Travelers' segmentation based on multimodality behaviors and attitudes

Marco Diana*
Dipartimento di Idraulica, Trasporti e Infrastrutture Civili
Politecnico di Torino
Corso Duca degli Abruzzi, 24 10129 Torino – ITALY
Tel. +39 011 564 5638 – Fax. +39 011 564 5699
marco.diana@polito.it

Patricia L. Mokhtarian
Department of Civil and Environmental Engineering, Institute of Transportation Studies
University of California, Davis
One Shields Avenue Davis, CA 95616 – USA
Tel. +1 530 752 7062 – Fax. +1 530 752 7872
plmokhtarian@ucdavis.edu

June 2008

Abstract

Using data collected from 164 French employees of a transportation institute and 1904 residents of the U.S. San Francisco Bay Area, we operationalize a segmentation of mobility patterns based on objective, subjective, and desired amounts of mobility by various modes and overall. We define a multimodality index from basic concepts of information theory, and we especially focus on the degree of multimodality in an individual's current modal mix and desired changes to that mix. The clusters that result showed some similarities and some differences across countries, where the latter are likely due to disparities in the sampling strategies and in the land use/transportation/ cultural milieux. In both cases, however, the clusters have useful policy implications, enabling us, for example, to distinguish car users who might be inclined to reduce car use and increase transit use from those who are largely content with their current modal baskets.

Keywords
cluster analysis, desired mobility, market segmentation, multimodality, public transport

* *Corresponding author*

1 Introduction

Transport systems are more and more confronted with the need to rationalize the use of resources and minimize environmental impacts, while maintaining or even improving current levels of service to support economic development. Encouraging a more consistent use of public transport is among the favorite strategies to achieve this goal, but results seem to lag behind both research and policy efforts.

One of the central problems seems to be the possibility of actually substituting trips using a motorized individual transport mode with trips using a transit service. Several studies have explored this mechanism with different methodologies and in different disciplinary contexts. From the point of view of individual differences, one of the most popular techniques consists of identifying groups of people who have different attitudes and/or behaviors regarding travel, in order to define a set of policy actions (rules enforcement, economic incentives, awareness campaigns etc.) that is targeted for each group. This user segmentation technique is commonplace in the marketing literature and several examples can be found in the transport sector, where segments are defined on the basis of various socioeconomic or personal characteristics (TRB, 1977; Dobson and Tischer, 1978; Tardiff, 1979; Gensch and Torres, 1980; Jensen, 1999; Outwater, Castleberry, Shiftan, Ben-Akiva, Zhou, and Kuppam, 2003; Elgar and Bekhor, 2004; Anable, 2005; Beirão and Cabral, 2007). An example of how to customize marketing in the bus industry is given by Beale and Bonsall (2007).

Beyond individual characteristics, the possibility of diverting trips is also influenced by physical constraints, pertaining for example to the pattern of activities across different locations and the related structure of the transport network. Hence an alternative methodology is to perform an analysis where the actual trips, more than the persons, are grouped on the basis of their suitability for modal transfer. Beyond the pioneering study of Nash and Hille (1968), this kind of analysis, which perhaps lies more in the purview of the transport engineering discipline, is proposed for example by Bovy (1999), Hensher and Reyes (2000) and Massot, Armoogum, Bonnel and Caubel (2006).

The present research offers a contribution that is complementary to the above mentioned segmentation analyses, by focusing on those aspects that are more linked to the individual degree of acquaintance with different transport modes. In this respect, Anable and Gatersleben (2005) have shown the importance of the "predictability" of a transport mode, which in their research was rated only slightly less important than other instrumental factors such as cost, convenience and flexibility, at least for some kinds of trips. Predictability is surely influenced by the performance of the mode itself, but also by the familiarity that the users have with it. The premise of this study is that familiarity with different transport means could be expressed through measures of the real and the perceived relative intensity of use of different modes. We hypothesize that this in turn could be related to the desire to change one's behavior toward a different balance of levels of mobility across the various transport means (which we call the "modal basket"). Our point of view is hence somewhat related to the study of the relationship between mode use habits and mode choices, that is the object of intensive research (see for example Verplanken, Aarts, van Knippenberg and van Knippenberg, 1994; Verplanken, Aarts and van Knippenberg, 1997; Aarts and Dijksterhuis, 2000; Gärling, Fujii and Boe, 2001; Garvill, Marell and Nordlund, 2003; Bamberg, Ajzen and Schmidt, 2003; Thøgersen, 2006).

Several interesting types of individuals can be postulated a priori. For example, among those currently using a variety of modes, some could desire maintaining a similar balance, some could desire to intensify their use of environmentally benign modes (transit, non-motorized) and diminish their use of the automobile, and some could desire the opposite. Conversely, some monomodalists may wish to maintain the primacy of their dominant mode, while others may wish to diversify their modal basket. By focusing on the basket of mobility levels rather than on separate mode-specific ones, we would like to ascertain its usefulness in assessing the desire to alter one's modal basket. This would be of help in giving a preliminary indication of the influence of self-related factors (specifically, stated desires) on the switching potential of field interventions, to be followed by a more thorough evaluation through the consideration of other psychological dimensions such as attitudes, opinions or beliefs. The relevance of the study of multimodal behaviors from a policy viewpoint is well shown by Nobis (2007), who offers also a socioeconomic and demographic characterisation of those segments of the population that are more likely to use different transport modes for their trips.

The objective of our work is then the definition of a new market segmentation approach, complementary to the other above-mentioned transport market studies, which is based on the joint consideration of one's actual and perceived mobility levels by the different transport modes, as well as the desire to change them in order to modify the modal basket. This goal is achieved through a comparative study of two different travel behavior survey datasets, one involving a random sample of individuals living in selected neighborhoods of the San Francisco Bay Area, and the other comprising employees of the French national transportation research institute INRETS. In both cases, the actual use (in terms of the distance traveled or trip frequency) of different transport modes was measured (via self-reports), as well as the subjective assessment of the quantity of travel performed and the desire to increase or reduce the use of any transport means. These field studies are described in section 2, along with the definitions of the variables that will be used for the market segmentation.

Given the potentially high number of modes to be considered in our basket, it is important to try to use synthetic indicators in order to have a manageable number of variables for subsequent analysis. With this objective in mind, in section 3 several kinds of "multimodality indexes" are proposed, on the basis of measures of objective, subjective and relative desired mobility levels, and their usefulness is discussed. These indexes are then used in several cluster analyses that allow for grouping the survey respondents on the basis of their desire to alter their modal basket, in relation to their real and perceived intensity of utilization of different transport means. Repeating the same kind of analysis on somewhat comparable datasets collected in widely different environments allows for greater insights on the matter, helping to understand the general validity of the analysis outcomes. In section 4, we present the most interesting cluster analyses that we performed, including an examination of the socioeconomic characteristics of the U.S. clusters. We draw some conclusions from our exercise in section 5.

2 Experimental settings and variable definition

The present research makes use of two datasets that are the outcome of two studies, independently conducted in radically different contexts. The first study was conducted in 1998 with a postal survey sent to 8000 residents in three different communities in the San Francisco Bay Area, which resulted in 1904 useful observations. The central purpose of the study was to analyze attitudes toward the act of travelling, and the survey obtained data on those attitudes (both general, and specific to various types of travel), as well as personality traits and lifestyle orientations, several measures of mobility, and conventional sociodemographic characteristics. The second study was an Internet survey that was administered to INRETS staff in 2004. INRETS is the French National Institute for Transport and Safety Research, with two larger premises in the Paris and Lyon metropolitan areas and two smaller branches near Lille and Marseille. 164 cases were considered for analysis out of about 550 workers and students that were contacted. The survey goals and contents were somewhat related to those of the U.S. survey, but the questions were focused on a specific trip that was randomly selected among those reported in the introductory part of the survey. This approach should facilitate the quantitative consideration of elements such as attitudes and personality traits within a modelling framework (Diana, 2008), along with the customary trip attributes.

The reader is referred to Mokhtarian and Salomon (2001) for a more detailed presentation of the U.S. survey context, purposes and contents, and to Diana (2005, 2006) for a thorough description of the French survey. In the following we focus instead on those variables that were built from the collected data and that we consider in our subsequent analysis. We select from the two datasets those variables that describe the actual, perceived and desired mobility levels of the respondents by different transport modes. Following the definitions given respectively by Mokhtarian, Salomon and Redmond (2001), Collantes and Mokhtarian (2007) and Choo, Collantes and Mokhtarian (2005), we will name these three entities Objective Mobility (OM), Subjective Mobility (SM) and Relative Desired Mobility (RDM).

2.1 Variables measuring Objective Mobility levels

The two datasets differ both in the way they measure those mobility levels and in the set of transport modes they consider. Concerning OM, people were asked in the U.S. survey for their typical weekly mileage by each of four transport modes, considering only trips up to 100 miles one way (long-distance OM was measured differently and is not considered in this study). The four modes were: (1) driver or passenger in any personal vehicle, (2) bus, (3) rail and (4) walking, jogging, cycling. The French survey asked for the mean frequency of utilisation of 10 ground transport modes (bicycle, motorbike, car driver, car passenger, bus, tram, metro, taxi, suburban train and long-distance train) over the previous 12 months through 5-point ordinal variables (never, sporadically, 1-3 times a month, 1-2 times a week, 3 times a week or more).

For our purposes it is important to try to measure OM in such a way that it can reflect the degree of familiarity of the respondent with the different means. From the point of view of an individual, his/her familiarity with a given mode is probably more related to the amount of time s/he spends in using it than to the distance travelled. We would say that, for example, a person using a bike two hours per *day* is more acquainted with bikes than a car driver using a car two hours a *week* is with his car, although the weekly mileage could be comparable given the different mean speeds of the two modes. Hence we estimate from the U.S. data the number of weekly hours spent in each of the four considered modes through a best guess of their typical speeds. Concerning the French data, we assess instead the number of monthly trips in each of the 10 considered modes from the reported frequencies. Those 14 variables are named with the prefix "OM_" followed by a label indicating the mode. We list them in the second column of block "A" of Tables 1 (for the French dataset) and 2 (for the U.S. dataset).

In the following we focus on the modal basket equilibrium between cars and public transport modes, so that we need to condense information regarding objective mobility levels of different modes. Keeping in mind that we look for an OM measure that is a proxy for the familiarity with the corresponding mode, we consider in the French dataset the sum of OM_CARDR and OM_CARPAX as a measure of the mobility level of the respondent with cars, and we name the new variable OM_CAR. We similarly consider the sum of the mobility levels by bus, trams and metro and we name it OM_PT. For the U.S. dataset we already have a single variable for car use (OM_CAR) and we define OM_PT as the sum of OM_BUS and OM_RAIL, keeping in mind that long-distance trips were not considered. Finally, OM_GLOB is the sum across all modes of the number of weekly hours (for the U.S. case) or of monthly trips (for the French case) spent in each mode. These derived variables are reported in the second column of block "B" of Tables 1 and 2.

2.2 Variables measuring Subjective and Relative Desired Mobility levels

Subjective Mobility and Relative Desired Mobility are measured in both datasets with ordinal variables. Concerning SM, it is important to stress the added value of considering it along with the OM measures in our cluster definition. These entities are not to be seen as two different ways to measure the same thing, since they are complementary in uncovering different aspects of mobility behaviours. Questions related to OM implied a more analytical review of the real behaviours of the respondents, and elicited a rational process to evaluate it in quantitative terms. SM questions have instead to do with feelings, so that mobility is measured on the basis of a personal norm regarding "ideal" mobility levels through a semantic scale.

