Transportation Research Part A 83 (2016) 14–29

Contents lists available at ScienceDirect

Transportation Research Part A

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

Multimodal travel groups and attitudes: A latent class cluster analysis of Dutch travelers ☆

Eric Molin [a,*], Patricia Mokhtarian [b,1], Maarten Kroesen [a,2]

[a] Transport Policy Group, Faculty of Technology, Policy and Management, Delft University of Technology, PO Box 5015, 2600 GA Delft, The Netherlands
[b] Transportation Systems Engineering, School of Civil and Environmental Engineering, Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332-0355, United States

ARTICLE INFO

Article history:
Received 3 February 2015
Received in revised form 4 September 2015
Accepted 9 November 2015
Available online 10 December 2015

Keywords:
Multimodality
Mode choice
Mode frequency
Attitude
Latent class cluster analysis
Factor analysis

ABSTRACT

For developing sustainable travel policies, it may be helpful to identify multimodal travelers, that is, travelers who make use of more than one mode of transport within a given period of time. Of special interest is identifying car drivers who also use public transport and/or bicycle, as this group is more likely to respond to policies that stimulate the use of those modes. It is suggested in the literature that this group may have less biased perceptions and different attitudes towards those modes. This supposition is examined in this paper by conducting a latent class cluster analysis, which identifies (multi)modal travel groups based on the self-reported frequency of mode use. Simultaneously, a membership function is estimated to predict the probability of belonging to each of the five identified (multi)modal travel groups, as a function of attitudinal variables in addition to structural variables. The results indicate that the (near) solo car drivers indeed have more negative attitudes towards public transport and bicycle, while frequent car drivers who also use public transport have less negative public transport attitudes. Although the results suggest that in four of the five identified travel groups, attitudes are congruent with travel mode use, this is not the case for the group who uses public transport most often. This group has relatively favorable car attitudes, and given that many young, low-income travelers belong to this group, it may be expected that at least part of this group will start using car more often once they can afford it. Based on the results, challenges for sustainable policies are formulated for each of the identified (multi)modal travel groups.

© 2015 Published by Elsevier Ltd.

1. Introduction

One of the central aims of studying travel behavior is helping policy makers and other stakeholders to develop policies that make travel behavior more sustainable (Banister, 2008; Van Wee et al., 2013). Among other things, this involves reducing car travel and reinforcing travel by public transport and bicycle. To develop those policies it is important to understand the behavioral patterns of travelers. While traditionally research focused on explaining differences in behavior between

☆ An earlier version of this paper appeared in the compendium of papers of the 94th Annual Meeting of the Transportation Research Board, January 2015, Washington, D.C.

* Corresponding author. Tel.: +31 15 278 8510; fax: +31 15 278 2719.
E-mail addresses: e.j.e.molin@tudelft.nl (E. Molin), patmokh@gatech.edu (P. Mokhtarian), m.kroesen@tudelft.nl (M. Kroesen).
[1] Tel.: +1 404 385 1443.
[2] Tel.: +31 15 278 7183; fax: +31 15 278 2719.

http://dx.doi.org/10.1016/j.tra.2015.11.001
0965-8564/© 2015 Published by Elsevier Ltd.

individuals, there is now a growing interest in analyzing the variability of behavior within individuals (e.g., Heinen and Chatterjee, 2015; Jones and Clarke, 1988; Kitamura et al., 2006; Lavery et al., 2013; Schlich and Axhausen, 2003). Of particular interest is answering the question of whether travelers always use the same mode or whether they use different modes, i.e., the extent to which travelers are multimodal. Multimodality in itself can be regarded as a reflection of the deliberate choice process of a traveler who chooses a mode depending on context, as opposed to a habitual traveler who exclusively uses a single mode irrespective of context (e.g. Aarts et al., 1998). From a policy perspective, it is important to understand the nature of the multimodal group in order to help facilitate more of such behavior. Kroesen (2014) found evidence that multi-mode users are indeed more likely than single-mode users to switch over time from one behavioral profile to another.

Multimodality has historically been neglected in studies of travel behavior, in part because of the difficulties in obtaining the requisite data and building more complex models (Clifton and Muhs, 2012). Lately, however, a small but growing body of literature is investigating traits that are associated with multimodal travel behavior. For example, Nobis (2007) finds that multimodality is disproportionately high among adolescents, older people, and residents of population centers. Blumenberg and Pierce (2014) find that lower-income Americans are *less* multimodal than those with higher incomes. Kuhnimhof et al. (2012) find that an observed increase in multimodality among Germans between ages 18 and 29 is accompanied by a declining use of cars within the same group, which helps explain the flattening trend of car use overall. In contrast to most studies that are mainly descriptive in nature, Vij et al. (2013) take a modeling perspective and estimate latent class choice models. They find that different multimodal travel styles are related to long-term travel decisions and travel time sensitivity. In addition, various researchers have stated that travelers who always use the same mode may have incorrect knowledge of other modes. For example, it is often found that people who solely use car overestimate travel times by public transport. In this respect, Van Exel and Rietveld (2009) found that, on average, car travelers' perceptions of public transport travel time exceeded objective values by 46%. In contrast, travelers who also use other modes gain experience with those modes and may to a lesser extent over- or underestimate their performance. In particular, it is suggested that strong car users who also travel by public transport (PT) may develop different attitudes towards PT than strong car users who solely use car (Diana and Mokhtarian, 2009a, 2009b).

However, the mode perceptions and attitudes of multimodal travelers have received little empirical attention so far. Diana and Mokhtarian (2009a) examined the nature of various multimodal clusters, but limited the analysis to a few socioeconomic characteristics. Diana (2010) included a latent variable capturing the multimodality orientation for a given trip, as an explanatory variable for the decision to switch modes for a hypothetical future trip of the same kind. The model incorporated mode-specific cognitive and affective attitudes as other explanatory variables, but did not directly relate the attitudinal variables to the multimodality orientation. The aim of this paper is to address this gap in the literature, in particular by examining the relations between mode-related perceptions and attitudes and belonging to a particular mono- or multimodal travel behavior group in addition to socio-demographic and other structural variables.

Multimodality is defined as the (flexible) use of various modes of transport within a certain time period, whereas intermodality is the use of multiple modes in the course of a single trip (Nobis, 2007). Hence, classifying a traveler as multimodal depends on the time period considered. Blumenberg and Pierce (2014) use a one-day diary for this purpose, Nobis (2007), Buehler and Hamre (2015), and Kuhnimhof et al. (2012) rely on one-week travel diary data, and Vij et al. (2013) use six-week travel diary data. Although six weeks may be considered as a relatively long period for collecting diary data, it cannot be ruled out that car drivers who use public transport less frequently may be misclassified as solo car drivers even with six-week data and especially with shorter-duration data, while as argued above, it is of particular interest to identify this group. Diana and Mokhtarian (2009a, 2009b) consider a longer time period by relying on self-reported frequency of use of various modes over the course of a year, as measured on five-point ordinal scales. Although these kinds of standard survey data are less detailed than diary data, these have the advantage that much longer time periods can be taken into account. In addition, multiple-week travel diary data typically have rather small samples (Buehler and Hamre, 2015), whereas it is easier to realize larger samples with standard survey methods. Following this approach, this paper bases the identification of multimodal travelers on self-reported frequency of driving a car, bicycling, using train and using BTM (bus, tram, metro) measured on an ordinal scale running from 'every day' to 'less than once a year'. Data are collected from a sample of 2548 travelers in the Netherlands.

The approach to distinguish the different travel groups is related to the work of Diana and Mokhtarian (2009a, 2009b), who applied cluster analysis to identify multimodal travel groups. A disadvantage of conventional cluster analysis is, however, that it deterministically assigns travelers to a single cluster, ignoring the possibility of misclassification into the wrong cluster. To overcome this shortcoming, latent class cluster analysis (LCCA) is applied in this paper (e.g. Kroesen, 2014). LCCA is a model-based approach that probabilistically assigns individuals to clusters and thus takes measurement error into account. This analysis technique involves estimation of two models simultaneously. First is a measurement model, which identifies internally homogeneous latent clusters based on simple indicators, which in this study are the frequency of use of various transport modes. Second is a membership model, which predicts the probability of belonging to each of the identified clusters by personal characteristics such as socio-demographic, work-related, and attitudinal variables.

To summarize, the contribution of this paper to the literature is two-fold. First, it is the first paper that identifies (multi)modal travel groups by applying LCCA in which the indicators are simple self-reported mode use frequencies considering a long period of time. Second, the insight into who belongs to multimodal travel groups is extended from relying entirely on structural variables, like socio-demographics, towards including relations with attitudinal variables.

The remainder of the paper is organized as follows. The applied methodology, including the LCCA method, is explained in the following section. This is followed by a results section in which the identified (multi)modal travel groups are described in detail. In the final section, a discussion on the results is provided, in which, among other topics, challenges are formulated for sustainable policies targeting each of the identified groups.

