of disturbances	Announcement effect					0.097	(1.56)
	Display effect					-0.022	(-0.35)
	Interaction effect					-0.018	(-0.30)
Information at the platform/bus stop in the event of disturbances	Announcement effect					-0.021	(-0.34)
	Display effect					0.101	(1.61)
	Interaction effect					0.035	(0.57)
On-board information concerning connecting trains/busses	Announcement effect					-0.097	(-1.58)
	Display effect					-0.030	(-0.50)
	Interaction effect					-0.038	(-0.62)
Number of observations		1010		945		760	
Number of parameters		14		14		18	
Log-likelihood value		-880.09		-851.69		-608.43	

Table 9 Concatenated Tri-modal Models (All Tri-modal Sub-Experiments)

		Concatenated model (scale parameters allowed to vary)		Concatenated model (scale parameters constrained to be equal)	
		Par.	(t-ratio)	Par.	(t-ratio)
Train constant		0.780	(11.99)	0.850	(16.29)
Car constant		1.361	(6.36)	1.494	(6.79)
Fare public transport		-0.180	(-11.03)	-0.194	(-11.75)
Time public transport		-0.037	(-12.44)	-0.040	(-14.91)
Price car		-0.130	(-10.27)	-0.143	(-10.36)
Time car		-0.039	(-10.69)	-0.041	(-11.48)
Quality of Connection		0.115	(7.56)	0.128	(8.86)
Comfort		0.110	(7.84)	0.128	(9.03)
Information		0.080	(5.88)	0.077	(5.41)
Punctuality	<i>On time or up to 3 minutes late</i>	0.264		0.268	
	3 to 10 minutes late	0.119	(1.62)	0.122	(1.65)
	10 to 20 minutes late	-0.383	(-4.94)	-0.390	(-5.06)
Interchanging	<i>0 interchanges</i>	0.360		0.369	
	1 interchange with guaranteed connection	-0.028	(-0.37)	-0.029	(-0.38)
	1 interchange without guaranteed connection	-0.332	(-4.31)	-0.340	(-4.37)
Frequency	<i>Every 30 minutes</i>	0.619		0.635	
	Every 60 minutes	0.163	(2.17)	0.165	(2.21)
	Every 120 minutes	-0.782	(-9.67)	-0.800	(-9.85)
Cleanliness of train/bus inside	<i>Clean</i>	0.293		0.290	
	Floor is sticky, paper lies around on the seats	-0.293	(-4.79)	-0.290	(-5.29)
Cleanliness of train/bus toilet	<i>Clean</i>	0.326		0.320	
	Dirty	-0.181	(-2.22)	-0.178	(-2.33)
	No toilet	-0.145	(-1.71)	-0.142	(-1.81)
Seat availability	<i>During the whole trip</i>	0.639		0.630	
	During half of the trip	-0.047	(-0.59)	-0.050	(-0.66)
	No free seat	-0.591	(-6.02)	-0.580	(-7.27)
Comfort of seats	<i>Comfortable</i>	0.051		0.051	
	Not comfortable	-0.051	(-0.91)	-0.051	(-0.95)
Timetable information at the platform/bus stop	<i>Available</i>	0.087		0.112	
	Not available	-0.087	(-1.87)	-0.112	(-1.91)
On-board information in the event of disturbances	Announcement effect	0.068	(1.48)	0.083	(1.42)
	Display effect	-0.016	(-0.34)	-0.019	(-0.31)
	Interaction effect	-0.011	(-0.23)	-0.006	(-0.11)
Information at the platform/bus stop in the event of disturbances	Announcement effect	-0.012	(-0.26)	-0.023	(-0.39)
	Display effect	0.072	(1.54)	0.079	(1.34)
	Interaction effect	0.024	(0.51)	0.030	(0.52)
On-board information concerning connecting trains/busses	Announcement effect	-0.068	(-1.43)	-0.079	(-1.35)
	Display effect	-0.026	(-0.56)	-0.033	(-0.57)
	Interaction effect	-0.027	(-0.59)	-0.035	(-0.60)
IV parameter (normalised at RU2)					
Quality of Connection		1.0	(fixed)		
Comfort		1.050	(0.50) <sup>1</sup>		
Information		0.733	(-4.18) <sup>1</sup>		
Number of observations		2715		2715	
Number of parameters		33		31	
Log-likelihood value		-2346.41		-2354.29	

<sup>1</sup> t-statistics are calculated for the hypothesis that the IV parameter equals 1.0.