SM scales range from "I feel I do not travel at all" to "I feel I travel a lot" (by that particular mode), whereas RDM scales go from "I would like to travel much less" to "I would like to travel much more", passing through the neutral point "I would like to travel the same amount as now" (with that particular mode). The U.S. scales have 5 points, whereas the French ones have 10 points when SM is measured and 11 points when RDM is measured.

The considered modes are the same as in the OM analysis, so that as for the OM measures we will have 14 SM variables and 14 RDM variables that are named following the same conventions. The 20 variables from the French dataset are reported in the third and fourth columns of block "A" of Table 1, whereas the 8 variables from the U.S. dataset are listed in the third and fourth columns of block "A" of Table 2.

Beyond those variables that are directly taken from the datasets, for our purposes new SM and RDM aggregate measures for car and public transport use need to be defined, corresponding to OM_CAR and OM_PT in the preceding subsection. However in this case we cannot simply sum SM and RDM values, since they are ordinal measures. Hence we use here a method that has been proposed by Wittkowski, Lee, Nussbaum, Chamian and Krueger (2004). According to this method, it is possible to combine *n* ordinal measures by defining a partial order among the observations. That is, we say that the observation A is "greater" than observation B if A is greater in at least one of the *n* ordinal measures and smaller in none. Then the combined ordinal measure for the observation A is given by the number of observations that are smaller than A minus the number of observations that are greater than A. The interested reader is referred to Wittkowski et al. (2004) for more details and for the study of the statistical properties of the combined measures. In parallel with what has been done for the OM case, we can thus define for the French dataset SM_CAR combining SM_CARDR and SM_CARPAX and SM_PT combining SM_BUS, SM_TRAM and SM_METRO, and similarly for RDM_CAR and RDM_PT. In the U.S. dataset we will define SM_PT and RDM_PT just considering the bus and the rail modes.

Finally, we compute the global SM and RDM mobility levels across all modes. In the case of the French dataset, SM_GLOB and RDM_GLOB are computed assuming that the ordinal measures pertaining to the 10 modes under consideration can be combined in successive steps, according to the hierarchy among these modes that is depicted in figure 1. In doing this we apply an extension of the above methodology that is presented in Diana, Song and Wittkowski (2008). In short, by taking into account the hierarchical structure of the data it is possible to decrease the number of pairwise orderings that are ambiguous (in the sense that the above described partial order cannot be defined), so that the resulting combined measures are much more informative, as shown in that paper. For the U. S. dataset, we combine the measures pertaining to buses and trains to derive the measures for the composite "transit" mode, and then we combine "walk", "car" and "transit" to obtain SM_GLOB and RDM_GLOB (see figure 2).

Figures 1 and 2

The composite measures for SM and RDM, as well as the global SM and RDM mobility levels, are reported in the third and fourth columns of block B of Tables 1 and 2 for the French and the U.S. datasets respectively.

Tables 1 and 2

3 Multimodality indexes

We can see from tables 1 and 2 that the number of considered variables is quite high. As a preliminary step, we propose then a methodology that tries to express the same information with some more synthetic indicators.

In order to define such indicators, we must keep in mind that from our perspective it is not so important to separately analyse mobility levels with respect to single modes, but to consider them jointly. Hence, beyond the global mobility level of individuals, we would like to define an index that measures the degree of multimodality in the individual travel behaviors, i.e. the "product differentiation" inside the modal baskets.

We accomplished this by borrowing some ideas from information theory, a branch of probability theory that allows for measuring the amount of information that is contained in a message, and then adapting those ideas to our framework. According to this discipline, given a probabilistic experiment having n possible results, with the respective probabilities $p_1, \ldots, p_n$ such that $\sum_{i=1}^{n} p_i = 1$, the amount of information H_n supplied by such an experiment is given by the amount of uncertainty that is associated with the experiment itself. Hence it can be measured by the entropy formula (Shannon, 1948):

$$H_n = -\sum_{i=1}^{n} p_i \log p_i = \sum_{i=1}^{n} p_i \log \frac{1}{p_i} \ . \tag{1}$$

The basis of the logarithm can be arbitrary, although it is customary to take it as 2. In this way, the amount of information in an experiment with two equally possible outcomes is equal to 1. Since such a case represents the smallest amount of information (and hence the largest amount of entropy) contained in an experiment with two outcomes, this measurement unit is then named "bit", and the information contained in any other such experiment can be measured in bits. Rényi (1970, p. 579) proposed a useful generalisation of this measure for incomplete probability distributions, i.e. when $\sum_{i=1}^{n} p_i < 1$:

$$H'_n = \frac{\sum_{i=1}^{n} p_i \log \frac{1}{p_i}}{\sum_{i=1}^{n} p_i} \ , \tag{2}$$

which can be seen as the probability-weighted average of the contribution to entropy of each included outcome.

Now, let us come back to our problem. From the outside, we can see the mode choice of an individual for a generic "average trip" as a probabilistic experiment, each mode having a given probability of choice. If we suppose that the characteristics of such a trip are unknown, then the best approximation to quantifying these probabilities is to consider the relative intensities of use of each mode, which in our case can be computed from the OM variables presented in the previous section. The entropy of such an experiment is a measure of the uncertainty of the outcome; the more evenly distributed the choice probabilities of the modes are, the greater the entropy H_n would be, thus indicating that the final choice is less easily predictable because the individual does not predominantly use any mode. Hence, the Shannon entropy of such an experiment can be seen as a measure of the multimodal behavior of the individual, irrespective of his/her mobility level. Moreover, by using the Rényi generalisation it is possible not to consider all the modes that the individual actually uses, but to arbitrarily define the modal basket so that it better fits our research purposes.

Once having defined the set of n modes under investigation and the corresponding frequencies of utilisation $f_1, \ldots, f_n$ (for example, number of trips or number of hours spent onboard in a week or a year, as in our OM variables defined in section 2), we compute a Pure multimodality Index *OM_PI* from the above entropy formulas (1) and (2):

$$OM\_PI = \sum_{i=1}^{n} \left(\frac{f_i}{\sum_{j=1}^{n} f_j} \log_n \left(\frac{\sum_{j=1}^{n} f_j}{f_i} \right) \right) . \tag{3}$$

We take the number of considered modes as the base of the logarithm, so that *OM_PI* lies between 0 and 1. If a given f_i is 0, then the addend of mode i in the above summation is set to zero. When *OM_PI* = 0 the individual uses only one mode among those being considered (or he/she does not use any mode at all), whereas when *OM_PI* = 1 the individual uses all these modes with the same intensity.

It is interesting to note that *OM_PI* is not influenced by the mean mobility level of the individual, since only relative frequencies enter in (3). One might wonder if this is a completely desirable characteristic in our framework, where we would like to use this index as an indicator of the degree of acquaintance with different transport modes. Thus we introduce a variant of *OM_PI* that is also sensitive to the mean mobility level of individuals. For this, we define M as the absolute maximum reported frequency of utilisation of any mode across all observations in the sample, and therefore nM as the potential maximum total frequency across all considered modes. Then, we could define a Mobility-level-sensitive multimodality Index, *OM_MI*, as follows:

$$OM\_MI = \sum_{i=1}^{n}\left(\frac{f_i}{nM}\log_n\left(\frac{nM}{f_i}\right)\right) = \sum_{i=1}^{n}\left(\frac{f_i}{nM}\left(\log_n n + \log_n \frac{M}{f_i}\right)\right) = \sum_{i=1}^{n}\left(\frac{f_i}{nM}\left(1+\log_n\left(\frac{M}{f_i}\right)\right)\right).$$

However, experimentation indicated that taking the natural logarithm seemed to give the best results in terms of an optimal blend between sensitivity to the mean mobility level and sensitivity to the differences of mobility levels across modes, so we define:

$$OM\_MI = \sum_{i=1}^{n}\left(\frac{f_i}{n\,M}\left(1+\ln\left(\frac{M}{f_i}\right)\right)\right) . \quad (4)$$

OM_MI also lies between 0 and 1, with greater values when the travel behavior is more multimodal. As before, if a given f_i is 0, then the addend of mode i in the above summation is set to zero. Moreover, if $M = 0$ then *OM_MI* = 0 (a theoretical circumstance, but with the logical outcome). *OM_MI* also increases with the mean mobility level, and *OM_MI* = 1 when an individual uses all the modes in the basket with frequency equal to M.

To have a more intuitive representation of the differences between the two indexes, we plot in Figures 3 and 4 the functions representing *OM_PI* and *OM_MI*, respectively, when only two modes are considered and M is equal to 15. Note that when $f_1 = f_2 \neq 0$, *OM_PI* always equals 1 (maximum entropy for equal shares, regardless of level), whereas *OM_MI* only increases to 1 as $f_1 = f_2$ approaches M. Note also the comparison between *OM_PI*(0, 1) and *OM_PI*(0, 15) versus

OM_MI(0, 1) and *OM_MI*(0, 15). *OM_PI*(0, 1) and *OM_PI*(0, 15) are both equal to zero (complete monomodality = minimum entropy, regardless of level), whereas the two *OM_MI* measures differ. On the one hand, for *OM_MI*(0, 15) the shares are more disparate (carry more information, i.e. less entropy) than for *OM_MI*(0, 1), but on the other hand, the mobility level of (0, 15) is much greater than that of (0, 1). The outcome of balancing those two factors is that *OM_MI*(0, 15) is greater than *OM_MI*(0, 1).

Figures 3 and 4

Up till now, the frequencies that entered in equations (3) and (4) are those measuring the Objective Mobility levels of individuals. The final step is to generalise the concepts of Pure and Mobility-level-sensitive multimodality indexes, by considering also “perceived” and “relative desired” frequencies, according to the variable definitions given in section 2. These frequencies are measured through the semantic ordinal scales there introduced, but here we treat those ordinal measures as ratio-scaled. Doing so is an approximation, but one that is well-attested and seemingly robust when there are five or more categories to the scale (Bollen and Barb, 1981).

Multimodality indexes based on SM measures, *SM_PI* and *SM_MI,* represent the balance in current usage across the considered modes as it is perceived by the respondent, whereas *RDM_PI* and *RDM_MI* are a measure of the degree of multimodality of the desired changes in the levels of utilisation of different modes. Hence we point out that RDM indexes do not measure the change in the degree of multimodality in the basket itself, should the desired changes take place, since RDM measures are relative to the respondents’ perceptions of current mobility levels and cannot give absolute estimations of multimodality changes if considered alone. To sum up, in the following we consider 6 different multimodality indexes, i.e. Pure and Mobility-level-sensitive multimodality indexes with OM, SM and RDM measures.