2. Methodology

2.1. Latent class cluster analysis

LCCA is used in this study to reveal single- and multimodal travel groups. The main idea of LCCA is that a discrete latent variable can account for the observed associations between a set of indicators, such that, conditional on the latent class variable, these associations become insignificant (Magidson and Vermunt, 2004; McCutcheon, 1987). This is generally called the assumption of local independence. The goal is to find the most parsimonious model, i.e. with the smallest number of latent classes, which can adequately describe the associations between the indicators.

While cluster research in the transport domain often relies on the ad-hoc and deterministic classification method of conventional cluster analysis to identify homogeneous clusters, LCCA is a model-based clustering technique which probabilistically assigns individuals to classes/clusters. This reduces misclassification biases. Additional benefits over deterministic cluster analysis are that: (1) statistical criteria can be used to judge the optimal number of classes, (2) the significance of the model parameters can be computed and assessed and (3) variables of mixed-scale type can be accommodated (hence, there is also no need to standardize variables) (Vermunt and Magidson, 2002).

Latent class cluster analysis (LCCA) should not be confused with latent class choice models (LCCMs). While both models contain a latent class variable in their formulation, the model structure is fundamentally different. Basically, an LCCM tries to capture taste or preference heterogeneity (i.e. diverse coefficients and/or other parameters) in the way a set of explanatory variables predicts a categorical outcome (e.g. a choice). Specifically, it is assumed that the heterogeneity can be expressed in a discrete number of ways (the latent classes) (see Vij et al., 2013 for an LCCM application to identifying multimodal travel groups). In LCCA, on the other hand, it is assumed that a latent class variable directly underlies the responses on a set of indicators. Hence, an LCCA is conceptually similar to a (confirmatory) factor model, which is later explained in more detail in the subsection titled "The model". Very simply stated, the key difference is that in LCCM, classification into segments is based on differences in the influence of explanatory variables on observed choices, while in LCCA, classification into clusters is based on differences in simple indicators (as in factor analysis).

LCCA has been applied in the transportation domain, however, the number of applications is rather limited. Without pretending to be exhaustive, the following studies have applied LCCA in the transport domain: Goulias et al. (2003) distinguished groups of travelers that travel alone or with others; Goulias and Henson (2006) distinguished groups of altruists and egoists in activity participation and travel; Deutsch and Goulias (2013) identified differences and similarities among different social network types; Kim et al. (2005) distinguished travel groups based on mode frequencies as observed in two-day diaries; Beckman and Goulias (2008) identified clusters of travel behavior of immigrants; Bamberg (2013) identified clusters of stages of behavioral change with respect to car use; and Depaire et al. (2008), Kaplan and Prato (2013) and Ona et al. (2013) all applied LCCA to identify clusters of injury severity of persons involved in road accidents.

2.2. Data collection and sample

Respondents were recruited from an existing large national panel in the Netherlands owned by Intomart GfK. This panel involves about 110,000 persons that on a regular basis are requested to fill out questionnaires, mainly for marketing research purposes. Intomart GfK claims that this panel is representative for the Dutch population in terms of regular background variables such as gender and age. A sample from all panel members older than 18 years was drawn from this panel; it is only semi-random in that public transport users were slightly oversampled to assure that sufficient respondents are familiar with public transport, but in all other respects those who were invited to take the survey were selected at random. Of course, since participation was voluntary, the usual non-response biases can be expected to apply. Persons who do not have a driver's license are not asked to report on the frequency of driving a car and to complete the car perception questions, which are therefore not included in the analysis presented in this paper. As a consequence, the subsample on which the analyses in this paper are based totals 2548 respondents.

The last column of Table 4 (provided below, in the "Results" section) presents the sample characteristics of all variables included in the analysis. On the socio-demographic variables for which population distributions are readily available (Statistics Netherlands), the sample distributions are fairly similar to the distributions in the Dutch population with respect to gender, paid job-holding, and residence in four large cities, while the sample somewhat underrepresents students, single-person households, and higher-educated people. Although the sample fairly well represents the Dutch population with respect to many socio-demographic variables, it should be noted that we included only persons with a driver's license, that PT travelers are slightly overrepresented, and that the latter group does not represent those PT users without a driver's license.

2.3. Measurements

As indicators of the multimodal travel clusters, we selected the frequency of use of each of the following four modes of transport: car, bicycle, train and BTM (bus, tram or metro). Train is distinguished from BTM as in the Netherlands the train network is mainly designed to connect cities (and consequently also connects all towns along those corridors) and is therefore especially used for intercity travel, hence, for traveling relatively long distances. The bus, tram and metro networks, on the other hand, are mainly designed to connect locations within cities and to connect cities with their neighboring municipalities, hence, BTM services are mainly used for traveling relatively short distances. Long distance bus lines are not available in the Netherlands. Bus, tram and metro are treated as a single category because metro services are only available in the two largest cities and some of their neighboring municipalities, tram services are only available in the four largest cities, whereas bus services are available in any municipality in the Netherlands. Respondents were asked to report how often they use each of the four distinguished modes on an eight-point ordinal scale ranging from (1) (almost) every day to (8) less than once a year. As shown in Table 4 below, a range of socio-demographic and work-related variables will be explored as covariates in our model to explain the membership of each of the identified multimodal traveler clusters.

In addition, we will explore the relations of car availability and city size with cluster membership. Unlike the socio-demographic variables, for which the causal direction of influence on the travel clusters is relatively clear, this is less the case for the variables city type and car availability. It may be argued that at least some travelers make choices that influence these variables because they prefer a certain travel style. Hence, self-selection may play a role with respect to these variables. For example, those who prefer to travel by train may choose a residence near a main railway station. Likewise, those who prefer to travel by car will purchase at least one car. To avoid any issues of endogeneity, namely that a predictor variable is (partly) dependent on the variable it aims to predict, the variables *city size* and *car availability* are included as inactive covariates in our model as opposed to the other covariates which are considered active. The inactive covariates did not influence the cluster probabilities, though the distribution of their categories across the clusters can be calculated. Moreover, the effects of the active covariates are obtained by controlling for other active covariates in the model, while those effects are not being controlled for the inactive covariates city size and car availability (see next subsection for a more detailed explanation of active and inactive covariates).

Furthermore, as discussed in the Introduction, one of the main interests of this paper is to explore the role of mode perceptions and attitudes on multimodal cluster membership. These were measured by formulating statements with which respondents could agree or disagree to a certain extent. Five-point Likert scales ranging from (1) completely agree to (5) completely disagree were used as response scales for all statements.

First, we measured how travelers perceive traveling by each mode. Statements started with 'How do you perceive driving a car/bicycling/traveling by train/traveling by bus?', which was asked separately for each of the four modes. For each mode, this was followed by 7 terms: *status giving*, *environmentally friendly*, *relaxing*, *comfortable*, *time saving*, *flexible*, and *pleasant*.

To examine whether any latent variables underlie the responses to the seven perception items, a common factor analysis was conducted. More specifically, the principal axis factoring method with oblimin rotation was chosen for this. To arrive at the same factor structure for each mode, the responses for each of the seven perception items were pooled and stacked mode-on-mode, that is, placed underneath each other. Thus, the analysis was conducted on the seven perception variables irrespective of mode. Two factors were identified in this analysis: "pleasant", on which the variables *pleasant*, *relaxing* and *comfortable* loaded strongly, and "convenient", on which the variables *time saving* and *flexible* loaded strongly. Factor scores were saved, standardized, then split back up by mode and added to the database as mode-specific variables. The items *environmentally friendly* and *status giving* did not load highly on the identified factors and were therefore excluded from the factor analyses. These two items were also standardized and added to the data file. That the scores are broken down per mode after standardization has the advantage that the means express how a mode is perceived relative to the overall mean across all modes. For example, as will become clear in the results section, 'car driving is pleasant' has a positive mean value, whereas 'traveling by bus is pleasant' has a negative value, indicating that overall driving a car driving is perceived as more pleasant than traveling by bus. In contrast, if factor scores were developed for each mode in isolation, the mean for every mode would be zero (since scores are standardized). In total, 16 mode-specific perceptions were created, four for each mode.

To examine the travelers' mode attitudes, a battery of 19 statements was formulated mainly about public transport, while a few were related to car driving. In a factor analysis we identified seven underlying factors: *PT transfer acceptability*, *PT waiting acceptability*, *car inexpensive*, *PT timeliness*, *PT seat availability*, *PT planning ease*, and *PT inexpensive.* A list of all statements and the constructed factors is presented in Table 1.

2.4. The model

Let y_{it} represent the response of subject i on indicator variable t and let m denote a particular category of y_{it}. In the present case, there are 4 indicator variables (respectively the use of the car, bicycle, train and BTM) with 8 categories each (ranging from (1) (almost) every day to (8) less than once a year), so $t \in \{1,2,3,4\}$ and $m \in \{1,2,\ldots,8\}$. In latent class analysis, it is assumed that there is a single nominal latent variable which can explain the associations between the indicator variables. Let this latent nominal variable be defined as x with K categories ($x \in \{1,2,\ldots K\}$). The categories of this nominal latent variable are typically called classes (or clusters). Each subject i is assumed to have a certain probability of belonging to each class (or cluster), which depends on the characteristics of the individual. These characteristics are called covariates and may for

Table 1
Perception and attitude measurements and constructed factors.