In the second stage, a concatenated model with scale parameters constrained to be equal was estimated. The results are displayed in Table 9. Again, this model showed the same results concerning significance levels and signs as the separate models. The log-likelihood value was -2354.29. The likelihood ratio test of the hypothesis of equal scale parameters led to  $\lambda_2 = -2[-2354.29 - (-2346.41)] = 15.76$ . This value was larger than 5.99, the critical value of the  $\chi^2$  distribution at 0.05 significance level with  $33 - 31 = 2$  degrees of freedom. Therefore, the hypothesis of equal scale parameters was rejected.<sup>13</sup>

## 6 Conclusion and Discussion

There are many different attributes that influence a passenger's mode choice. Common stated choice analysis only allows for a very limited number of attributes to be included in choice experiments, since too many attributes lead to information overload and biased results. The Integrated Hierarchical Information Integration (HII-I) approach overcomes this problem by summarising similar attributes into constructs. In separate sub-experiments, one construct is described by its attributes while the other constructs are included by summarising construct values. This approach allows for testing of process equality in order to know if the different sub-experiments may be concatenated into an overall model.

The aim of this paper was to apply HII-I to model mode choice in passenger transport and to test if the separate sub-experiments can be concatenated into an overall model. In sub-experiments, respondents had to choose between a train alternative, a (hypothetical) bus alternative, and – if available – a car alternative. All transport modes were described by time and fare/price. In addition, the two public transport modes were described by the three constructs 'Quality of Connection', 'Comfort', and 'Information' either as attributes or as summarising construct values. Choice data were analysed separately for the bi-modal and the tri-modal experiments using multinomial logit (MNL) models. Taste parameters and scale parameters were found to be equal in the bi-modal experiments supporting process equality across sub-experiments. This empirical finding suggests that it is possible to construct separate sub-experiments while arriving at a single concatenated model. In the tri-modal experiments only taste parameters were found to be equal while differences in scale were found making rescaling necessary.

Concerning the selection of attributes, price and total travel time can be derived straight forward from economic theory. Lancaster's new economic theory of consumer behaviour allows for further attributes to characterise the transport modes. However, the question of which attributes to be included in the choice experiments remains an empirical one. Studies which have applied HII-I so far based their selection of attributes and grouping of attributes into constructs on literature studies and expert interviews. Studies which have based their selection of attributes on an empirical pre-study (Bos 2004, Bos et al. 2004, and Molin and van Gelder 2004) applied the HII and not the HII-I approach which means that process equality could not be tested.

---

<sup>13</sup> In this paper only the test of process equality of all the three sub-experiments at the same time is reported. However, process equality across each pair of sub-experiments was tested. For each pair of sub-experiments, the hypothesis of equal taste parameters could not be rejected. The hypothesis of equal scale was not rejected for the pair of sub-experiments 'Quality of Connection' & 'Comfort' but was rejected both for the pair of sub-experiment 'Quality of Connection' & 'Information' and 'Comfort' & 'Information'.

Following the approach proposed by Bos et al. (2002, 2003), this study is based on an empirical pre-study in which travellers were asked to group attributes according to their perceived similarities. Multidimensional Scaling was applied to derive constructs from many attributes of service quality of transport modes. Even though this method is a heuristic method, it probably has contributed to a better selection of constructs. Consequently, respondents probably better understood the relationship between attributes and constructs in the HII-I choice experiments. This may be a reason for the empirical finding of this paper.

Differences in variance and therefore in scale parameters may have various reasons. Possible reasons may be due to differences in the complexity of the three sub-experiments or in heterogeneity across respondents. So far, data were only analysed separately for bi-modal and tri-modal experiments (i.e. car availability). Apart from that, homogeneity across the respondents was assumed but not tested. Not only scale but also taste parameters might vary across different groups of respondents. For example information related attributes might be more important for respondents that rarely use public transport, while they are less important for commuters. This might be the reason of the insignificant information attributes. On the one hand, variance decreases within a sub-experiment with increasing homogeneity of a group. On the other hand, the more data from respondents are split into more homogenous sub-groups, the smaller is the number of available data sets in each group. Separate analysis will therefore be done with ongoing data collection.

MNL models which were used in this paper are based on the assumption of independence of irrelevant alternatives (IIA) stating that the ratio of the probabilities of two alternatives should not be affected by the presence or absence of a third alternative. This can be tested among others by means of the Hausman-McFadden test of IIA (Hausman and McFadden 1984 and Hensher et al. 2005). A nested logit (NL) model allows for partial relaxation of the IIA assumption. Results for the Hausman-McFadden test of IIA for the tri-modal experiments indicate that a NL model might be a more appropriate model. While the two-stage likelihood ratio test is widely used when combining MNL models (often used for the combination of stated and revealed preference data (Hensher and Bradley 1993, Hensher et al. 1998), there are only a few applications for the combination of NL models (Swait and Bernardino 2000). Assumptions for the scale parameters of the different levels need to be taken (Wen 2010).