4 Cluster analyses

Several cluster analyses have been performed with the variables defined in sections 2 and 3. In the following we present and comment on two of the most interesting types of clusters that were investigated. The type "A" cluster solution considers 9 variables, namely the objective, subjective and relative desired mobility levels in terms of mean values for (1) all ground transportation modes, (2) car, and (3) public transport. Hence it focuses on the general balance between car and transit within the actual modal basket, as well as respondents' perceptions and desired direction and magnitude of change of their baskets. The type "B" cluster solution involves 36 variables for the French case and 18 for the U.S. one, i.e. the mean values for every considered mode (the corresponding variables are reported in block A of Tables 1 and 2), the overall means OM_GLOB, SM_GLOB and RDM_GLOB and finally the pure multimodal (PI) indexes. This analysis allows for a more complete consideration of the whole range of ground transport modes.

Conducting a cluster analysis even on 9 variables, let alone 36, yields cluster centroids that are too complex to be readily cognitively processed. To simplify the interpretation of the results, clusters are built only considering the three variables related to objective mobility for the A-type analysis, and the three overall means (OM, SM, and RDM) and the three PI indexes for the B-type analysis. Then we look at the means across groups for the remaining variables of interest to each cluster. The methodology we employed is a nonhierarchical or k-means procedure where the number of clusters is selected by the researcher. After several trials it emerged that in both the A- and B-type analyses the solutions with four clusters are the most interesting and informative given our research goals.

4.1 A-type cluster analysis

The A-type cluster analysis aims at clarifying the relationships among the actual (OM), subjective (SM) and desired use (RDM) of (1) transport modes in general, (2) cars and (3) transit. It can hence be seen as a bidimensional analysis where different transport mode uses (1-2-3) and different ways to intend the term "use" (OM-SM-RDM) are jointly considered. As such, our work builds on previous research that has already investigated the relationships within these two separate dimensions. Considering the first dimension, pertaining to the relationships among the uses of different transport modes, there is already a vast literature that deals with "intermodal effects" or "modal diversion", where the variation in use of a given mode is related to the variation in use of other modes (see for example Golob, Van Wissen and Meurs, 1986 and Golob and Meurs, 1987 for longitudinal studies; Ben-Akiva and Morikawa, 1990 and Hirobata and Kawakami, 1990 for discrete choice models or Hensher, 2001 for a review on cross-elasticities).

As for the second of the above dimensions, which explores the relationships among objective, subjective and relative desired mobility, previous research results are also available for the same U.S. data set analyzed here. Collantes and Mokhtarian (2007) and Ory, Mokhtarian and Collantes (2007) study the determinants of subjective mobility and its relationship with actual mobility behavior, whereas Choo et al. (2005) analyze models where relative desired mobility is a function of objective and subjective mobility, among others. More disaggregated results are presented by Curry (2000, p. 53), which shows that the correlation between objective mobility and relative desired mobility is generally low for motorized transport modes, and is about 0.3

when walking and bicycling are considered. Using different data, Stradling, Meadows and Beatty (2004) present some interesting crosstabulations of RDM and the expectation of use of a given means in the future.

The goal of our A-type analysis is to offer additional insights into the mixed use of different transport modes that are important from a policy viewpoint. It would be, for example, of interest to understand if people who intensively use cars desire to reduce their use proportionately more or less than people who usually travel by public transport, or to determine which transport modes are more likely to be less used by those who are fed up with travelling. For these kinds of analyses it is useful to jointly consider the above defined dimensions.

4.1.1 Cluster centroids and mobility levels

Tables 3 and 4 present the resulting A-type cluster centroids for the French and the U.S. datasets respectively. We can label the resulting clusters based on the three objective mobility variables around which the clusters were formed, whereas the other values in italics are group means that were computed a posteriori. Turning first to the French dataset, we have a large first group of strong car users, a second group of strong public transport users, a small third group of weak users of both modes and a fourth group of strong users of both modes. We recall that for the French dataset the reported OM measures are the sum of the number of monthly trips taken by all the considered modes. Hence it is not meaningful to make "vertical" comparisons between, say, OM_CAR and OM_PT, since the former is the sum of the number of trips taken by two modes and the latter the sum of three. For the same reason, French and U.S. OM measures are also not directly comparable, given the different number of considered modes.

OM measures in table 3 seem to be unreasonably low, but those averages are biased downward by the fact that the underlying variables are composites of ordinal OM measures whose highest grade is “more than 3 times a week” (see section 2) so that higher mobility levels are flattened around that value. However we believe that this is not a big concern, since for our research purposes OM variables do not measure the exact mobility levels of individuals, but rather represent the degree of acquaintance with a given mode, as stated in the introduction. In that sense, we believe that above a certain threshold utilization level for a given mode, the corresponding familiarity with that mode does not increase proportionally, so that underestimations of higher mobility levels are not so relevant in our case.

Tables 3 and 4

In tables 3 and 4, subjective mobility and relative desired mobility measures are the cluster means of the individual SM and RDM scores that were computed according to the procedure that was presented in subsection 2.2. The individual scores can theoretically range from $-(n-1)$ to $(n-1)$ when the partial order among the n observations is complete, lower scores indicating lower SM and RDM levels of that individual in comparison with the whole sample. If the order is not complete, then the range of the scores is narrowed in proportion to the number of ambiguous pairwise orderings. However, it is important to understand that, for example, an individual RDM score of zero does not mean that the individual does not wish to alter his/her mean mobility level across the considered modes. It simply means that an equal number of cases are greater than this one (i.e. have RDM measures at least as high on all categories, and strictly higher on at least one) as are smaller than this one. Moreover, the SM and RDM means that are shown in the tables are to be interpreted in relative terms across the different clusters: specifically, comparisons can only meaningfully be done among numbers in the same row, i.e. horizontally reading the data.

With that in mind, it is interesting to compare the following results with those of Diana and Mokhtarian (2007), that report the same clusters but with a different method to combine SM and RDM measures, which hence leads to a different cluster interpretation. Their method was based on heuristic approximations that keep the information on a case's position relative to the neutral point in a bipolar scale such as the RDM one. In other words, in that paper it was possible to understand if the respondents within a given cluster tended to actually like to travel more or less than what they actually do. The cluster interpretations that can be found in that paper hence give information that is complementary with the one that is reported in the following, since here the focus of the analysis is on the above mentioned horizontal comparisons.

If we "horizontally" read the subjective mobility measures, we can see that they are quite correlated with the corresponding OM measures, confirming the results of Collantes and Mokhtarian (2007). On the other hand, RDM measures show interesting patterns among clusters in the French dataset. Those who intensively use a given mode (i.e. car for groups 1 and 4 and public transport for groups 2 and 4) would like to travel less than the general average across the four groups by such a mode and (except for group 4) would like to travel more by the other mode. Group 3 seldom uses both modes and is thus willing to increase their use more than the average. The four groups appear to be less distinct in terms of global RDM levels, so that the desire to travel more (or less) overall seems less strongly related to the composition of the modal basket.

Repeating the same kind of analysis on the U.S. dataset gives us partly different results. This is not surprising, given the nature of the cluster analysis technique and the radically different experimental settings. Three out of four groups predominantly use cars, comprising 89% of the sample, compared to the 54% of respondents belonging to group 1 in the French dataset. People who predominantly use public transit fall from 29% in the French sample (Groups 2 and 3) to

11% in the U.S. sample (Group 3). The latter group is much less monomodal toward transit than the French group 2 and has OM patterns rather similar to the French group 3. However a cautionary note must be considered when comparing OM measures across datasets, since these measures are different, as explained in subsection 2.1. Another interesting disparity between the two samples is that the U.S. analysis did not distill a group that intensively uses both means.

Subjective mobility measures have slightly lower correlations with OM measures compared to the French sample. This could be due to the coarser SM scale used by the U.S. sample, but behavioral differences probably also have an influence. In particular, the U.S. group 1 travels slightly more than group 3 but it “feels” it travels less. However we do not observe this discrepancy when looking at the French groups 1 and 3. This could be due to different attitudes regarding cars and perhaps even more so, public transport in the two samples, since the same differences between the datasets can be detected when comparing OM_GLOB and RDM_GLOB for groups 1 and 3.

As explained above, RDM scores do not allow us to determine if people belonging to a group desire to travel more or less on average. Hence we had a closer look at the RDM measures of the two considered transit modes in the U.S. dataset, namely the bus and the train. Only 4% wanted to use both transit modes less. In fact, however, the majority of the sample (55%) wanted to travel either by bus or by train the same amount as now, and 80% of those were currently traveling little or not at all by transit. Only 27% of the sample actually wanted more travel by train, and 11% wanted more by bus. Overall though, latent demand for travel seems greater in the U.S. dataset, a result that is probably linked to the different socioeconomic composition of the sample. This could also be due to the lower levels of use of transit, so that the desired modal baskets of the two samples are probably closer to each other than is the actual use of different means.

4.1.2 Mobility patterns among different clusters

The similarities and differences between the two datasets that we analysed are useful for drawing some more general conclusions from our exercise. A preliminary consideration is related to the existence of individual thresholds concerning the traveling activity, that can of course be interpreted in view of the long-standing debate on the existence of travel time budgets (Zahavi and Talvitie, 1980; Zahavi and Ryan, 1980; Supernak, 1982). Several studies have been published on this matter supporting different points of view, but more recent reviews that draw on a vast number of works (Mokhtarian and Chen, 2004; Joly, 2005, 2006) tend to conclude that unobserved travel time “desired budgets” exist (in contrast to observed travel time expenditures), but vary across individuals, being related to several factors such as socioeconomic status, kind and patterns of activities at different locations and characteristics of residential areas.

Following this line of research, our results can offer preliminary insights on another possible source of variation of individual budgets, namely the actual use of different modes. In fact, the values in tables 3 and 4 and the lack of a clear socioeconomic differentiation between clusters, as we will see in the following subsection 4.1.3, suggest to us that people have different ideal levels of use of the different modes. We can see this by separately considering the clusters that show strong monomodal behavior (groups 1 and 2 of the two datasets) and then some clusters with comparable global mobility levels (namely, groups 1 and 3).

As a general rule, strong users of a given mode would like to bring more balance to their basket by decreasing the use of this mode more than the average, and increasing the use of the alternative mode. Hence it is important to jointly look at the use of the different modes to understand people’s desires. For example, French groups 2 and 4 have the same level of use of public transport, and U.S. groups 3 and 4 are the same for cars, but the corresponding RDM levels are quite different.

However, turning our attention to objective and relative desired global mobility levels, other interesting patterns emerge. For example, the U.S. group 3 desires to travel dramatically less than the U.S. group 1, even if the two OM_GLOB values are comparable. The decisive difference is that the group 3 uses more transit than cars, so that the ideal level of use of the two modes is clearly different. We do not observe the same phenomenon when considering the French groups 1 and 3, so that ideal levels are clearly different also for the two samples.

Our methodology does not allow us to quantify the above two gaps in the ideal levels of use (between different modes and between different datasets); nevertheless we can infer that (1) the ideal levels of use of the different transport means may be linked more with the balance in the modal basket than with one's global mobility levels and that (2) when the modal basket is equilibrated, a comparison of the two samples shows differences in the relative attractiveness of the modes, that are probably due to the different experimental contexts. For example, the French sample, comprising employees of a transportation research institute, would be more sensitive to the negative externalities imposed by the automobile, and perhaps ideologically more inclined toward environmentally-benign travel modes, than would the more general sample of the U.S. dataset (although the northern Californians comprising the U.S. sample would in turn be expected to be more environmentally aware and proactive than the country as a whole).

