

## Grouping travelers on the basis of their different car and transit levels of use

Marco Diana · Patricia L. Mokhtarian

Published online: 26 April 2009

© The Author(s) 2009. This article is published with open access at Springerlink.com

**Abstract** Market segmentation studies in travel behavior research are ordinarily based on socioeconomic characteristics and personality traits. This study explores the usefulness of a different approach, where the actual overall mobility levels across different ground transportation modes, along with desired changes in the use of cars and transit, are used as clustering variables. Using a given mode can in fact influence the personal representation of that mode, which in turn has been proven to be a key element in transport behaviours. We form such multimodality-based clusters from two field studies, one involving employees of the French transportation research institute INRETS and the other a representative sample of residents of the US San Francisco Bay Area. We find that strong users of a given mode would like to bring more balance to their “modal consumptions” by decreasing the use of this mode more than the average, and increasing the use of the alternative mode. However, concerning ground transport travel budgets, the desire to travel more (or less) *overall* seems less strongly related to the composition of the modal balance. The US dataset shows also a greater latent demand for travel than the French one. Socioeconomic characteristics of the clusters could not explain the patterns that were found, confirming the importance of taking into account multimodality issues in travel behavior research. Some policy implications from these findings are finally reported.

**Keywords** Cluster analysis · Desired mobility · Market segmentation · Multimodality

---

M. Diana (✉)

Dipartimento di Idraulica, Trasporti e Infrastrutture Civili, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Turin, Italy  
e-mail: marco.diana@polito.it

P. L. Mokhtarian

Department of Civil and Environmental Engineering, Institute of Transportation Studies,  
University of California, Davis, One Shields Avenue, Davis, CA 95616, USA  
e-mail: plmokhtarian@ucdavis.edu

## Introduction

Helping stakeholders to set up transport policies effective in making travel behaviors more sustainable is one of the central issues of current research agendas. In this endeavour, researchers often rely on methods from disciplines where the study of the determinants of individual behaviors, and of the most effective ways to influence them, has been undertaken for a long time.

One analytical technique that is increasingly gaining attention is market segmentation, or customer classification, where subsets of people that share some common characteristics are identified within a larger sample, in general through multivariate analysis techniques. The researcher has to select the characteristics that define the groups on the basis of the objectives of the study. In the transport field, the key policy interest of such marketing studies is in general to understand which people could be more inclined to shift from individual motorized transport modes to transit or non-motorized ones. Segments are subsequently defined on the basis of various socioeconomic characteristics on one hand, or personality traits on the other (e.g., Transportation Research Board 1977; Dobson and Tischer 1978; Tardiff 1979; Gensch and Torres 1980; Jensen 1999; Outwater et al. 2003; Elgar and Bekhor 2004; Anable 2005). However, if the final research goal is to cluster travellers in order to maximize the differences between groups concerning the responses to policy actions aimed at modal diversion, one might wonder if also other clustering variables could be relevant.

The present study considers the role that the degree of acquaintance with different transportation modes could have in influencing modal choices. Our perspective is related to previous research dealing with the role of habit on mode choice, yet a conceptual difference is detectable. Habit can be seen as a behavioural mechanism that goes against the rational decision-making process as depicted in standard microeconomics theory, since it tends to make consumer choices less deliberate. Specifically, the sufficient degree of acquaintance with all the modes in a choice set can be seen as a basic element of one of the assumptions of that theory, namely the complete knowledge of all the alternatives by the decision maker. By contrast, it is in fact well known that, for example, exclusive car users tend to overestimate travel times by public transport, and in general to have a biased knowledge of the potential transit alternatives for their trips (O'Farrell and Markham 1974; Fujii et al. 2001; Beale and Bonsall 2007).