Factor labels	Statements
Mode perceptions (mode specific)	
Pleasant[a]	Pleasant (+) Relaxing (+) Comfortable (+)
Convenient	Time saving (+) Flexible (+)
Single item	Status giving (+)
Single item	Environmentally-friendly (+)
Mode attitudes	
PT transfer acceptability	Traveling by PT is a lot less attractive if it includes a transfer (−) Transferring in PT is not annoying (+) I am annoyed by waiting for PT (−)
PT waiting acceptability	If you are protected from the weather, waiting for PT does not matter (+) As long as you are well informed about waiting times, it does not matter to wait a little longer due to delays of bus, tram, train or metro (+) Services (like shops) at a railway station make waiting more pleasant (+)
Car inexpensive	Car fuel is expensive (−) Travel costs of the car are negligible (+) If I travel by car I am conscious of the fuel costs that are incurred (−)
PT timeliness	Traveling by PT makes you worry about being on time at your destination (−) If you travel by train it is less likely that you are late than if you travel by car (+)
PT seat availability	It is annoying not knowing beforehand whether you will be able to find a seat in PT (−) It is no problem not having a seat in the train from time to time (+) Usually it is possible to find a seat in the train (+)
PT planning ease	Planning a trip with PT is complicated (−) It is difficult to use the PT-chipcard in PT (−) If, because of disruptions in PT, you have to travel an alternative route it is difficult to find out how to travel (−)
PT inexpensive	Usually PT is more expensive than car (−) Traveling by PT should be cheaper (−)

[a] The non-bold terms in the first column denote the labels of the identified factors in the factor analysis; the second column provides the high loading statements, with the highest loading mentioned first; a (+) denotes a positively formulated statement, while (−) denotes a negatively formulated statement.

example be socio-demographic variables and attitudes. The covariates are assumed to predict class membership. In other words, a subject's class membership is conditional on the included covariates. Let z_{ir} represent the response of subject i on covariate variable r and $\boldsymbol{z}_i$ (in bold face) the complete vector of covariate values of subject i.

Assuming that the local independence assumption (that correlations between indicators can be explained by the latent clusters) indeed holds, the probability of observing a complete response pattern (a particular sequence of responses on the four indicators) can be formulated as follows (following the notation in Vermunt and Magidson (2005)):

$$P(y_{i1} = m_1, y_{i2} = m_2, y_{i3} = m_3, y_{i4} = m_4) = \sum_{x=1}^{K} P(x|\boldsymbol{z}_i) \prod_{t=1}^{4} P(y_{it} = m_t|x). \tag{1}$$

As can be seen from this formulation, the latent class cluster model basically consists of two kinds of probabilities: the probabilities of belonging to a certain latent class given an individual's covariate values, the $P(x|\boldsymbol{z}_i)$'s (or mixing weights), and the probabilities of particular responses on the indicator variables given latent class membership, the $P(y_{it} = m_t|x)$'s (or mixture densities). Given that the probability of observing a complete response pattern is determined by a (weighted) product of individual response probabilities, i.e. one for each y_{it}, this formulation implies that the indicator variables are mutually independent given that the individual belongs to a certain latent class. This is typically called the local independence assumption.

In the present case, the indicator variables represent ordinal variables. To model these indicators an ordinal logit model is estimated, in which the response probabilities can be parameterized using the multinomial logit formula:

$$P(y_{it} = m|x) = \frac{\exp(\beta^t_m + \beta^t_{mx})}{\sum_{m'=1}^{M} \exp(\beta^t_{m'} + \beta^t_{m'x})}, \tag{2}$$

with the additional restriction that:

$$\beta^t_{mx} = \beta^t_x * m. \tag{3}$$

Hence, the restricted association term (3) yields an adjacent-categories ordinal logit model for response variable y_{it} (Agresti, 2002). Note that the formulation in (2) implies that for each indicator variable t a set of category-specific parameters (the β^t_m's) and a set of class-specific parameters is estimated (the β^t_x's). The category-specific parameters (the β^t_m's) can be interpreted as alternative-specific constants, which indicate travellers' 'baseline' preference with respect to specific categories (while accounting for the effects of the classes). The class-specific parameters (the β_x^t's), on the other hand, indicate the unattractiveness of mode t for individuals belonging to class x: the greater the value of β_x^t, the greater the value of $\beta_{mx}^t = \beta_x^t * m$, and the less frequently individuals belonging to class x are likely to use mode t. For identification purposes, all parameters are restricted using effect coding ($\sum_{m=1}^{M} \beta_m^t = 0$ and $\sum_{x=1}^{K} \beta_x^t = 0$).

Finally, the probability of belonging to a particular latent class x given person i's covariate values $\boldsymbol{z}_i$ is also parameterized by means of a multinomial logit model:

$$P(x|\boldsymbol{z}_i) = \frac{\exp\left(\gamma_x + \sum_{r=1}^{R} \gamma_{xr} * z_{ir}\right)}{\sum_{x'=1}^{K} \exp\left(\gamma_{x'} + \sum_{r=1}^{R} \gamma_{x'r} * z_{ir}\right)}. \tag{4}$$

Hence, for each category of the latent class variable (x) one intercept (γ_x) and a set of regression parameters will be estimated (the γ_{xr}'s). Again, these parameters are restricted using effect coding ($\sum_{x=1}^{K} \gamma_x = 0$ and $\sum_{x=1}^{K} \gamma_{xr} = 0$).

The latent class cluster model in Eq. (1) can also be represented graphically. This is done in Fig. 1. In our specific application, the categories of the latent class variable represent different (latent) travel patterns. Each of these is associated with particular levels of using each of the four modes. The model assumes that the latent travel patterns can effectively explain the associations between the indicators. In this sense, a latent class cluster model is similar to a (confirmatory) factor model. The only difference is that the underlying latent factor is assumed to be categorical instead of continuous in nature. Yet, both models can be regarded as measurement models, aiming to (indirectly) measure an underlying latent construct via the use of multiple indicators.

By adding (active) covariates the model is expanded with a structural part. As mentioned before, these covariates are assumed to predict class (or in this case more specifically travel pattern) membership. For example, younger subjects will likely have a higher probability of belonging to a travel pattern in which the bicycle is used often. Hence, it is likely that age will influence class membership. In addition to socio-demographic variables, this study explores a range of mode-related perceptions and attitudes as explanatory variables of travel pattern membership. Finally, two inactive covariates are included as well (city size and car availability). As mentioned in the previous section, these are not allowed to be active predictors of class membership in the statistical model, but are only included to additionally profile the classes.[3]

2.5. Model estimation

To determine the appropriate number of latent classes, models were first estimated without active covariates in order to assess only the measurement part of the model. Table 2 presents the fit of consecutive models starting with a model with one class up to a model with ten classes. The dedicated software package Latent Gold was used to estimate the models (Vermunt and Magidson, 2005).

Various approaches exist to evaluate which number of latent classes is statistically optimal. A common approach is the chi-squared goodness-of-fit test (based on, for example, the likelihood-ratio chi-squared statistic L^2), in which the observed cell frequencies are compared with the model-implied cell frequencies for the various response patterns under the null hypothesis that the difference between the two sets of frequencies is zero. However, if there are many possible response patterns, which is the case here (with $8 \times 8 \times 8 \times 8$ = 4096 possible patterns), many observed cell frequencies will be zero and the chi-squared statistic will no longer approximate a chi-squared distribution. As can be observed from Table 2, based on the likelihood-ratio chi-squared statistic (L^2) all models would in fact be rejected.

The most typical approach to assessing model fit in the case of sparse data is the use of an information criterion, which weighs both model fit and parsimony (i.e. the number of estimated parameters). In the context of latent class analysis, the Bayesian information criterion (BIC) has been shown to perform well (Nylund et al., 2007). However, the guideline is to select the solution for which the BIC is lowest, and as the results in Table 2 show, the BIC values consistently decrease as K increases, indicating that the optimal model (based on this criterion) is one with at least 10 classes.

Since such a solution would be difficult to interpret and also too complicated to communicate, we relied on multiple local (instead of a global) measures of model fit to determine the optimal number of classes, namely the bivariate residuals. These residuals are estimates of the improvement in model fit (L^2) if a direct effect between the indicators were included (thus, if the local independence assumption is relaxed) (Vermunt and Magidson, 2005). Since these measures are chi-squared distributed, values greater than 3.84 (with one degree of freedom) indicate that significant covariation remains between a pair of indicators.

Based on this criterion (see the last six columns in Table 2), the 5-class model was found to be the first which could adequately account for the observed associations between the indicators (only one just-significant residual remained between the car and the train indicator). Based on these results the 5-class solution was selected as optimal.

After establishing the so-called measurement part of the model, the individual background characteristics and attitudes (Table 1) were included as active covariates. To assess how well one can predict class memberships based on the covariates,

[3] For each class the profile (distribution) of the inactive covariates is estimated based on the model's posterior probabilities.