These findings show the importance of considering multimodal behaviors, beyond the absolute mobility levels, in order to understand the direction and magnitude of potential modal shifts. This fact prompted us to more carefully look into the modal basket mechanism, through the consideration of the multimodality indexes in the B-type analyses presented in section 4.2.

4.1.3 Socioeconomic characteristics of the U.S. clusters

A closer look at the socioeconomic characteristics of the different clusters allows us to gain further insight concerning the above mentioned classification. This kind of analysis is not so meaningful on the French dataset, given the very particular characteristics of the sample that does not allow for an easy generalization of the results and the small number of observations for three out of the four clusters. In contrast, characterizing the clusters of the U.S. dataset is an effective way to understand how personal traits relate to multimodality attitudes and behaviors. Concerning the U.S. A-type clusters, it is particularly interesting to study the differences between the first two clusters on one hand (heavily versus somewhat car-oriented) and between car-oriented and transit-oriented clusters on the other. Light travelers would presumably have greater within-group variability in terms of socioeconomic characteristics than the other clusters but offer less insight concerning multimodality behaviors.

Since the socioeconomic variables we consider are ordinal rather than continuous, and (more importantly) are not always approximately normally-distributed, a standard analysis of variance (ANOVA) test of whether the means on each variable are equal across clusters is not necessarily appropriate. Accordingly, we perform the conceptually similar non-parametric Kruskal-Wallis test to ascertain whether the group differences that we found are simply due to random variation. The corresponding p-values are always equal to or below .005, so that we can safely reject the null hypothesis of no difference among the group populations.

We consider at the outset the number of available cars in the household and the percentage of time in which a vehicle is available to the respondent, that are respectively shown in figures 5 and 6. Both values are not surprisingly lower for transit-oriented than for car-oriented people. However it is interesting to note that there is almost no difference between the two car-oriented clusters in these two aspects, so that their distinction is not due to the “car availability” instrumental factor. A detectable difference is instead related to sex, since 65% of the heavily car-oriented people are male, whereas for group 2 this figure is lowered to 53%, roughly the same as for the transit-oriented cluster. Probably the greater symbolic value of cars for male drivers (Steg, 2005) can explain part of this difference.

Figures 5 and 6

The share of persons who did not attend graduate school is roughly equal to 35% across the four clusters, but transit-oriented persons are much more likely to have completed a 4-year degree (figure 7). On the other hand, these latter have a lower income than the average (figure 8); the well-known positive relationship between income and educational level is thus not reflected in our multimodality-based clusters. Strong car users are more likely to be full-time workers, but other socioeconomic characteristics do not display much difference between the two car-oriented groups, as shown in figures 5 to 8. To sum up, the use of a multimodality-based clustering technique cannot easily be represented by segmentations based on socioeconomic characteristics, since patterns of modal usage are generated from both instrumental and affective factors.

Figures 7 and 8

4.2 B-type cluster analysis

The B-type analysis should clarify the added value of considering the proposed multimodality indexes within our framework. Our expectation in performing this analysis is to have a closer look at the relationship between the fact of being more or less acquainted with a variety of transport modes and the desire to alter one's modal basket, while controlling for the actual, perceived and desired change in the level of use of several different modes. The interest in separately looking at the mean mobility level and the multimodal behavior pushed us to consider "*_GLOB" variables and PI indexes in the analysis. Given the definitions of our indexes, an analytically similar procedure would be to define clusters only on the basis of MI variables. A variant of a B-type analysis that uses MI instead of "*_GLOB" variables and PI has been performed, and the cluster solutions are almost equivalent.

4.2.1 Cluster centroids and mobility levels

Tables 5 and 6 present the resulting B-type cluster centroids for the French and the U.S. datasets respectively. Clusters are presented by ascending OM_GLOB values. The means for the six variables on which the cluster formation is based are the ones that are not italicized in the tables. Their values are standardized since the range of the raw PI variables is up to three orders of magnitude smaller than that of the "*_GLOB" variables. The numbers in the tables represent thus the number of standard deviations the original value is from the mean. This must be taken into account in the interpretation of the results, since for example the 0.1 values for RDM_GLOB in the French groups 2 and 3 do not mean that the respondents in these groups desire to keep their amount of travel nearly constant on average, but that on average they desire to change this amount only slightly more than the mean of the whole sample.

The tables report also in italics the group means of the other variables defined in section 2, that have been directly observed. Their analysis can help to define the characteristics of the distilled groups. SM scales have a rather intuitive meaning, so that the group means of the questionnaire scales are directly reported. RDM scales are instead somewhat more cumbersome, since RDM mean values of the cluster centroids are often quite close to the neutral point (that is "3" in the U.S. dataset and "6" in the French one). This would hamper an easy appreciation of the differences among clusters that actually exist. Hence we decided to define a new 7-point scale using the symbols "– – –", "– –", "–" (representing the desire to decrease the use of a given mode), "=" (neutral point), "+", "+ +" and "+ + +" (representing the desire to increase the use of a given mode). Then we relate this scale to the mean values computed from the scales that we have in our datasets through the definition of thresholds. These thresholds, reported in table 7, have been judgementally set in such a way that the differences among clusters are highlighted. Moreover, the relative width of each corresponding point in the two 7-point scales is the same. For example, the width of extreme point "– – –" is 3 in the French dataset where the original scale has 11 points, of which 5 are negative. In the U.S. scale we have 5 points of which 2 are negative, hence the width of that extreme point will be 3*2/5 = 1.2.

Tables 5, 6 and 7

It can be seen that the analysis gave different clusters in the two samples. For France, the first two clusters have both their mobility level and their OM_PI below average, but differ in that the first cluster contains largely monomodal car users, whereas the second cluster is more multimodal albeit still auto-dominated. The third and fourth clusters are more multimodal, differing in that the third group has a mixture of mode-specific desires with respect to altering their travel, whereas the fourth group is surfeited with nearly every mode. Overall, the first three groups would like both to increase their amount of travel more than the average and to alter their modal basket composition by diversifying their desired changes in various modes, whereas group 4 wants the opposite. The U.S. dataset produced four clusters where car use always prevails and respondents would even increase it. Group 1 has a low mobility level and uses only cars, but is moderately open to increase the use of rail. Groups 2 and 4 have higher mobility and multimodality levels, but the former would like to sharply decrease its transit usage and thus the quantity of trips, whereas the latter is mildly deprived with respect to all modes except bus, even though it is the maximum user of all four considered modes (car, train, bus and non-motorized). Group 3 is somewhat the opposite, since it travels less than average, feels it travels more than that but would like to travel even more both by car and by train.

4.2.2 Mobility patterns among different clusters

B-type clusters can be further interpreted by looking at some other properties of the groups, namely the values of the italicised variables in the tables. Concerning the French dataset, we preliminarily notice that these groups are (obviously) different from that of the A-type analysis, since we have now two groups (1 and 2) that almost exclusively use cars and two others with a more or less marked predominant use of public transport, as we mentioned above. Group 2 shows also an occasional use of public transport, so that its members feel themselves to be less monomodal and have a smaller desire to diversify their modal basket compared to group 1. Groups 3 and 4 travel more than the average and would like to reduce trips taken by both private and public transport modes. The distinction is that group 4's desire to reduce travel is nearly across the board (with the low diversity in mode-specific RDM scores shown in the low value of RDM_PI), whereas group 3's desire is more selective (and hence, the greater diversity showing as a higher RDM_PI).

Another interesting and policy-relevant observation is the strong attractiveness of cycling for all groups, and the desire to travel by car more as a passenger and less as a driver. Thus it seems surely worthwhile for policy makers to take those actions that can transform such desires into intentions and, hopefully, actual choices. The measures concerning desired use of non-motorised modes could be affected by a social desirability bias, beyond the often already-large gap between desires (which are influenced by factors such as health concerns) and intentions (which can take into account more practical aspects such as scheduling problems). On the other hand, our finding concerning the differential attractiveness of the car-driver and car-passenger modes is less trivial and hopefully less affected by such problems, and may be able to give an indication of the effectiveness of initiatives such as car-sharing and car-pooling. In more general terms, the strong attractiveness of a less intensively used mode such as the bike helps explain how the first three groups would like to change their modal basket in a more polarised way. The last group, as indicated above, contains those who are fed up with travelling and would like to drastically reduce their trips by every motorized transport means, increasing only cycling Thus, their desired changes favour cycling exclusively, though as mentioned earlier, this may not mean that the desired modal basket contains only cycling – it could simply involve greater balance among all modes.

Since the various modes are considered in a more aggregate way we cannot match in the U.S. dataset some of the findings of the analysis on the French data. It is however evident that the attractiveness of nonmotorized transport modes is confirmed for all four U.S. groups, whereas the analysis did not discern a group that is willing to decrease or keep constant the use of cars. Groups 1 and 2 are similar to the corresponding French ones in that (1) they both have an intensive use of cars and (2) the second group has more familiarity with public transport than the first one. However here we do not see a replication of the above finding of a desire to increase the use of public transport for the U.S. group 2. This can be explained by looking at the value of RDM_GLOB, which is now negative. This strong desire to decrease the global mobility level comes at the expense of the level of utilisation of transit.

On the other hand, two of the four clusters, comprising 58% of the sample, show an interest in increasing their rail usage slightly (group 4, 17%) or moderately (group 3, 41%), while not decreasing their (modest) bus usage. Whether this represents a desire to be in a different location, where rail is a practical alternative, or a desire specifically for the perceived higher levels of speed and comfort of rail compared to bus, or simply a bias toward rail as a generally more attractive mode, it also points to the infeasibility of saturating a metropolitan area in the U.S. (even one as relatively transit-oriented as San Francisco) to the extent of satisfying the latent demand for that means of travel. And, as already pointed out, this desired increase in rail travel would apparently *not* be realized as a substitute for car travel, since respondents in those two groups tend to want to increase their (already substantial) levels of car travel as well.

We finally point out that the B-type clusters here presented are qualitatively quite similar to those described in Diana and Mokhtarian (2007), although the computation of SM and RDM global mobility levels is different. This makes us more confident in the stability of our findings, given the well-known high sensitivity of the cluster analysis technique to the definition of the clustering variables.

4.2.3 Socioeconomic characteristics of the U.S. clusters

As previously done for the A-type clusters (see section 4.1.3), to see how the clusters might differ on socioeconomic traits, we restrict our analysis to the U.S. dataset. Concerning sex, 60% of group 4 is male, whereas for groups 2 and 3 this percentage is 44% and 42% respectively. Characterizing these clusters with the same socioeconomic variables that were analyzed in section 4.1.3, namely the number of vehicles in the household, the percentage of time a vehicle is available to the respondent, educational level and income (figures 9 to 12 respectively), other interesting patterns emerge. In this case, we see that groups 3 and 4 have roughly the same car availability but rather different multimodal behaviors and desires, as pointed out in the preceding section. Not surprisingly, group 1 has a higher than average car availability and the opposite is true for group 2. Group 1 is less educated and group 4 is more educated than average, whereas groups 2 and 3 are poorer and group 4 is richer.