To the best of our knowledge, very few works in the published literature have investigated the importance that being familiar with different transport means can have in shaping desires and ultimately influencing behaviors. Related previous research findings have nevertheless shown the decisive role of the predictability (Anable and Gatersleben 2005) and of individual representations (Guiver 2007) of a transport means as key elements to understand mode choice. We can argue that both of these are strongly influenced by the relative familiarity with a means, i.e., the degree of acquaintance with a given mode compared to the degree of acquaintance with the alternative modes. This multimodal perspective is essential but it is seldom considered, even though, for example, a strong car user who travels also by transit is likely to develop car and transit attitudes that are different from those of a strong and exclusive car user. This paper focuses exclusively on these elements, not considering several other factors regarding mobility behaviors whose importance is well illustrated by previous research, such as individual attitudes, socioeconomic status, situational variables (availability of personal vehicles, accessibility to public transport systems, land use and activity patterns) and instrumental ones (performances of the different means in terms of costs, travel times, service quality). We only

look at some of the most influential of the above elements, namely car ownership and availability, education and income, in the socioeconomic characterization of the clusters that we later present.

In the following we present a segmentation study that seeks to operationalize such a multimodal perspective by considering the actual intensity of use of different modes as clustering variables. Then we look at the perceived levels of use of such modes and at the desire to change them, within each cluster. Consistently with previous research in the field (Choo et al. 2005), we will name these three different mobility measures Objective Mobility (OM), Subjective Mobility (SM) and Relative Desired Mobility (RDM). Hence we define the groups solely on the basis of OM, a measure that can be found in any mobility survey, so that our clustering procedure can be more easily applied in different contexts. In doing that, we assume that the degree of acquaintance with a mode can be measured by its actual level of use, but then we assess whether the segments that we defined are relevant from a policy point of view by looking at the SM and RDM means. These three variables are of course subject to measurement errors and potential bias. We preliminarily notice that in our research framework SM is not a proxy for OM, since it is actually measuring a different construct. SM deals with the personal evaluation of the amount of travel that has been consumed; as such, it is obviously related to OM. However, different individuals could be differently “fed up” with the same mobility level, and SM is aimed at investigating such difference. OM in turn is based on self-reporting, as we detail in the following section, so that a more accurate measurement method for this variable would involve the use of mechanical devices such as odometers or GPS receivers, which were not used in the present research.

Our methodology should give us insights into a number of questions of interest. For example, if the groups show different desires concerning their future levels of use of a set of modes, then such different desires can be traced back to the actual modal consumptions. This would allow us to better understand how mode use habits are associated with desires, which in turn can be seen as one of the determinants of behaviors. On the other hand, the OM-SM relationship could be differentiated across the groups, thus pointing to a perception of the actual modal use that is differently distorted according to the “modal mix”. It would then be possible to check if the intensity of use of some modes is systematically over- or understated for some specific level of use, thus giving interesting insights into potential mode-specific biases that should be accounted for in modelling practices.

## Datasets and clustering variables

Cluster analysis is an exploratory statistical technique that allows for grouping observations in a dataset, and it is often used in segmentation studies. However, the distilled solutions are not unique for a given dataset, nor is there a commonly-accepted goodness of fit measure to discriminate among them, so that the generalizability of the results is always a critical matter. To have a richer perspective, we consider two datasets coming from radically different experimental settings. The older study was conducted in 1998 with a postal survey sent to 8,000 residents in three different communities in the San Francisco Bay Area—Concord and Pleasant Hill, representing two different kinds of suburban neighborhoods comprising about half the sample, and an area defined as North San Francisco representing an urban neighborhood—which resulted in 1904 useful observations (Mokhtarian and Salomon 2001). Then we consider a dataset from an Internet survey that was administered in 2004 to people working at the French National Institute for

Transport and Safety Research (INRETS), with two larger premises in the Paris and Lyon metropolitan areas and two smaller branches near Lille and Marseille. 164 completed surveys were thus gathered out of about 550 workers and students who were contacted (Diana 2005).