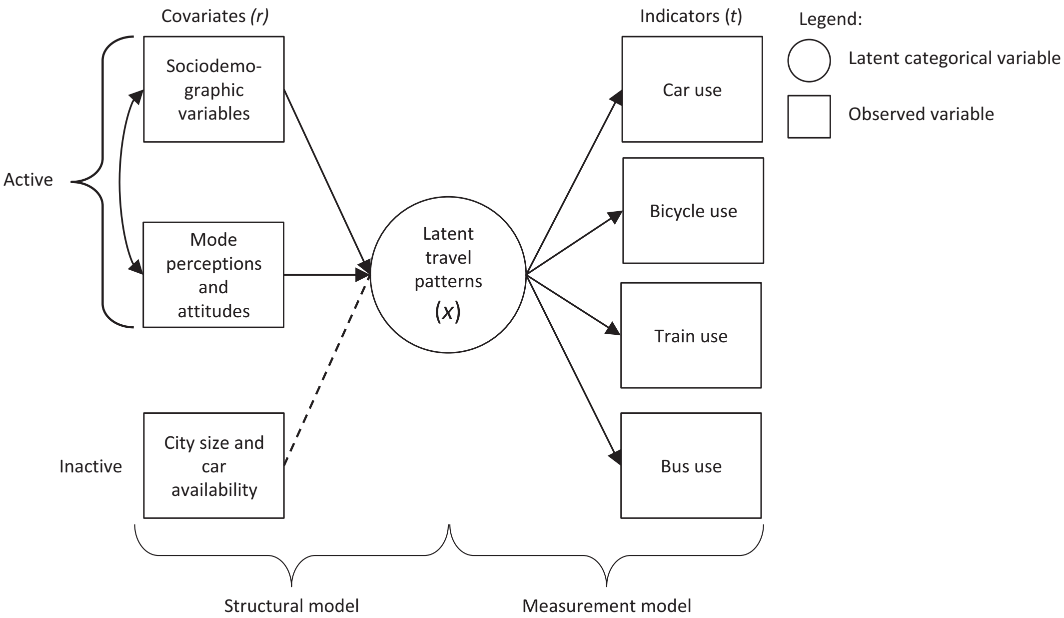

Fig. 1. Graphical representation of the latent class cluster model with covariates.

Table 2
Model fit of the latent class cluster analysis models.

No. of classes	Npar	LL	L^2	df	BIC(LL)	p-value	Bivariate residuals					
							Car-train	Car-BTM	Car-bicycle	Train-BTM	Train-bicycle	BTM-bicycle
1	27	−16382.8	4475.5	2210	32973.9	0.00	179.2	211.5	365.3	638.3	29.5	1.9
2	32	−15896.9	3503.7	2205	32040.7	0.00	15.0	0.5	108.2	95.1	17.3	56.5
3	37	−15727.4	3164.7	2200	31740.2	0.00	0.0	2.4	3.7	53.0	0.0	0.7
4	42	−15645.9	3001.7	2195	31615.8	0.00	0.4	1.5	7.2	6.5	0.6	0.3
5	47	−15584.0	2877.8	2190	31530.4	0.00	4.0	0.5	0.2	2.3	1.1	0.1
6	52	−15537.3	2784.4	2185	31475.7	0.00	3.2	0.0	0.2	3.3	1.9	0.3
7	57	−15498.6	2707.1	2180	31436.9	0.00	1.4	0.2	1.1	1.9	0.3	0.8
8	62	−15476.2	2662.2	2175	31430.5	0.00	0.8	0.2	0.6	0.9	1.4	0.3
9	67	−15455.2	2620.3	2170	31427.2	0.00	0.5	0.0	0.0	1.6	1.9	0.4
10	72	−15431.5	4475.5	2165	31418.3	0.00	0.7	0.0	0.0	0.7	1.1	0.3

LL = final log-likelihood of the model.
BIC(LL) = Bayesian information criterion (based on log-likelihood).
Npar = number of parameters.
L^2 = likelihood-ratio chi-squared statistic.
df = degrees of freedom.

Latent Gold provides several pseudo R-squared statistics. In the model with covariates, the standard R-squared measure (a qualitative variance-based measure (Magidson, 1981)) yields a value of 0.29, indicating that a substantial portion of the variability in class membership is explained by the included covariates. The parameters of the final model are presented in Table 3. Note that no parameters are estimated for inactive covariates, which is the reason why city size and car availability are not included in this table. For the interpretation of the results we mainly rely on the within-cluster distributions of the indicators and covariates as presented in Table 4.

3. Results

3.1. Five identified multimodal classes

The top of Table 4 (indicators) presents the distributions of the frequencies of use of each mode within each cluster. Though the model is based on all eight frequency categories, to ease interpretation the results were joined into four categories only. We identify and label each cluster based on these mode use shares. The first term in the cluster label is based

Table 3
Parameters and z-values of the estimated latent class cluster analysis model with covariates.

Prediction of the indicators (the β_x^t's in the measurement model)	Cl.1	z-value	Cl.2	z-value	Cl.3	z-value	Cl.4	z-value	Cl.5	z-value	Wald	p-value
Car use frequency	−0.590	−8.934	0.936	15.959	0.191	3.805	−1.241	−14.210	0.706	11.787	279.460	0.000
Bicycle use frequency	0.442	11.505	−0.873	−12.430	−0.332	−6.185	0.657	15.785	0.107	2.299	302.299	0.000
Train use frequency	−0.557	−7.269	−0.828	−10.999	1.085	8.811	1.797	12.773	−1.496	−15.958	275.972	0.000
Bus use frequency	−0.146	−2.570	−0.360	−6.977	0.859	10.468	1.197	13.323	−1.551	−10.625	259.770	0.000

Intercepts (the β_m^t's)	Car use		Bicycle use		Train use		Bus use	
Wald	539.011		734.796		513.161		388.396	
p-value	0.000		0.000		0.000		0.000	
	Overall	z-value	Overall	z-value	Overall	z-value	Overall	z-value
(Practically) every day	0.671	7.439	1.379	16.737	−4.482	−15.770	−3.964	−15.064
5–6 days per week	1.169	12.023	0.589	6.245	−2.024	−11.075	−2.171	−14.076
3–4 days per week	1.776	16.962	1.176	14.837	−0.808	−6.972	−0.322	−2.480
1–2 days per week	1.817	20.633	1.217	19.595	0.187	2.279	0.820	6.585
1–3 days per month	0.588	8.286	0.478	7.902	2.024	20.940	1.779	17.788
6–11 days per year	−1.334	−12.301	−0.800	−9.049	2.230	12.868	1.692	15.711
1–5 days per year	−2.128	−15.735	−1.935	−15.226	2.370	12.182	1.573	13.069
Less than 1 day per year	−2.559	−15.203	−2.105	−15.340	0.502	3.397	0.593	4.498

Prediction of latent class membership (the structural model)	Cl.1	z-value	Cl.2	z-value	Cl.3	z-value	Cl.4	z-value	Cl.5	z-value	Wald	p-value
Intercept (γ_{x0})	0.315	0.991	1.010	2.162	−1.858	−4.051	−2.693	−6.064	3.226	8.098	111.699	0.000
Covariate coefficients (the γ_{xr}'s)												
Male	0.120	2.039	−0.273	−3.084	−0.181	−2.270	0.018	0.250	0.316	3.881	23.0	0.000
Age	0.010	2.077	−0.004	−0.620	0.028	4.113	0.014	2.264	−0.048	−6.846	60.6	0.000
Up to middle vocational	−0.207	−2.328	−0.429	−3.797	0.490	3.494	0.533	4.197	−0.386	−3.241	59.5	0.000
Higher sec. & vocational	−0.215	−2.748	−0.161	−1.728	0.017	0.128	0.369	3.154	−0.010	−0.108		
University	0.422	3.917	0.590	4.698	−0.506	−2.375	−0.902	−4.633	0.396	3.043		
Minimum income	−0.651	−1.903	1.212	4.705	0.360	0.997	−1.581	−2.550	0.660	2.532	52.9	0.000
Minimum – modal income	−0.047	−0.334	−0.213	−1.667	0.141	0.928	0.542	2.393	−0.423	−3.034		
Above modal income	0.253	1.790	−0.661	−4.559	−0.043	−0.252	0.726	3.150	−0.276	−1.881		
Missing value income	0.444	3.085	−0.338	−2.329	−0.459	−2.672	0.313	1.327	0.040	0.285		
Single person in household	0.119	1.105	0.272	2.110	−0.415	−2.446	−0.225	−1.622	0.248	1.763	41.4	0.000
Couple with children	0.051	0.547	−0.193	−1.636	0.239	1.734	−0.023	−0.200	−0.074	−0.594		
Couple without children	−0.253	−2.465	−0.324	−2.266	0.592	4.060	0.214	1.811	−0.229	−1.746		
Other household composition	0.083	0.540	0.245	1.315	−0.417	−1.524	0.033	0.174	0.055	0.313		
Paid work	0.278	1.522	−0.432	−1.716	0.312	1.051	−0.065	−0.232	−0.095	−0.398	29.3	0.000
Student	−0.169	−0.699	0.442	1.654	−0.786	−1.710	−0.377	−1.075	0.890	3.865		
Else (not paid work or student)	−0.110	−0.668	−0.011	−0.053	0.474	1.770	0.442	1.896	−0.796	−3.867		
Number of working days per week	−0.007	−0.140	−0.016	−0.177	−0.274	−3.539	0.340	4.337	−0.043	−0.546	22.8	0.000
Fixed work location	−0.192	−2.796	0.057	0.601	0.467	4.178	−0.269	−3.397	−0.062	−0.705	24.8	0.000
Driving a car is pleasant	0.288	3.688	−0.603	−7.091	−0.166	−1.765	0.674	6.944	−0.194	−2.094	82.8	0.000
Driving a car is convenient	0.144	1.694	−0.228	−2.242	0.158	1.418	0.090	0.900	−0.164	−1.481	11.0	0.026
Driving a car is environmentally friendly	0.267	2.909	−0.048	−0.428	−0.332	−2.989	0.032	0.293	0.081	0.677	14.0	0.007