Figures 9, 10, 11 and 12

Looking at these figures, we can conclude that the generally more affluent respondents of group 4 are quite happy with their higher than average mobility levels and modal basket composition, even if they are open to increasing their already relatively high use of train. Interestingly, group 2 is in the opposite situation, concerning both socioeconomic conditions and mobility desire. It is discouraging from a policy point of view to see that what they want, to improve their satisfaction with travel, is to sharply cut their transit use and thereby become much more monomodal with respect to the car. This contrasts with group 4, which displays the same levels of use of transit but is happy with that. This may be a case of the less-affluent desiring a lifestyle it perceives to characterize the upper class, when the upper class in reality is living, by choice, the lifestyle that the former group aspires to shed. On the other hand, given the correlations that exist among income, residential location, and transit supply, these results may well reflect differences in the latter characteristics between the clusters. For example, the transit use of cluster 4 may more likely take place on comfortable commuter trains and express buses, while that of cluster 2 may occur on crowded trams, metro rail, and local buses.

5 Conclusions

This paper examined various mode-specific and overall measures of mobility – objective, subjective, and desired – for a sample of transportation institute employees in France, and a sample of the general population of the San Francisco Bay Area in the U.S. We focused on analyzing the modal mixes, or "baskets", used, perceived, and/or desired by individuals, with a view to identifying population segments having similar modal baskets, and better understanding the role of modal mix in personal travel patterns and desires.

We conducted two cluster analyses, each resulting in four distinct clusters for each sample. Although the relative cluster sizes are not necessarily representative of the population at large, and certainly size-based comparisons across the two samples are unjustified, the types of clusters identified are of general interest, and have useful policy implications. As a general comment, the results suggest that different people have different ideal targets for the use of each mode; it is valuable to explore further the types of people wanting to decrease their car use, increase their transit use, and/or increase the multimodality of their modal baskets. The second cluster analysis of the French sample, for instance, produced two clusters of predominantly car users. Group 1 is largely monomodal and would like to increase the use of transit. Group 2 uses public transport to a certain degree, wants to use it more and to decrease the use of cars more sharply than the other group. Then it would presumably be more effective to market transit to the second group of car users than to the first.

We also identified people wanting to decrease car driving but increase their use of car as a passenger; they are potential targets of ridesharing or carsharing campaigns. Finally, a desire to increase the use of non-motorized modes such as cycling and walking appeared across *all* of the clusters from both countries. There may be a large measure of "wishful thinking" in such responses (knowing it is “good for you” but not willing to commit the time and effort to change), and even if change on that dimension is realized, it may take more the form of recreation rather than replacing very much auto travel. Nevertheless, the presence of such a desire is at least an important starting point for behavior change that would have positive health effects even if limited transportation effects.

The preliminary analyses reported here could be extended in several different ways. With the same datasets, it would be valuable to explore the identified clusters further, using other variables in each dataset to flesh out the skeletal picture we currently have of the types of people in each cluster. The usefulness of the multimodality indexes that we defined could further be assessed by using them also in different kinds of analyses, for example generalizing the paper of Choo et al. (2005) (which modeled the relative desired mobility for personal vehicles, alone among the modes surveyed for short-distance travel) to explore more systematically how relative desired mobility for various modes differs with degree of multimodality (among other explanatory variables). For future data collection efforts, it would be desirable to customize the measurement of mobility-related variables to be optimally-suited for the current purpose. It would also be highly desirable to conduct identical (as much as possible) surveys of the general population in several different countries, to enable a more systematic exploration of cultural differences relevant to travel behavior. In any case, the concept of multimodality appears to make a useful contribution to the study of travel patterns, and one which could be profitably explored further.

References

Aarts, H. and Dijksterhuis, A. (2000) The automatic activation of goal-directed behaviour: The case of travel habit, *Journal of Environmental Psychology*, 20(1) pp. 75-82.

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

Anable, J. and Gatersleben, B. (2005) All work and no play? The role of instrumental and affective factors in work and leisure journeys by different travel modes, *Transportation Research Part A*, 39(2-3), pp. 163-181.

Bamberg, S., Ajzen, I. and Schmidt, P. (2003) Choice of travel mode in the theory of planned behavior: the roles of past behavior, habit, and reasoned action. *Basic and Applied Social Psychology*, 25(3), pp. 175–187.

Beale, J.R. and Bonsall, P.W. (2007) Marketing in the bus industry: A psychological interpretation of some attitudinal and behavioural outcomes, *Transportation Research Part F*, 10(4), pp. 271-287.

Beirão, G. and Cabral, J. S. (2007) A market segmentation analysis using attitudes towards travel behaviour. *11th World Conference on Transport Research*, Berkeley, USA, (presentation).

Ben-Akiva, M. and Morikawa, T. (1990) Estimation of switching models from revealed preferences and stated intentions, *Transportation Research Part A*, 24A(6), pp. 485-495.

Bollen, K.A. and Barb, K.H. (1981) Pearson's R and coarsely categorized measures, *American Sociological Review* 46(2), pp. 232-239.

Bovy, P.H. (1999) Urban structure and modal distribution, *Public Transport International*, 48(1), pp. 8-15.

Choo, S., Collantes, G.O. and Mokhtarian, P.L. (2005) Wanting to travel, more or less: exploring the determinants of the deficit and surfeit of personal travel, *Transportation*, 32(2), pp. 135-164.

Collantes, G.O. and Mokhtarian, P.L. (2007) Subjective assessments of personal mobility: What makes the difference between a little and a lot?, *Transport Policy*, 14(3), pp. 181-192.

Curry, R.W. (2000) *Attitudes toward travel: The relationships among perceived mobility, travel liking, and relative desired mobility*, Master's Thesis, Institute of Transportation Studies, University of California, Davis, 252pp.

Diana, M. (2005) *An exploratory web-based attitudinal travel survey administered to INRETS staff*. Les Collections de l'INRETS – "Outils et méthodes" n°12, Arcueil, 187pp, ISBN 2-85782-623-0.

Diana, M. (2006) Utilité primaire des déplacements et multimodalité : conception et réalisation d'un outil d'enquête novateur, *Recherche Transports Sécurité*, 93, pp. 279-294.

Diana, M. (2008) Making the "primary utility of travel" concept operational: a measurement model for the assessment of the intrinsic utility of reported trips, *Transportation Research Part A*, 42(3), pp. 455-474.

Diana, M. and Mokhtarian, P.L. (2007) The individual desire to change a "modal basket" and its relationship with the actual and perceived use of different transport means, *11th World Congress on Transport Research*, Berkeley, USA, 41 pp.

Diana, M., Song, T. and Wittkowski, K.M. (2008) Studying travel-related individual assessments and desires by combining hierarchically structured ordinal variables, *Research paper available from the Authors*.

Dobson, R. and Tischer, M.L. (1978) Perceptual market segmentation technique for transportation analysis, *Transportation Research Record 673*, pp. 145-152.

Elgar, A. and Bekhor, S. (2004) Car-rider segmentation according to riding status and investment in car mobility, *Transportation Research Record 1894*, pp. 109-118.

Gärling, T., Fujii, S. and Boe, O. (2001) Empirical tests of a model of determinants of script-based driving choice, *Transportation Research Part F*, 4(2), pp. 89-102.

Garvill, J., Marell, A. and Nordlund, A. (2003) Effects of increased awareness on choice of travel mode, Transportation, 30(1), pp. 63-79.

Gensch, D.H. and Torres, P.T. (1980) Perceived-difference segmentation model for mass transit marketing, *Transportation Research Record 765*, pp. 16-22.

Golob, T.F. and Meurs, H. (1987) A structural model of temporal change in multi-modal travel demand, *Transportation Research Part A*, 21A(6), pp. 391-400.

Golob, T.F., Van Wissen, L. and Meurs, H., (1986) A dynamic analysis of travel demand, *Transportation Research Part A*, 20A(5), pp. 401-414.

Hensher, D.A. (2001) Modal diversion, in K.J. Button and D.A. Hensher (eds.), *Handbook of Transport Systems and Traffic Control*, Pergamon, Amsterdam, pp. 107-123.

Hensher, D.A. and Reyes, A.J. (2000) Trip chaining as a barrier to the propensity to use public transport, *Transportation*, 27(4), pp. 341-361.

Hirobata, Y. and Kawakami, S. (1990) Modeling disaggregate behavioural modal switching models based on intention data, *Transportation Research Part B*, 24B(1), pp.15-25.

Jensen, M. (1999) Passion and heart in transport – a sociological analysis on transport behavior, *Transport Policy*, 6(1), pp. 19-33.

Joly, I. (2005) *L'allocation du temps au transport – De l'observation internationale des budget-temps de transport aux modèles de durées*, Ph.D. Thesis, Université Lumière Lyon 2, Lyon, VI+477pp.

Joly, I. (2006) Stability or regularity of the daily travel time in Lyon ? Application of a duration model, *International Journal of Transport Economics*, 33(3), pp. 369-400.

Massot, M.-H., Armoogum, J., Bonnel, P. and Caubel, D. (2006) Potential for car use reduction through a simulation approach: Paris and Lyon case studies, *Transport Reviews*, 26(1), pp. 25-42.

Mokhtarian, P.L. and Chen, C. (2004) TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets, *Transportation Research Part A*, 38(9-10), pp. 643-675.

Mokhtarian, P.L. and Salomon, I. (2001) How derived is the demand for travel? Some conceptual and measurement considerations, *Transportation Research Part A*, 35(8), pp. 695-719.

Mokhtarian, P. L., Salomon, I. and Redmond, L. S. (2001) Understanding the demand for travel: It's not purely 'derived', *Innovation: The European Journal of Social Science Research*, 14(4), pp. 355-380.

Nash, A.N. and Hille, S.J.K. (1968) Public attitudes toward transport modes: A summary of two pilot studies, *Highway Research Record 233*, pp. 33-46.

Nobis, C. (2007) Multimodality: facets and causes of sustainable mobility behavior, *Transportation Research Record 2010*, pp. 35-44.

Ory, D.T., Mokhtarian, P.L., and Collantes, G.O. (2007) Exploring the cognitive and affective mechanisms behind subjective assessments of travel amounts, *Environment and Behavior*, 39(4), pp. 494-528.

Outwater, M.L., Castleberry, S., Shiftan, Y., Ben-Akiva, M., Zhou, Y.S. and Kuppam, A. (2003) Attitudinal market segmentation approach to mode choice and ridership forecasting – Structural Equation Modeling, *Transportation Research Record 1854*, pp. 32-42.

Rényi, A. (1970) *Probability theory*, North-Holland, Amsterdam, 670pp.

Shannon, C.E. (1948) A mathematical theory of communication, *The Bell System Technical Journal*, 27, pp. 379-423; 623-656.

Steg. L. (2005) Car use: lust and must. Instrumental, symbolic and affective motives for car use, *Transportation Research Part A*, 39(2-3), pp. 147-162.

Stradling, S., Meadows, M. and Beatty, S. (2004) Who will reduce their car use – and who will not?, in T. Rothengatter and R. D. Huguenin (Eds.), *Traffic and Transport Psychology – Theory and Application – Proceedings of the ICTTP 2000*, Elsevier, Amsterdam, Netherlands, pp. 459-467.

Supernak, J. (1982) Travel-time budget: a critique, *Transportation Research Record 879*, pp. 15-28.

Tardiff, T.J. (1979) Attitudinal market segmentation for transit design, marketing, and policy analysis, *Transportation Research Record 735*, pp. 1-7.

Thøgersen, J. (2006) Understanding repetitive travel mode choices in a stable context: A panel study approach, *Transportation Research Part A*, 40(8), pp. 621-638.