More recent publications report that the two-stage likelihood ratio test is biased because it is not able to account for potential correlation of observations (Hensher et al. 2008, Hensher 2008). Instead, estimation with an error component mixed logit model is proposed.

## Acknowledgements

We thankfully acknowledge the funding of the project entitled “Qualitäts-simulationsmodell im Regionalverkehr (QUASIMIR)” within the *FHprofUnd* programme of the German Federal Ministry of Education and Research and the support of three local transport companies in Westphalia.

We are also thankful to Eric Molin for general discussion on Integrated Hierarchical Information Integration and the two anonymous reviewers for helpful comments on a previous version of this paper.

## References

- Bech, M. and D., Gyrd-Hansen, 2005. Effects coding in discrete choice experiments. *Health Economics*, 14(10), 1079–1083.
- Bos, D. M., 2004. Changing seats. A behavioural analysis of P&R use (Netherlands TRAIL Research School, Delft, the Netherlands).
- Bos, D., E. J. E. Molin and R., van der Heijden, 2002. User-based definition of higher-order decision constructs to model preferences for P&R transfer points, in L. J., Sucharov, ed., *Urban transport VIII. Urban transport and the environment in the 21st century* (WIT; Computational Mechanics, Southampton, Boston, Billerica), 417–426.
- Bos, I., E. Molin, H., Timmermans and R., van der Heijden, 2003. Cognition and relative importance underlying consumer valuation of park-and-ride facilities. *Transportation Research Record*, 1835, 121–127.
- Bos, I., R., van der Heijden, E., Molin and H., Timmermans, 2004. The Choice of park and ride facilities: an analysis using a context-dependent hierarchical choice experiment. *Environment and Planning A*, 36(9), 1673–1686.
- Hausman, J. and D., McFadden, 1984. Specification tests for the multinomial logit model, *Econometrica*, 52(5), 1219–1240.
- Hensher, D. A., 2008. Empirical approaches to combining revealed and stated preference data: some recent developments with reference to urban mode choice. *Research in Transportation Economics*, 23(1), 23–29.
- Hensher, D. A. and M., Bradley, 1993. Using stated response choice data to enrich revealed preference discrete choice models. *Marketing Letters*, 4(2), 139–151.
- Hensher, D., J. Louviere, and J. D. Swait, 1999. Combining sources of preference data, *Journal of Econometrics*, 89(1-2), 197–221.
- Hensher, D. A., J. M. Rose, and W. H., Greene, 2005. *Applied choice analysis* (Cambridge University Press, Cambridge).
- Hensher, D. A., J. M., Rose, and W. H., Greene, 2008. Combining RP and SP data: biases in using the nested logit ‘trick’ – contrasts with flexible mixed logit incorporating panel and scale effects, *Journal of Transport Geography*, 16(2), 126–133.
- Louviere, J. J., 1984. Hierarchical information integration: a new approach for the design and analysis of complex multiattribute judgment problems. *Advances in Consumer Research*, 11, 148–155.
- Louviere, J. J., D. A., Hensher and J. D., Swait, 2000. *Stated choice methods. Analysis and applications* (Cambridge University Press, Cambridge, Madrid).
- Louviere, J. and H., Timmermans, 1990. Hierarchical Information Integration Applied to Residential Choice Behavior. *Geographical Analysis*, 22(2), 127–144.
- Molin, E. and H., Timmermans, 2009. Hierarchical information integration experiments and integrated choice experiments. *Transport Reviews*, 29(5), 635–655.
- Molin, E. and M., van Gelder, 2008. Freeway Access to Public Transport: A Hierarchical Multimodal Choice Model. *Transportation Research Record*, 2076, 106–113.
- Oppewal, H., J. J., Louviere, and H. J. P., Timmermans, 1994. Modeling hierarchical conjoint processes with integrated choice experiments. *Journal of Marketing Research*, 31(1), 92–105.

- Regulation (EC) No 1371/2007 of the European Parliament and of the Council of 23 October 2007 on rail passengers' rights and obligations, in: *Official Journal of the European Union*, 14–41.
- Swait, J. and A., Bernardino, 2000. Distinguishing taste variation from error structure in discrete choice data. *Transportation Research Part B*, 34(1), 1–15.
- Swait, J. and J., Louviere, 1993. The role of the scale parameter in the estimation and comparison of multinomial logit models. *Journal of Marketing Research*, 30(3), 305–314.
- van de Vyvere, Y., H., Oppewal, and H., Timmermans, 1998. The validity of hierarchical information integration choice experiments to model residential preference and choice. *Geographical Analysis*, 30(3), 254–272.
- Wen, C.-H., 2010. Alternative tree structures for estimating nested logit models with mixed preference data. *Transportmetrica*, 6(4), 291–309.