(*continued on next page*)

Table 3 (*continued*)

Prediction of latent class membership (the structural model)	Cl.1	*z*-value	Cl.2	*z*-value	Cl.3	*z*-value	Cl.4	*z*-value	Cl.5	*z*-value	Wald	*p*-value
Bicycling is pleasant	−0.512	−6.942	0.805	7.045	0.690	5.894	−0.467	−5.256	−0.515	−5.386	112.1	0.000
Bicycling is convenient	−0.298	−4.145	0.586	6.302	0.391	4.235	−0.606	−7.127	−0.073	−0.756	83.2	0.000
Traveling by train is pleasant	0.116	1.415	0.384	3.757	−0.189	−1.778	−0.363	−3.747	0.052	0.478	27.9	0.000
Traveling by train is environmentally friendly	0.214	2.374	0.130	1.170	−0.331	−3.091	−0.180	−1.803	0.167	1.407	18.6	0.001
Traveling by bus is pleasant	−0.149	−1.921	−0.029	−0.325	0.054	0.527	−0.128	−1.353	0.252	2.514	9.6	0.047
PT waiting acceptability	0.006	0.090	−0.035	−0.426	−0.367	−4.531	−0.176	−2.380	0.572	6.190	47.8	0.000
PT timeliness (no worries being on time)	−0.019	−0.299	0.190	2.522	−0.090	−1.080	−0.173	−2.173	0.092	1.089	11.3	0.023
PT planning is easy	0.096	1.560	0.470	5.791	−0.424	−5.498	−0.292	−4.063	0.150	1.910	69.8	0.000
PT is inexpensive	−0.006	−0.099	0.247	3.263	−0.213	−2.679	−0.159	−2.197	0.131	1.666	21.0	0.000

Table 4
The within-cluster distributions of indicators and covariates.

	Cluster 1 CAR MM	Cluster 2 BIKE MM	Cluster 3 BIKE + CAR	Cluster 4 CAR MOSTLY	Cluster 5 PT MM	Sample total
Indicators						
Car use						
• 5–7 days a week	**55%**	3%	19%	**82%**	6%	34%
• 1–4 days a week	42%	**41%**	**64%**	18%	**53%**	**43%**
• At least once a month	3%	22%	13%	0%	21%	11%
• A few days a year or less	0%	34%	4%	0%	20%	12%
Bicycle use						
• 5–7 days a week	15%	**84%**	**60%**	8%	32%	**41%**
• 1–4 days a week	**45%**	15%	33%	**37%**	**47%**	34%
• At least once a month	21%	1%	5%	23%	13%	13%
• A few days a year or less	19%	0%	1%	32%	7%	12%
Train use						
• 5–7 days a week	4%	9%	0%	0%	**40%**	9%
• 1–4 days a week	16%	24%	0%	0%	35%	15%
• At least once a month	36%	**36%**	4%	1%	19%	22%
• A few days a year or less	**44%**	30%	**95%**	**99%**	6%	**54%**
BTM use						
• 5–7 days a week	1%	2%	0%	0%	38%	6%
• 1–4 days a week	19%	27%	2%	1%	**48%**	19%
• At least once a month	31%	33%	8%	4%	11%	20%
• A few days a year or less	**49%**	**38%**	**90%**	**95%**	2%	**55%**
Covariates						
Personal characteristics						
Male	**59%**	39%	42%	**58%**	**56%**	**51%**
Female	41%	**61%**	**58%**	42%	44%	49%
Age (in years)	48	45	**52**	45	34	46
Household composition						
• Single	23%	**39%**	13%	17%	**33%**	26%
• Couple with kids	**45%**	34%	**47%**	37%	26%	**39%**
• Couple without kids	24%	18%	37%	**39%**	25%	28%
• Other household types	8%	9%	3%	7%	15%	8%
Level of education						
• Up to middle vocational	37%	29%	**64%**	**50%**	22%	40%
• Higher sec. & vocational	**39%**	**40%**	31%	46%	**54%**	**41%**
• University	24%	30%	5%	5%	24%	19%
• Missing value	25%	20%	19%	19%	31%	23%
Income						
• Minimum	2%	13%	3%	0%	13%	6%
• Minimum – modal	26%	**36%**	**41%**	33%	26%	32%
• Above modal	**47%**	31%	37%	**47%**	**30%**	**39%**
Work status						
• Paid job	**69%**	**62%**	**54%**	**80%**	**67%**	**66%**
• Student	3%	8%	1%	1%	23%	6%
• Other	29%	29%	46%	19%	10%	28%
Commuting days per week	2.7	2.5	1.9	**3.5**	2.7	2.6
Fixed working location (not home)	37%	40%	41%	**46%**	43%	41%
City size <Inactive>						
• The big four city	14%	22%	6%	6%	24%	14%
• Middle size cities	23%	33%	18%	17%	28%	24%
• Small cities and villages	**64%**	**45%**	**76%**	**77%**	**48%**	**62%**
Car availability <Inactive>						
• Car always available	**82%**	36%	**70%**	**92%**	32%	**64%**
• After arrangements	16%	26%	26%	8%	**38%**	22%
• No car available	2%	**38%**	4%	0%	30%	15%
Perceptions						
Driving a car is pleasant	0.52	−0.06	0.29	**0.70**	0.30	0.34
Driving a car is convenient	0.93	0.62	0.90	**1.02**	0.79	0.85
Driving a car is environmentally friendly	−0.91	−1.21	−0.99	**−0.86**	−1.01	−1.00
Bicycling is pleasant	0.15	**0.87**	0.80	0.07	0.16	0.43
Bicycling is convenient	−0.07	**0.64**	0.44	−0.31	0.19	0.19
Traveling by train is pleasant	**0.01**	0.40	−0.09	−0.56	0.00	−0.02

(*continued on next page*)

Table 4 (*continued*)

	Cluster 1 CAR MM	Cluster 2 BIKE MM	Cluster 3 BIKE + CAR	Cluster 4 CAR MOSTLY	Cluster 5 PT MM	Sample total
Traveling by train is environmentally friendly	0.26	**0.41**	0.04	−0.05	0.26	0.20
Traveling by bus is pleasant	−0.81	**−0.61**	−0.70	−1.09	−0.67	−0.77
Mode-specific attitudes						
PT waiting acceptability	−0.06	0.15	−0.14	−0.36	**0.27**	−0.03
PT timeliness (no worries being on time)	0.00	**0.27**	−0.17	−0.33	−0.01	−0.03
PT planning is easy	0.08	**0.38**	−0.46	−0.34	0.13	−0.02
PT is inexpensive	0.06	**0.26**	−0.22	−0.28	0.06	0.00

For categorical variables, the highest frequency within each cluster is in bold face. For continuous variables, the highest mean value across all clusters is in bold face.

on the mode which is most frequently used by that cluster. A suffix MM of multi-modal is added if a considerable share of the cluster members uses PT at least once a month. The suffix MM rather than PT is used to incorporate the fact that all clusters make use of the bicycle to some extent. The five identified clusters were labeled 'CAR MM', 'BIKE MM', 'BIKE + CAR', 'CAR MOSTLY' and 'PT', which are described in detail in the following. For each cluster, we first describe its travel behavior as measured by the reported travel frequencies (indicators). This is followed by a description of the personal characteristics and attitudes that influence cluster membership, which is based on the within-cluster distributions of the covariate variables that are presented in the second part of Table 4 (covariates).

3.2. The CAR MM group

The first group is labeled CAR MM, which has a share of 27% in the sample.[4] Of all the modes, this group uses car most often: 97% uses car at least once a week, while 55% uses car (almost) every day. Also bicycle is used relatively often: 60% uses it at least once a week. PT is used less often, but there is a fair share: 56% uses train and 51% uses BTM (bus, tram, metro) at least once a month.

This group is characterized by the following personal characteristics. It has a relatively large share of males (59%) and consequently a smaller share of females (41%). The average age is slightly higher (48) than the sample average (46) and the propensity to belong to it increases with age as indicated by the positive age parameter (see Table 3). With respect to household composition, the 'couples with children' category is overrepresented (45% versus 39% in the sample). The share of university graduates is relatively high (24%), and related to this, more persons belong to the 'above modal' income group.