TRB – Transportation Research Board (1977) Preferences, perceptions, and market segments in travel behavior, *Transportation Research Record 649*, 48pp.

Verplanken, B., Aarts, H., van Knippenberg, A. and van Knippenberg, C. (1994) Attitude versus general habit: Antecedents of travel mode choice. *Journal of Applied Social Psychology*, 24(4), pp. 285–300.

Verplanken, B., Aarts, H. and van Knippenberg, A. (1997) Habit, information acquisition, and the process of making travel mode choices, *European Journal of Social Psychology*, 27(5), pp. 539-560.

Wittkowski, K.M., Lee, E., Nussbaum, M., Chamian, F.N. and Krueger, J.G. (2004) Combining several ordinal measures in clinical studies, *Statistics in Medicine*, 23(10), pp. 1579-1592.

Zahavi, Y. and Ryan, J.M. (1980) Stability of travel components over time, *Transportation Research Record 750*, pp. 19-26.

Zahavi, Y. and Talvitie, A. (1980) Regularities in travel time and money expenditures, *Transportation Research Record 750*, pp. 13-19.

List of figures

List of tables

Table 1
List of variables from the French dataset

Travel mode	Objective evaluation of the travel amount	Subjective estimation of the travel amount	Desired change in the travel amount
A. Variables present in the dataset			
Bicycle	OM_BIKE	SM_BIKE	RDM_BIKE
Motorbike	OM_MOTO	SM_MOTO	RDM_MOTO
Car as driver	OM_CARDR	SM_CARDR	RDM_CARDR
Car as passenger	OM_CARPAX	SM_CARPAX	RDM_CARPAX
Bus	OM_BUS	SM_BUS	RDM_BUS
Tramway	OM_TRAM	SM_TRAM	RDM_TRAM
Metro	OM_METRO	SM_METRO	RDM_METRO
Taxi	OM_TAXI	SM_TAXI	RDM_TAXI
Suburban train	OM_SUBURB	SM_SUBURB	RDM_SUBURB
Long-distance train	OM_TRAIN	SM_TRAIN	RDM_TRAIN
B. Variables built from those above			
Car	OM_CAR	SM_CAR	RDM_CAR
Public transport	OM_PT	SM_PT	RDM_PT
Composite over all modes	OM_GLOB	SM_GLOB	RDM_GLOB
Measures	*Metric variables: estimated number of monthly trips*	*"I feel I don't travel" (1) … "I feel I travel a lot" (10)*	*"I'd like to travel much less" (– – –) … "I'd like to travel the same" (=) … "I'd like to travel much more" (+++)*

Table 2
List of variables from the U.S. dataset

Travel mode	Objective evaluation of the travel amount	Subjective estimation of the travel amount	Desired change in the travel amount
A. Variables present in the dataset			
Personal vehicle	OM_CAR	SM_CAR	RDM_CAR
Bus	OM_BUS	SM_BUS	RDM_BUS
Train/Light rail	OM_RAIL	SM_RAIL	RDM_RAIL
Walk/jogging/cycling	OM_WALK	SM_WALK	RDM_WALK
B. Variables built from those above			
Public transport	OM_PT	SM_PT	RDM_PT
Composite over all modes	OM_GLOB	SM_GLOB	RDM_GLOB
Measures	*Metric variables: estimated number of weekly hours spent on that mode*	*"I feel I don't travel" (1) … "I feel I travel a lot" (5)*	*"I'd like to travel much less" (– – –) … "I'd like to travel the same" (=) … "I'd like to travel much more" (+++)*

Table 3
A-type cluster solution and group means, French dataset

Variable	Group 1 Car-oriented	Group 2 Transit-oriented	Group 3 Neither-oriented	Group 4 Both-oriented
OM_GLOB	25	40	19	55
OM_CAR	18	4	3	16
OM_PT	3	30	10	30
SM_GLOB	*-5*	*10*	*-39*	*24*
SM_CAR	*31*	*-60*	*-82*	*13*
SM_PT	*-44*	*66*	*-6*	*62*
RDM_GLOB	*0*	*2*	*17*	*-10*
RDM_CAR	*-18*	*22*	*43*	*8*
RDM_PT	*26*	*-30*	*-1*	*-45*
Number of cases (164)	88	33	14	29

NB: The first three rows report the group means of the sum of the number of monthly trips taken with the considered modes; the remaining rows the group means of the scores obtained combining SM and RDM ordinal measures.

Table 4
A-type cluster solution and group means, U.S. dataset

Variable	Group 1 Heavily car-oriented	Group 2 Rather car-oriented	Group 3 More transit-oriented	Group 4 Light travellers
OM_GLOB	17	9	15	4
OM_CAR	13	7	3	3
OM_PT	1	0	9	1
SM_GLOB	*267*	*-26*	*321*	*-118*
SM_CAR	*615*	*270*	*-589*	*-221*
SM_PT	*-79*	*-328*	*1242*	*-15*
RDM_GLOB	*-2*	*95*	*-236*	*-21*
RDM_CAR	*-289*	*-49*	*28*	*100*
RDM_PT	*188*	*156*	*-530*	*-43*
Number of cases (1904)	185	696	205	818

NB: The first three rows report the group means of the sum of the hours spent using the considered modes, the remained the group means of the scores obtained combining SM and RDM ordinal measures.

Table 5
B-type cluster solution and group means, French dataset

	Group 1 Very light travelers, monomodal car users, deprived re other modes	Group 2 Light travelers, car-dominated but multimodal, deprived re other modes	Group 3 Moderate travelers, highly multimodal, mixed satisfaction	Group 4 Heavy travelers, highly multimodal, generally surfeited
OM_GLOB (std)	-0.9	-0.6	0.7	1.2
OM_PI (std)	-1.5	-0.4	0.7	0.8
OM_BIKE	*1*	*2*	*3*	*1*
OM_MOTO	*0*	*0*	*1*	*1*
OM_CARDR	*11*	*11*	*7*	*10*
OM_CARPAX	*4*	*3*	*4*	*7*
OM_BUS	*1*	*2*	*9*	*11*
OM_TRAM	*1*	*0*	*3*	*4*
OM_METRO	*2*	*4*	*10*	*11*
OM_TAXI	*0*	*0*	*1*	*0*
OM_SUBURB	*0*	*0*	*1*	*4*
OM_TRAIN	*0*	*1*	*2*	*1*
SM_GLOB (std)	-1.5	-0.1	0.6	0.4
SM_PI (std)	-1.7	0.0	0.6	0.3
SM_BIKE	*1*	*3*	*4*	*3*
SM_MOTO	*1*	*1*	*2*	*2*
SM_CARDR	*7*	*8*	*6*	*8*
SM_CARPAX	*4*	*4*	*5*	*6*
SM_BUS	*2*	*4*	*6*	*7*
SM_TRAM	*1*	*2*	*3*	*4*
SM_METRO	*2*	*4*	*7*	*7*
SM_TAXI	*1*	*2*	*2*	*2*
SM_SUBURB	*1*	*2*	*3*	*3*
SM_TRAIN	*2*	*3*	*5*	*4*
RDM_GLOB (std)	0.2	0.1	0.1	-0.9
RDM_PI (std)	0.5	0.2	0.2	-2.3
RDM_BIKE	++	++	++	++
RDM_MOTO	+	+	+	=
RDM_CARDR	–	– –	–	– –
RDM_CARPAX	+	+	=	–
RDM_BUS	+	+	=	– – –
RDM_TRAM	+	+	+	– –
RDM_METRO	+	+	–	– – –
RDM_TAXI	=	=	=	– –
RDM_SUBURB	=	=	=	– – –
RDM_TRAIN	+	+	=	–
Number of cases	25	59	63	17

NB: std = standardized

Table 6
B-type cluster solution and group means, U.S. dataset

	Group 1 Light travelers, monomodal car users, generally deprived	Group 2 Moderate travelers, multimodal but auto-dominated, transit-surfeited	Group 3 Light travelers, multimodal but auto-dominated, generally deprived	Group 4 Heavy travelers, multimodal but auto-dominated, generally deprived
OM_GLOB (std)	-0.3	0.3	-0.4	1.4
OM_PI (std)	-0.8	0.3	0.3	0.7
OM_WALK	*1*	*2*	*2*	*4*
OM_CAR	*6*	*5*	*4*	*8*
OM_BUS	*0*	*2*	*1*	*2*
OM_RAIL	*0*	*1*	*0*	*2*
SM_GLOB (std)	-0.9	0.6	0.3	0.7
SM_PI (std)	-1.0	0.7	0.5	0.6
SM_WALK	*2*	*3*	*3*	*3*
SM_CAR	*4*	*4*	*4*	*4*
SM_BUS	*1*	*2*	*2*	*2*
SM_RAIL	*1*	*2*	*2*	*2*
RDM_GLOB (std)	-0.2	-0.6	0.3	0.0
RDM_PI (std)	0.3	-2.7	0.2	0.1
RDM_WALK	+ +	+	+ +	+ +
RDM_CAR	+	+	+ +	+
RDM_BUS	=	– – –	=	=
RDM_RAIL	+	– –	+ +	+
Number of cases	642	159	771	332

NB: std = standardized

Table 7
Thresholds in defining the 7-point RDM scale

Grade of the scale	Range of RDM means from the French dataset (the original scale has 11 points)	Range of RDM means from the U.S. dataset (the original scale has 5 points)
– – –	[1.00 ; 4.00[	[1.00 ; 2.20[
– –	[4.00 ; 5.30[	[2.20 ; 2.72[
–	[5.30 ; 5.80[	[2.72 ; 2.92[
=	[5.80 ; 6.20]	[2.92 ; 3.08]
+	]6.20 ; 6.70]	]3.08 ; 3.28]
+ +	]6.70 ; 8.00]	]3.28 ; 3.80]
+ + +	]8.00 ; 11.00]	]3.80 ; 5.00]

Figure 1. Hierarchical structure of the modes in the French dataset; grey modes are the observed ones

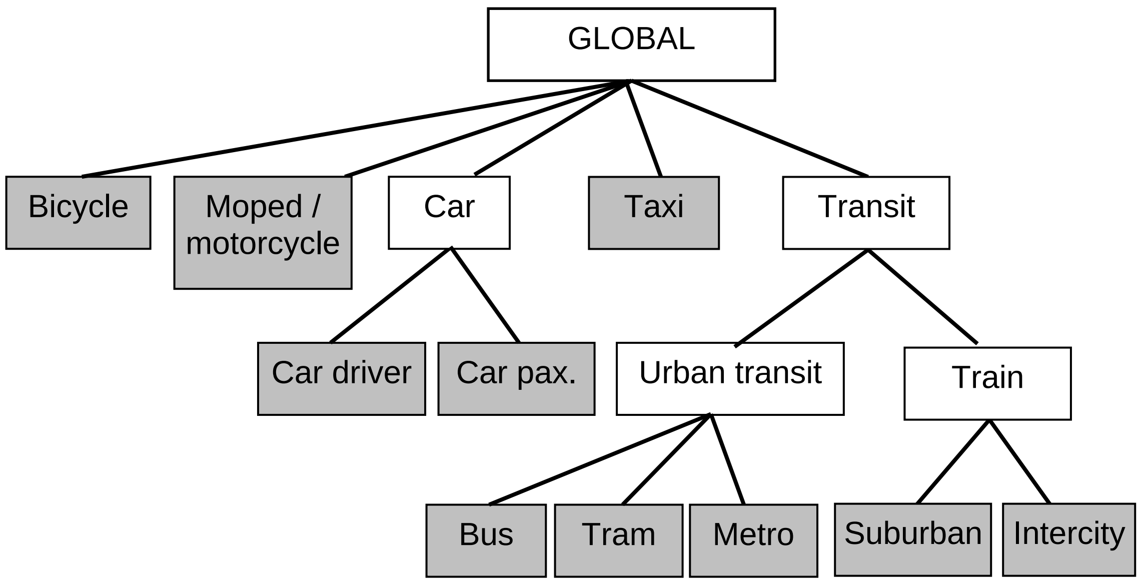

Figure 2. Hierarchical structure of the modes in the US dataset; grey modes are the observed ones