With respect to work status and city size, this group does not stand out. The group however clearly differs with respect to car availability: a much higher percentage (82%) than in the overall sample (64%) always has a car available, while only a very small share does not have a car available. The average number of commuting days per week is slightly higher than average and of all groups, this group has the lowest share with a fixed working location.

Compared to the sample as a whole, the perception variables show that the CAR MM group more than average: (i) derives pleasure from car driving, (ii) perceives car driving as more convenient, and (iii) is slightly less negative about the car being environmentally friendly. This group is not favorable towards bicycling: it clearly derives less pleasure from bicycling and also perceives bicycling as less convenient. On the other hand, its PT perceptions and attitudes are about average. It slightly more than average perceives train traveling as environmentally friendly, while it derives less pleasure from traveling by bus. This group scores about average on the PT-specific variables 'waiting time acceptability', 'PT timeliness', 'planning is easy' and 'PT is inexpensive'.

To summarize, the car MM group has a positive attitude towards car, a negative attitude towards bicycling and an about-average attitude towards PT. Hence, its attitudes are congruent with its travel behavior, although we do not make any inference about the direction of causality, i.e. whether a positive experience with the mode improved users' attitudes toward it, or whether those with favorable attitudes at the outset were naturally inclined to use the mode more often.

3.3. The BIKE MM group

The second group is labeled BIKE MM, and has a share of 24%. Of all groups, this group uses bicycle most often: 84% (almost) every day. Like the CAR MM group, also this group is not afraid to use PT and its use is even a little higher: 70% uses train and 62% uses BTM at least once a month. Of all groups, this group clearly uses car the least: only 3% uses car (almost) every day and only 44% once a week (or more).

Of all groups, this one has the largest share of females (61%). Its average age is slightly lower than the sample average (45 vs. 46), and the propensity to belong to it decreases with age, although the age parameter is not statistically significant.

[4] Readers are reminded that due to the sampling choices described earlier, these shares are not necessarily representative of the distribution of cluster membership in the adult Dutch population as a whole.

Noticeable is that many of its members are single (39%), while the share that is part of a couple, both with children and especially without children, is clearly smaller than in the overall sample. This group has the highest share of university graduates (30%), however, this is not reflected in the income distribution: the share of 'above modal' income groups is smaller than in the sample, while the share of the lowest income group is relatively high (13%). This is only to a very limited extent explained by the share of students (8%), only slightly higher than the sample share (6%).

Furthermore, its average number of commuting days and the share that has a fixed work location does not differ much from the sample as a whole. Many of its members live in the four big cities while a minority lives in smaller cities and villages and especially a large share (33%) lives in middle-sized cities. The latter result corroborates findings in Germany reported by Nobis (2007). That many of the BIKE MM members live in cities is in line with the large share of singles, who tend to live more in larger cities. Only a relatively small share (36%) of the BIKE MM group always has a car available and a large share (38%) never has, while 26% has a car at one's disposal after making arrangements.

Of all groups, the BIKE MM group derives the least pleasure from driving a car, scores the lowest on car convenience and is the most negative about 'car is environmentally friendly'. On the other hand, of all groups, it scores most positive on 'bicycling is pleasant' and 'bicycling is convenient'. In addition, of all groups, the BIKE MM group is most positive about 'traveling by train is pleasant' and 'traveling by train is environmentally friendly' and it is the least negative about 'traveling by bus is pleasant'. In addition, it scores highest on 'PT timeliness', 'PT planning is easy' and 'PT is inexpensive', and scores higher than average on 'PT waiting time acceptability'. Remarkably, the BIKE MM group is more favorable towards PT than is the PT MM group (the fifth distinguished travel group).

In sum, of all groups, the BIKE MM group is most negative about car driving, and most positive about both bicycling and PT. Its attitudes are congruent with its travel behavior.

3.4. The BIKE + CAR group

The third group is labeled BIKE + CAR, and has a sample share of 18%. As in the previous group (BIKE MM), bicycle is used most often: 93% uses bicycle at least once a week, while 60% uses bicycle (almost) every day. However, car use is also relatively high in this group: 83% uses the car at least once a week, and 19% (almost) every day. In contrast, the use of PT is very low: only 4% uses train and only 10% uses BTM at least once a month.

As is the case for the BIKE MM group, also the BIKE + CAR group has a large share of females (58%). In contrast to many other countries like the US, females cycle more than males in the Netherlands. Compared to the BIKE MM group, this group is much older (52) and the propensity to belong to it increases with age. Also the household composition differs considerably from that of the other bicycle group. This group has the smallest share of singles (13%), and the largest share of 'couples with children' (47%), which at least partly explains the relatively high car use in the BIKE + CAR group in addition to the bicycling. That it is this BIKE + CAR group who has the most couples with children and not either of the car groups, suggests that the presence of children is not as compelling a reason for car use as might be expected. Furthermore, whereas the BIKE MM group had the highest share of university graduates, the BIKE + CAR group has the smallest share (5%). This group is especially characterized by a very high share of low-educated (64%). It is noteworthy that both groups with the lowest PT use (BIKE + CAR and CAR MOSTLY, see next group) are dominated by the lowest education level, while the groups with the highest PT use have the largest shares of highly-educated (especially the BIKE MM group). The low education level in the BIKE + CAR group is only partly reflected in the income distribution: especially the middle income group (above minimum up to modal) is overrepresented.

This group has the smallest share of 'paid job' (54%) and a high share of 'other work status' (46%) members, which probably consists in large part of retirees (given the relatively high age), while another large part may consist of females without a job outside the home that mainly take care of children. This is reflected in the low average number of commuting days per week. Only a small share of this group lives in the four big cities (6%) and in middle sized cities (18%), while the vast majority lives in the small cities and villages (76%). A relatively large share always has a car available (70%), while another 26% may have one available after making arrangements.

The BIKE + CAR group has about average car perceptions: they derive slightly less than average pleasure from car driving and perceive driving slightly more than average as convenient. As may be expected, its members derive much pleasure from cycling and they find cycling more convenient than average, but especially the latter opinion is less strong than in the BIKE MM group. For convenience this group seems to lean more towards using the car. This group perceives traveling by train as somewhat more environmentally friendly than average, while pleasure derived from traveling by train and bus is about average. On the other hand, it clearly scores lower than average on the mode-specific attitudes, particularly for 'PT planning is easy'. Possibly the latter finding is related to the low educational level of this group, but possibly also related to living in the small cities/villages (76%), where PT service is lower and less convenient.

To summarize, the BIKE + CAR group has average car attitudes, positive bicycle attitudes and negative PT attitudes. Its attitudes are congruent with its travel behavior. While many members of this group have a car available, they prefer to cycle because it gives them pleasure. Possibly, this is caused by a relatively large group of retirees that bicycle for recreational purposes. On the other hand, for convenience they more lean towards using the car. This may partly be related to homemakers transporting children.

3.5. The CAR MOSTLY group

The fourth group is labeled CAR MOSTLY and has a share of 17%. Of all groups, this group uses car most often: 82% (almost) every day. In contrast, bicycle use in this group is clearly the lowest of all groups: only 45% uses bicycle at least once a week. This group also clearly uses public transport the least, which especially applies to train: only 1% uses train more than once a month and only 5% uses BTM at that frequency.

As is the case for the other car group, more males (58%) than females (42%) belong to this group. The mean age is slightly lower than the sample average (45 vs. 46), but the propensity to belong to it increases with age. It has a relatively small share of singles (17%), while the group 'couple without children' is overrepresented (39% vs. 28% in the sample). On the other hand, the share of 'couples with children' is even slightly lower than the sample share. The low-educated are clearly overrepresented (50%), while the share of university graduates is low (5%). This is not reflected in the income distribution which indicates that the higher incomes are overrepresented. This characteristic is shared with the CAR MM group, which suggests that car use is (still) dependent on income.

This group has by far the highest share of paid jobs (80%) and the highest average number of commuting days. This suggests some possible reasons for their reliance on the car: busier schedule, greater need to be at places at certain times, and therefore (together with income) a higher value of time. The CAR MOSTLY group shares its city size distribution with the BIKE + CAR group: this group tends to live in the smaller cities and villages. Its car availability is highest of all groups: 92% always has a car available while the other 8% has one available after making arrangements.

Of all groups, this group is clearly the most favorable towards car: it scores highest on both car pleasance and convenience, while it is less negative on the car's environmental friendliness. This group clearly scores lowest on both bicycle pleasance and convenience and also on pleasance derived from traveling by both train and bus and on train environmental friendliness. In addition, it scores the least positive on three of the four PT-specific mode variables.

In sum, this group has the strongest attitudes: very positive towards car, negative towards bicycle and very negative towards PT. Its attitudes are congruent with its travel behavior.

3.6. The PT MM group

The fifth group is labeled PT MM and has a share of 14%. Of all groups, this group uses public transport most often and also uses bicycle relatively often: 38% uses BTM and 40% uses train (almost) every day, while 75% uses train and 86% uses BTM at least once a week. Bicycle use is higher than in both car groups, but lower than in both bicycle groups. But this group is not entirely devoted to PT and bicycle: at least 59% uses the car at least once a week.

More males (56%) than females (44%) belong to this group. This group is by far the youngest group (34) and the propensity to belong to it decreases with age. The low average age is reflected in this group having the smallest share of 'couples with children', while on the other hand, the singles and 'other household type' are considerably overrepresented. This group also by far has the largest share of students (23%). This result may be explained by the free PT-card that is given to all students in the Netherlands, which provides them with free PT travel either on the weekend or during the week. Furthermore, this group has the smallest share of 'other work status' individuals, while the share of paid job holders is about equal to the sample share, which also applies to number of commuting days and fixed work location. This group has the smallest share of low-educated, while the middle education category is well represented. It has a relatively large low-income group (13%) and especially many persons with a missing value on income (31%).

As is the case in the BIKE MM group, relatively many of its members live in middle-sized and especially the four big cities. The distribution of car availability also resembles that of the BIKE MM group: only 30% always has a car available, and 38% has one available after making arrangements, a higher share than in any other group. This suggests that PT may be used to cope with a single car in the multi-person household: one household member uses the car while the other uses PT. Such strategies may also be applied, though to a lesser extent, in both bicycle groups (in which 26% has a car available after making arrangements).

The PT MM group has about average car attitudes: the scores on both car pleasance and car convenience are only slightly lower than the sample average. On the other hand, it scores relatively low on bicycling pleasance, while it has an average value on cycling convenience. The train attitudes, on the other hand, are about average, while the score on bus pleasance is slightly higher than average. The PT MM group is clearly the most positive group with respect to 'PT waiting time acceptability'. This suggests that waiting time acceptability is an important determinant for PT use, hence, those who have a high waiting time acceptability use PT more. Furthermore, the PT MM group is also more positive than average about PT planning being easy. On the other hand, its attitudes with respect to 'PT timeliness' and 'PT is inexpensive' are about average.

In sum, concerning attitudes, the PT MM group is the least extreme group as it has about average attitudes towards all modes. This suggests that its attitudes are not completely congruent with its travel behavior. This group has similarities to the 'reluctant riders' identified by Anable (2005). Striking is the relatively low pleasure this group derives from cycling, which is comparable to both car groups. This while in the Netherlands about 40% of the train riders use bicycle as an access mode to the station (Van Exel and Rietveld, 2009). This suggests that a large share of this group uses bicycle not because it gives them pleasure but out of necessity. This is confirmed by the PT MM group's average score on bicycling convenience, while its score is on bicycle pleasure is below average. Also noticeable is that this group is more favorable towards car driving than towards bicycling. Together with the findings that this is a young group with membership propensity decreasing with age and income is not very high (yet), these results suggests that at least some of the PT MM group members will start using car more often once they can afford it.

4. Conclusions

In this paper latent class cluster analysis was applied to identify multimodal travel groups and to explore the effects of socio-demographic, work-related and perception and attitudinal variables on the probability of belonging to each of the five identified classes. As bicycle is a widely used mode of transport in the Netherlands and to a considerable extent used in all identified clusters, all identified clusters are in fact multimodal. We found two groups who mainly use car and two groups that mainly use bicycle, while they differ in the use of public transport. Finally, a public transport group is identified. All included socio-demographic and work-related variables had statistically significant relations with the probability of belonging to the identified segments, as did a considerable portion of the measured attitudinal variables. All multimodal travel clusters and the significant covariates can be interpreted well. Furthermore, with respect to the socio-demographic variables our results agree with earlier research (Buehler and Hamre, 2015): Multimodal travel is associated with: young persons, high education, small households, and car availability, which gives confidence in the results.

Support is found for the suggestion provided in the literature that car drivers who also use public transport have more favorable public transport attitudes than car drivers who do not use public transport. More generally, we find that perceptions and mode-specific attitudes are congruent with mode use, that is, those who use a mode more frequently also have more favorable attitudes towards that mode. However, this does not apply for the PT MM group: this group has only average attitudes towards PT and also about average attitudes towards car. At the same time it has well below average attitudes towards bicycle, which is remarkable as many train travelers (40%) use bicycle as access mode. On the other hand, the (more) exclusive car users not only hold the most favorable attitude towards car, but are also most negative towards other modes, namely bicycle and public transport. To some extent this finding corroborates results from other studies showing that (exclusive) car users generally have biased perceptions towards other modes. Although to be more precise, in the context of this paper it is not so much *biased* perceptions, as *more negative* perceptions and attitudes. An alternative explanation is that the rather negative attitudes towards PT are caused by the fact that for these groups, PT is a relatively bad alternative due to limited access or because PT is not convenient given their activity program.

Based on the results presented in this paper we next discuss the challenges for developing sustainable transport polices with respect to each identified travel group. We roughly rank-order the clusters with respect to the sustainability of their travel behavior. For this, we assume that car is the least sustainable transport mode, followed by PT. As an oversimplification, we further assume that the average number of kilometers driven per car trip does not differ much among the groups. On the one hand, those who are more car dependent may tend to make longer trips by car than others (who might take the train instead), but they may also tend to make shorter trips by car than others (who might bicycle or walk instead). Thus, it is plausible that the average length of car trips does not depend on the extent of multimodality, and therefore that the number of car trips (or, frequency of use) is a reasonable proxy indicator of the total amount of car travel.

The CAR MOSTLY group clearly is the least sustainable group. This group hardly uses any public transport for at least two reasons. First, this group largely lives in smaller cities and villages which have more limited access to PT, especially to train. Second, they have a clear negative attitude towards traveling by train and public transport in general. It is not clear to what extent the negative PT attitude of this group was caused by limited access to good levels of PT service, but given their negative attitudes now, stimulating sustainability by promoting public transport travel may not be very effective for this group, at least not in the short run. And although this group bicycles to some extent, of all groups it is least favorable towards cycling. Hence, this group does not seem very susceptible to bicycle-promoting policies either. As the car orientation is so strong in this group, possibly only a switch to less polluting cars may increase the sustainable travel behavior of this group.

The second least sustainable group is the CAR MM group. This group shares with the other car group a not-so-positive attitude towards bicycling. However, compared to the other car group, this group uses public transport more often and has a more positive attitude towards traveling by PT. As this group is not afraid to use public transport, they may be more susceptible to public transport promotion.

Next is the BIKE + CAR group. This group hardly uses any PT, has rather negative attitudes towards PT and typically lives in smaller places with limited PT access. Hence, promoting PT for this group is probably to no avail. This group mainly uses bicycle: it derives much pleasure from using it, but it less values its convenience. This may explain its relatively frequent car use, to which a large share has always access, at least partly for traveling with children. A further stimulation of bicycle use probably has the most potential for further increasing sustainable travel behavior in this group. It may be worthwhile to find out how bicycle convenience could be increased for this group.

Then there is the PT MM group, which already has a high PT use. Remarkable is the relatively high pleasure this group derives from driving a car, a PT attitude which is only average, and not-so-positive attitudes towards bicycling. Also remarkable is the large share in this group that has to make arrangements with other persons to have access to a car, suggesting that especially in this group household members share a car: one household member uses the car while the other uses PT. The negative relationship with age and the relatively large share of low incomes suggest that part of this group may switch to a more car oriented travel behavior style once they can afford it. The main policy challenge for this group therefore seems to be to prevent the latter from happening as much as possible, perhaps starting by investigating the reasons why their perceptions of transit are not more positive than those of the CAR MM group, and seeing what remedies to this disparity there may be. Examining whether their relatively low bicycle attitudes are caused by low perceived quality of bicycle parking facilities, cycle ways and priorities at crossings may be an interesting avenue for further research.

Finally, the BIKE MM group, who uses bicycle most often and has a strong PT orientation: it has positive bicycle and PT attitudes. In addition, the BIKE MM group has a strong dislike for car traveling and has only very limited access to car. This group clearly has the most sustainable travel behavior and therefore is of the lowest priority to target with sustainable policies.

A limitation of this study is that we did not explicitly measure the availability of all distinguished modes. We did measure and included car availability in the results. With respect to bicycle, we argue that almost all inhabitants of the Netherlands have access to bicycle (our data indicated that 97% of the households possesses at least a single bicycle), and those who have not could easily purchase a bicycle to gain access. With respect to public transport, we believe it is fair to state that almost all residences in the Netherlands have access to at least a bus service accessible either by foot or within a short bicycle ride, though the level of service may differ considerably between different locations. However, we did not measure access to train, though we argued before that city size may serve as a rough proxy: the four big cities and all middle sized cities are well served by train, whereas many of the small towns and villages are not but a fair share is. The consequence of not fully measuring PT access and its level of service is that we are not sure of the meaning of the negative PT attitudes we found for some groups: are these based on objectively poor levels of service experienced in the past or present, or do they represent biased perceptions arising from ignorance of PT's actual level of service and/or a generalized dislike of PT? Hence, further research on multimodality should explicitly measure availability of all modes and also include the level of PT service to gain more insight into this matter.