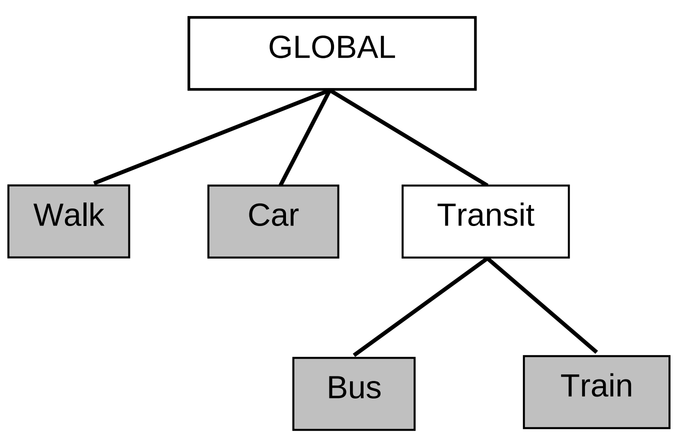

Figure 3. Plot of OM_PI for two modes and maximum frequency M = 15

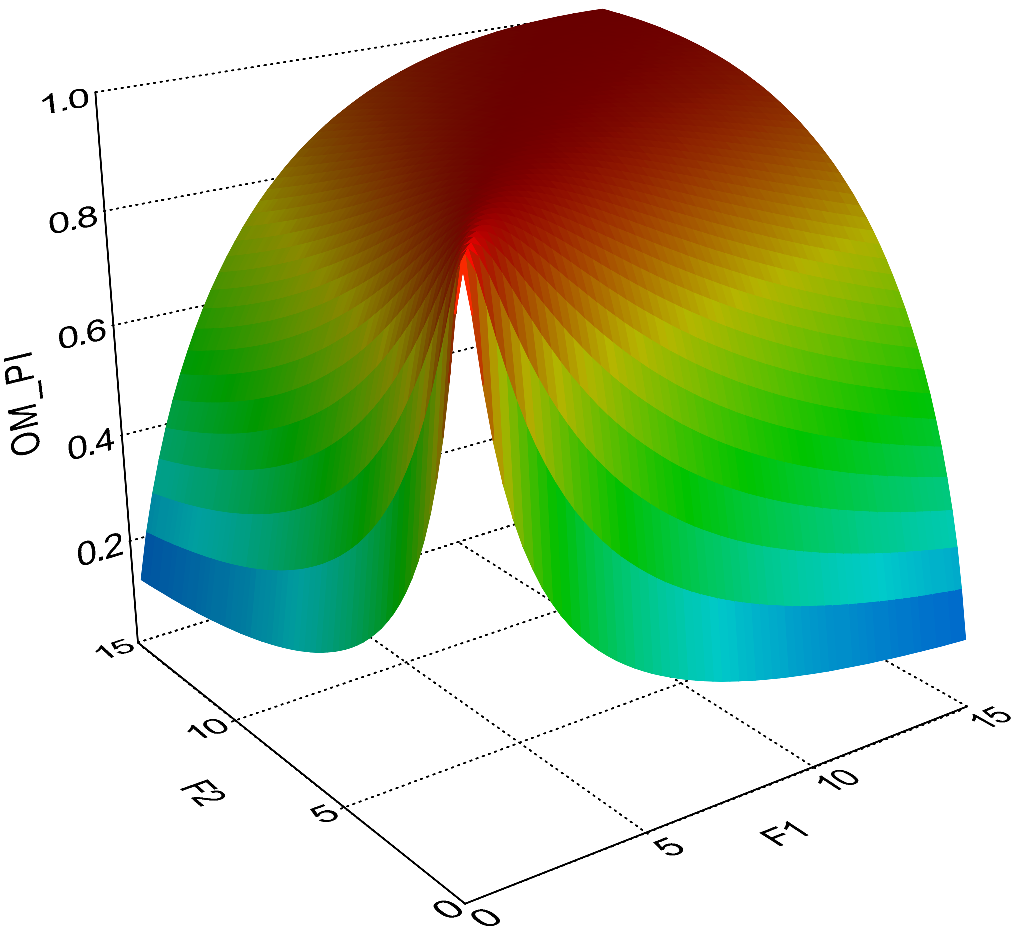

Figure 4. Plot of OM_MI for two modes and maximum frequency M = 15

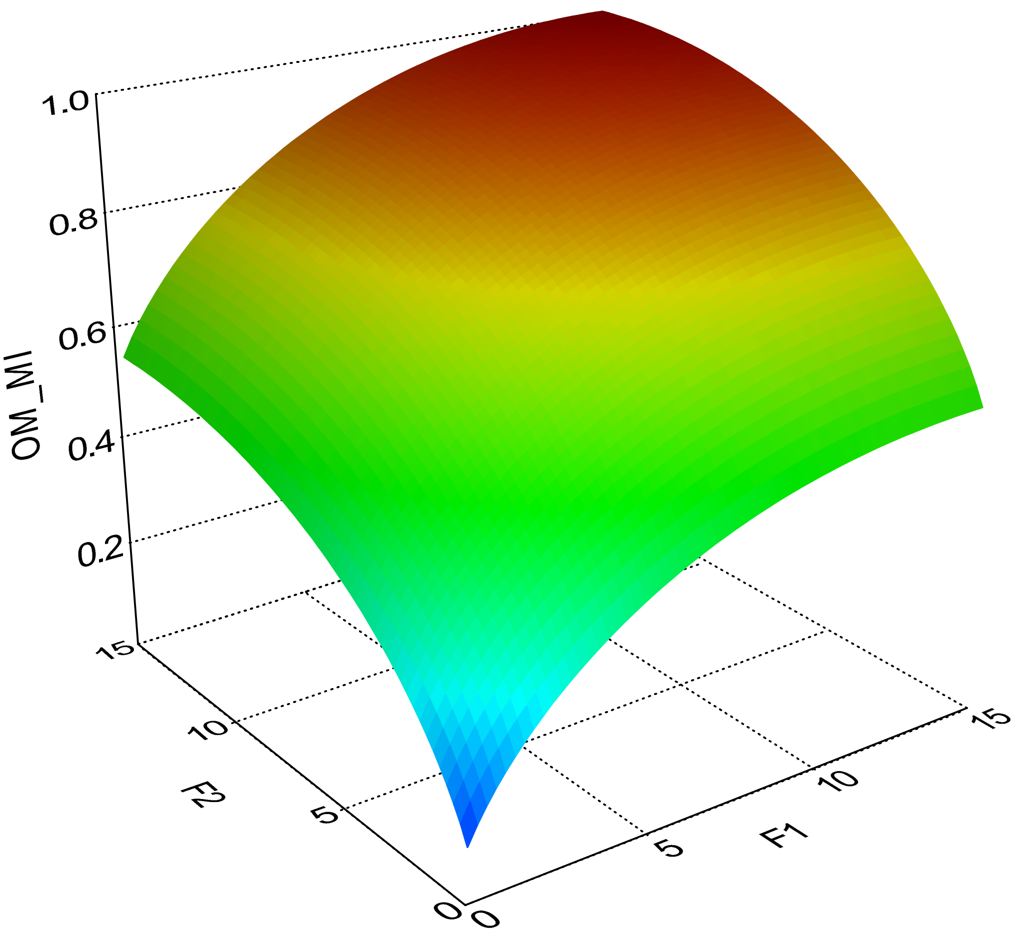
1.0
0.8
0.6
0.4
0.2
OM_MI
15
10
5
0
F2
0
5
10
15
F1

Figure 5. Number of personal vehicles per household across the A-type clusters

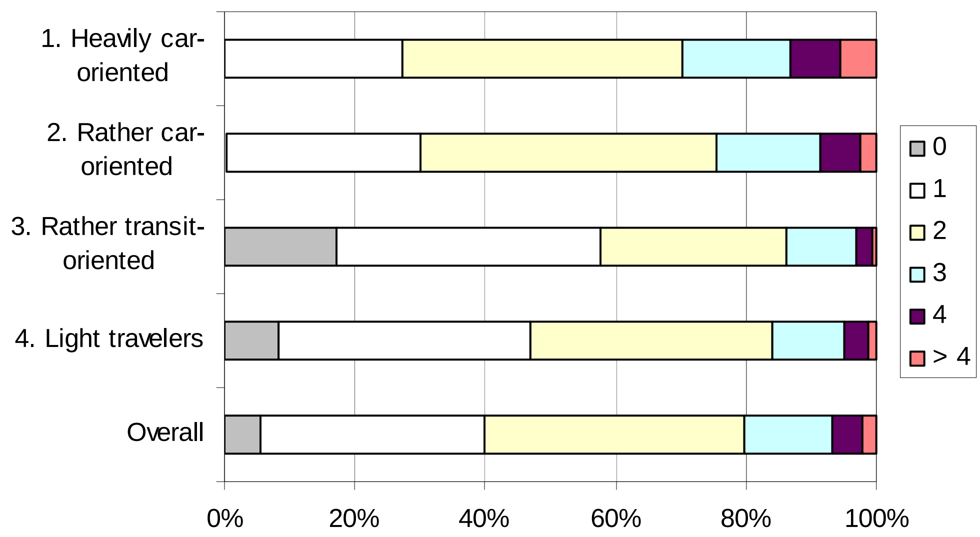

Figure 6. Percentage of time a personal vehicle is available to the respondent across the A-type clusters