A related point is our decision to include the attitudinal variables as covariates in our model to examine the extent to which these can explain the probability of belonging to behavioral clusters. This implicitly assumes that attitudes influence behavior, as most theories (and applications) do. However, as the data are cross-sectional, we are not able to empirically disentangle causes and effects. Hence, the question of whether attitudes influence behavior or behavior influences attitudes cannot be answered with our data. It is likely that they mutually influence each other over time. To gain more insight into these causal processes and to determine which causal direction is strongest requires panel data, as used by Kroesen (2014). However, the data used in the latter study does not contain attitudinal variables. Hence, measuring attitudes together with travel behavior in panel data would therefore be a valuable direction for future research.

We conclude this paper by providing two remarks on the indicators of the multimodal travel clusters. First, for this we selected the frequency of use of four different modes, with 8-point ordinal response scales. With these self-reported frequency questions, less information about travel behavior is collected compared to using travel diaries or GPS devices that record all trips and lengths of trips by each mode within a specified period of time, typically ranging from one day to a week and occasionally a couple of weeks. Nevertheless, the multimodal clusters found in the current study are fairly comparable to the clusters found by Kroesen (2014) that were based on one-week diaries, except that we did not identify a light-travel cluster. Although travel diaries and GPS devices collect more detailed information about the travel behavior that in principle allows clustering on distances traveled, the drawback of those methods is the limited period over which data can be collected. These data may miss the occasional PT use of regular car drivers and may therefore not be able to identify this interesting group that may be targeted by PT-stimulating policies. On the other hand, the four simple mode frequency questions applied in the present paper cover a longer period of time (up to a year) that thus allows us to identify the just-mentioned group. A further advantage of using self-reported frequency-of-use questions is that these are simple to understand and do not take much response time, hence, the questions can easily be included in any standard travel behavior survey, which may stimulate more research on multimodality.

The second remark we would like to make is that instead of choosing mode frequency questions as the only indicators for identifying multimodal groups, attitudes may be added as indicators. This may help to further distinguish the multimodal user groups based on the extent to which attitude and behavior agree or not (Diana and Mokhtarian, 2009a, 2009b), for example to distinguish the 'carless crusaders' from the 'reluctant riders' (Anable, 2005) within the PT MM group, which is therefore an interesting direction for further research.

Acknowledgements

The authors would like to thank Dr. Wendy Bohte for constructing the questionnaire and organizing the data collection. The research leading to these results has received funding from the Sustainable Accessibility Randstad program of NWO (Dutch organization of scientific research).

References

Aarts, H., Verplanken, B., Van Knippenberg, A., 1998. Predicting behavior from actions in the past: repeated decision making or a matter of habit? J. Appl. Soc. Psychol. 28 (15), 1355–1374.

Agresti, A., 2002. Categorical Data Analysis, second ed. Wiley, New York.

Anable, J., 2005. 'Complacent car addicts' or 'aspiring environmentalists'? Identifying travel behaviour segments using attitude theory. Transp. Policy 12 (1), 65–78.

Bamberg, S., 2013. Applying the stage model of self-regulated behavioral change in a car use reduction intervention. J. Environ. Psychol. 33, 68–75.

Banister, D., 2008. The sustainable mobility paradigm. Transp. Policy 15 (2), 73–80.
Beckman, J.D., Goulias, K.G., 2008. Immigration, residential location, car ownership, and commuting behavior: a multivariate latent class analysis from California. Transportation 35 (5), 655–671.
Blumenberg, E., Pierce, G., 2014. Multimodal travel and the poor: evidence from the 2009 National Household Travel Survey. Transp. Lett. 6 (1), 36–45.
Buehler, R., Hamre, A., 2015. The multimodal majority? Driving, walking, cycling and public transportation use among American adults. Transportation 42 (6), 1081–1101.
Clifton, K., Muhs, C.D., 2012. Capturing and representing multimodal trips in travel surveys: review of the practice. Transp. Res. Rec., J. Transp. Res. Board Natl. Acad., Washington, DC (2285), 74–83.
Depaire, B., Wets, G., Vanhoof, K., 2008. Traffic accident segmentation by means of latent class clustering. Accid. Anal. Prev. 40 (4), 1257–1266.
Deutsch, K., Goulias, K.G., 2013. Decision makers and socializers, social networks and the role of individuals as participants. Transportation 40 (4), 755–771.
Diana, M., 2010. From mode choice to modal diversion: a new behavioral paradigm and an application to the study of the demand for innovative transport services. Technol. Forecast. Soc. Change. 77 (3), 429–441.
Diana, M., Mokhtarian, P.L., 2009a. Desire to change one's multimodality and its relationship to the use of different transport means. Transp. Res. Part F 12 (2), 107–119.
Diana, M., Mokhtarian, P.L., 2009b. Grouping travelers on the basis of their different car and transit levels of use. Transportation 36 (4), 455–467.
Goulias, K.G., Henson, K.M., 2006. On altruists and egoists in activity participation and travel: who are they and do they live together? Transportation 33, 447–462.
Goulias, K.G., Kilgren, N., Kim, T., 2003. Decade of longitudinal travel behavior observation in the Puget Sound region: sample composition, summary statistics, and a selection of first order findings. Paper Presented at the 10th International Conference on Travel Behavior Research, Moving through Nets: The Physical and Social Dimensions of Travel, Lucerne, 10–14 August 2003.
Heinen, E., Chatterjee, K., 2015. The same mode again? An exploration of mode choice variability in Great Britain using the National Travel Survey. Transp. Res. Part A 78, 266–282.
Jones, P., Clarke, M., 1988. The significance and measurement of variability in travel behaviour. Transportation 15 (1), 65–87.
Kaplan, S., Prato, C.G., 2013. Cyclist–motorist crash patterns in Denmark: a latent class clustering approach. Traffic Injury Prevent. 14, 725–733.
Kim, T.-G., Goulias, K.G., Burbidge, S.K., 2005. Travel behavior comparisons of active living and inactive living lifestyles. Paper Included in the Compendium of Papers of the 85th Annual Meeting of the Transportation Research Board, Washington, DC.
Kitamura, R., Yamamoto, T., Susilo, Y.O., Axhausen, K.W., 2006. How routine is a routine? An analysis of the day-to-day variability in prism vertex location. Transp. Res. Part A 40 (3), 259–279.
Kroesen, M., 2014. Modeling the behavioral determinants of travel behavior: an application of latent transition analysis. Transp. Res. Part A 65, 56–67.
Kuhnimhof, T., Buehler, R., Wirtz, M., Kalinowska, D., 2012. Travel trends among young adults in Germany: increasing multimodality and declining car use for men. J. Transp. Geogr. 24, 443–450.
Lavery, T.A., Páez, A., Kanaroglou, P.S., 2013. Driving out of choices: an investigation of transport modality in a university sample. Transp. Res. Part A 57, 37–46.
Magidson, J., 1981. Qualitative variance, entropy, and correlation ratios for nominal dependent variables. Soc. Sci. Res. 10, 177–194.
Magidson, J., Vermunt, J.K., 2004. Latent Class Models. The Sage Handbook of Quantitative Methodology for the Social Sciences. Sage Publications, pp. 175–198.
McCutcheon, A.L., 1987. Latent Class Analysis (No. 64). Sage Publications.
Nobis, C., 2007. Multimodality: facets and causes of sustainable mobility behavior. Transp. Res. Rec., J. Transp. Res. Board Natl. Acad., Washington, DC (2010), 35–44.
Nylund, K.L., Asparouhov, T., Muthén, B.O., 2007. Deciding on the number of classes in latent class analysis and growth mixture modeling: a Monte Carlo simulation study. Struct. Equ. Model. 14 (4), 535–569.
Ona, J. de, Lópeze, G., Mujalli, R., Calvo, F.J., 2013. Analysis of traffic accidents on rural highways using Latent Class Clustering and Bayesian Networks. Accid. Anal. Prev. 51, 1–10.
Schlich, R., Axhausen, K.W., 2003. Habitual travel behavior. Evidence from a six-week travel diary. Transportation 30, 16–36.
Statistics Netherlands (CBS). <www.statline.cbs.nl>.
Van Exel, N.J.A., Rietveld, P., 2009. Could you also have made this trip by another mode? An investigation of perceived travel possibilities of car and train travelers on the main travel corridors to the city of Amsterdam, The Netherlands. Transp. Res. Part A 43 (4), 374–385.
Van Wee, B., Annema, J.A., Banister, D. (Eds.), 2013. The Transport System and Transport Policy. An Introduction. Edward Elgar, Cheltenham, UK/ Northampton, Massachusetts.
Vermunt, J.K., Magidson, J., 2002. Latent class cluster analysis. In: Applied Latent Class Analysis, pp. 89–106.
Vermunt, J.K., Magidson, J., 2005. Technical Guide for Latent GOLD 4.0: Basic and Advanced. Statistical Innovations Inc., Belmont (Mass.)
Vij, A., Carrel, A., Walker, J.L., 2013. Incorporating the influence of latent modal preferences on travel mode choice behavior. Transp. Res. Part A 54, 164–178.