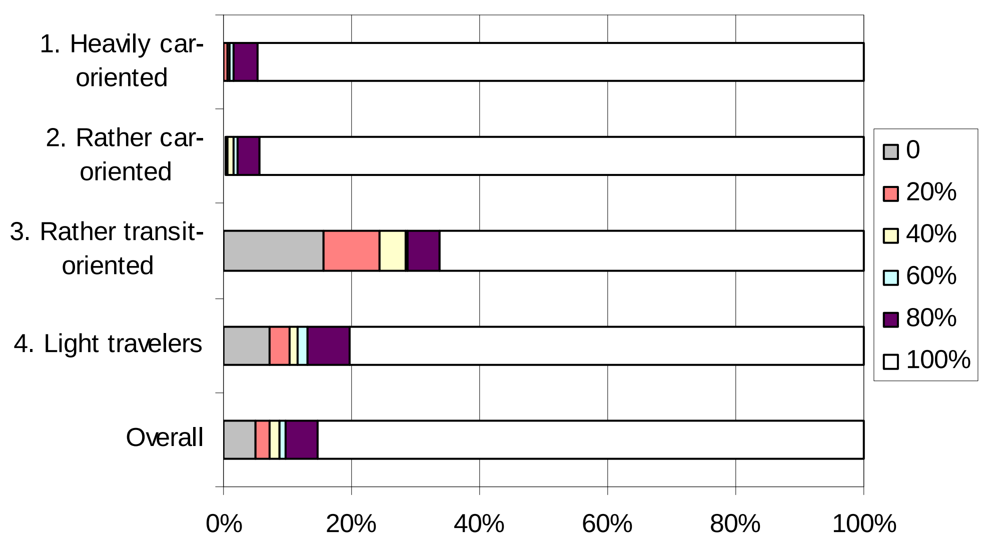

Figure 7. Educational level of the respondent across the A-type clusters

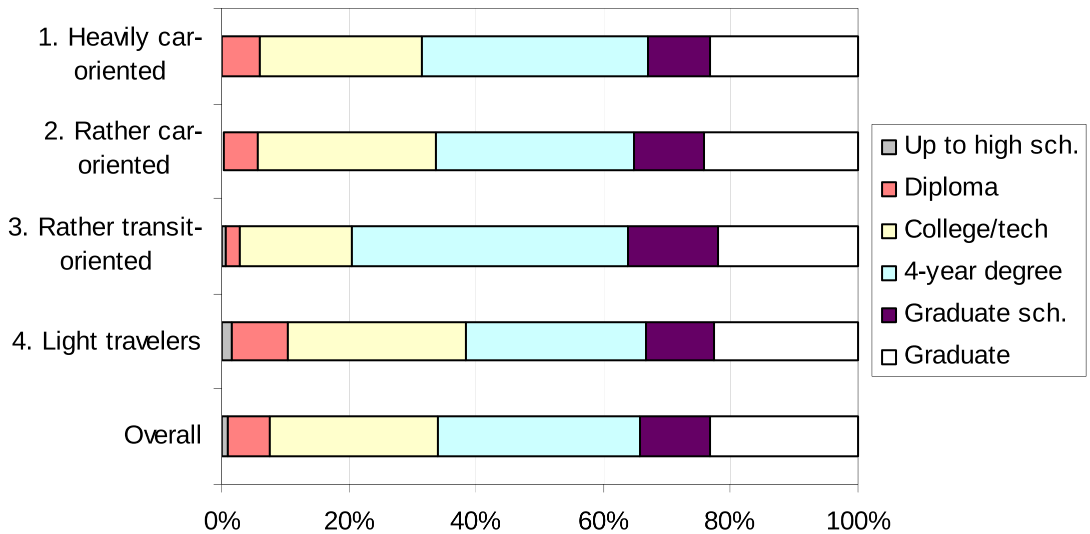

Figure 8. Household income across the A-type clusters

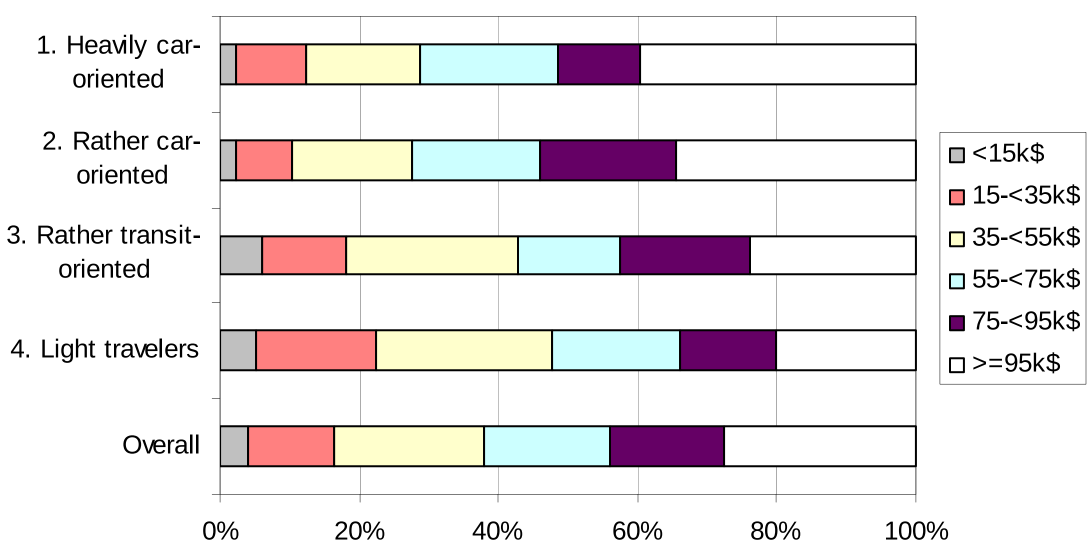

Figure 9. Number of personal vehicles per household across the B-type clusters

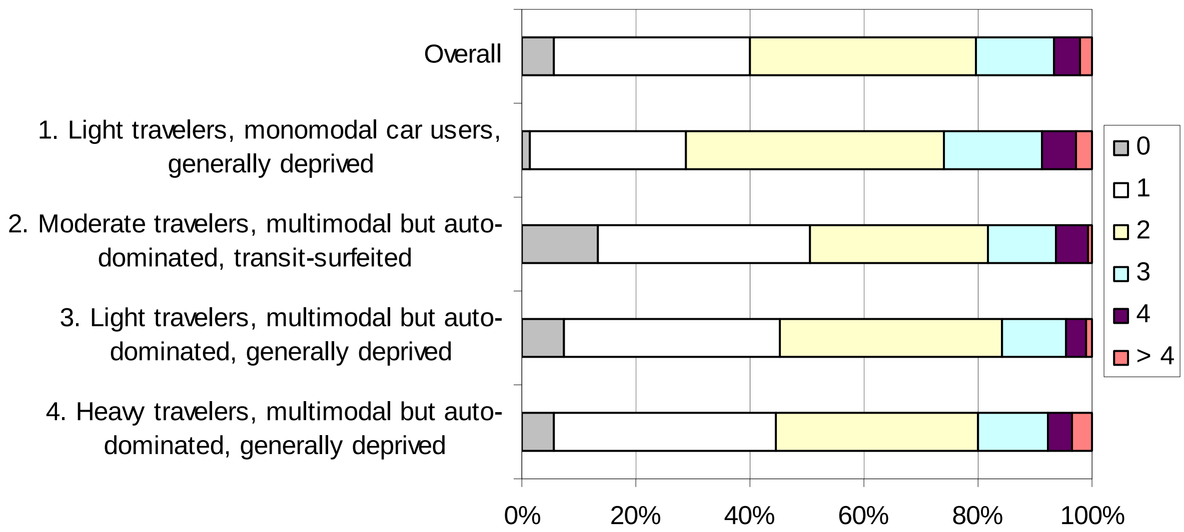

Figure 10. Percentage of time a personal vehicle is available to the respondent across the B-type clusters

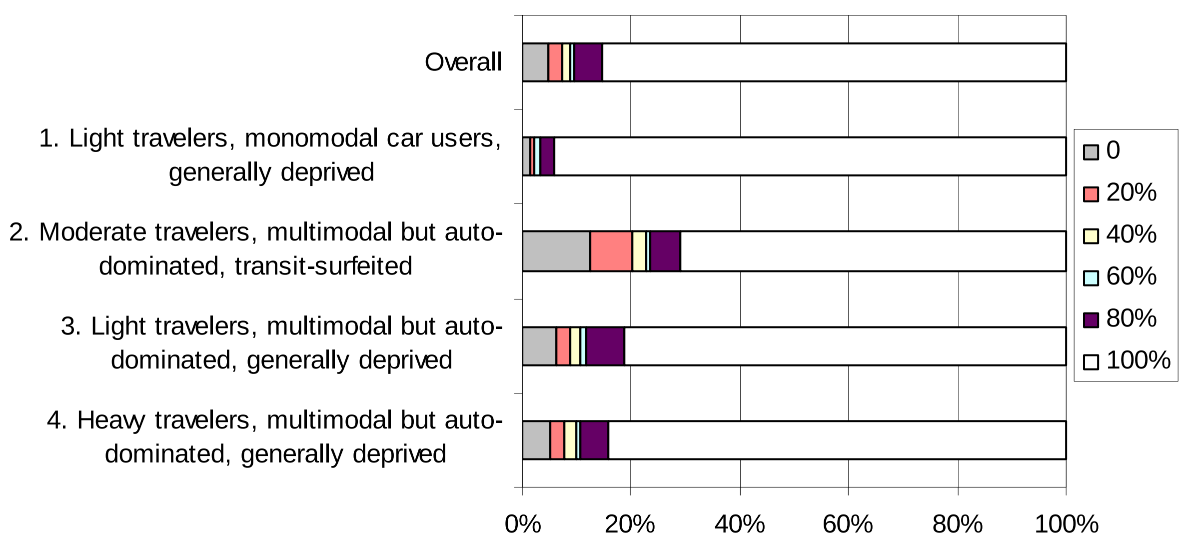

Figure 11. Educational level of the respondent across the B-type clusters

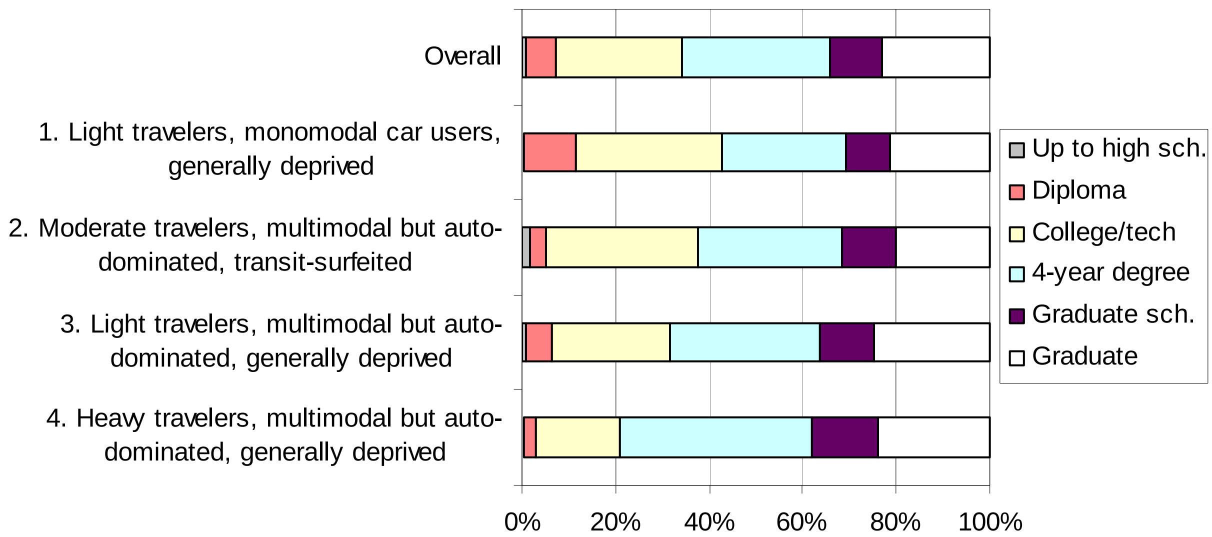

Figure 12. Household income across the B-type clusters