Both datasets investigated the previously defined actual (OM), perceived (SM) and desired (RDM) mobility levels of the respondents by different transport modes, among other things. The variables that we enter in our cluster analysis are then derived from the variables of both datasets as follows.

Considering first OM measures, people were asked in the US survey for their typical weekly mileage by each of the following four transport modes: (1) driver or passenger in any personal vehicle, (2) bus, (3) rail and (4) walking, jogging, cycling. Only trips up to 100 miles one way were examined in this way, since longer ones were measured differently and are not considered here. Then we use these data to estimate the number of weekly hours spent in each mode through a best guess of their typical speeds. This is because a time-based OM measure is probably a better proxy of the familiarity of the respondent with a transport mode than a distance-based one. We would say that, for example, a person using a bike 2 h per day is more acquainted with bikes than a car driver using a car 2 h a week is with his car, although the weekly mileage could be comparable given the different mean speeds of the two modes.

The French survey asked for the mean frequency of use of a larger set of transport modes 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). We take the following five modes among those considered in the survey: car driver, car passenger, bus, tram and metro, and we compute for each of these the number of monthly trips.

For the sake of simplicity, in the following we focus on the relative levels of use of cars and public transport modes; so that we need to condense the above described more disaggregate information regarding the objective mobility levels of different modes. Considering a larger set of modes would be interesting, but it is advisable to keep the number of variables that we consider in cluster analysis as low as possible in order to ease the interpretation of the results in this exploratory research. Extensions of the present work, which we further discuss in the conclusions, will be aimed at setting up a method that allows for taking into account a larger number of transport means. Hence we consider in the French dataset the sum of the frequency of driving a car and traveling by car as a passenger 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. In doing that, reported frequencies on ordinal scales have been transformed to monthly frequencies by assuming that “3 times a week or more” can be considered as 15 trips a month, “1–2 times a week” 8 times a month, “1–3 times a month” 2 times a month and “sporadically” 0.5 times a month. For the US dataset we already have a single variable for car use (OM\_CAR) and we define OM\_PT as the sum of the number of hours spent traveling by bus and by train, keeping in mind that long-distance trips were not considered. Finally, OM\_GLOB is the sum of the number of weekly hours (for the US case) or of monthly trips (for the French case) spent in each ground transport mode that was investigated in the survey (i.e., four modes for the US dataset and ten modes for the French one).

Turning now our attention to SM and RDM measures, we preliminarily notice that they were measured in both datasets through rating scales. 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 considered modes were obviously the same as for OM. The US scales have 5 points, whereas the French ones have 10 points when SM is measured and 11 points when RDM is measured.

Also in this case, the SM and RDM variables that we use in the analysis are then derived by aggregating the information available for each of the considered modes. However, one important difference is given by the fact that we cannot simply sum the observations coming from different modes, as done in the OM case, since we are now dealing with ordinal variables. Hence we adopt an aggregation method for ordinal measures, based on their ranks, that is available in the published literature (Wittkowski et al. 2004) and is now being applied also in the transport sector (Diana et al. 2009). In this approach, the combined ordinal measure for an observation A is a score given by the number of observations in the sample that are smaller than A minus the number of observations that are greater than A. According to this method, observation A is greater than B if A has measures at least as high as B on all categories (in our case, the transport modes), and strictly higher on at least one. For the sake of brevity, we refer the reader to the above-mentioned two references for more explanation. We can thus build for the French dataset a new variable, namely SM\_CAR, by combining the two measures relative to driving a car and being a passenger in a car. Analogously, SM\_PT is defined through the combination of the SM measures for bus, trams and metros, and similarly for RDM\_CAR and RDM\_PT. In the US dataset we already have unique SM and RDM measures for cars and we will define SM\_PT and RDM\_PT just considering the bus and the rail modes.

We show in Table 1 the nine variables that we defined and that we will use in the subsequent analysis. The table also shows how they are derived from the original variables of the datasets and their measurement units. Note that OM measures are already available from any travel diary survey. SM and RDM questions similar to the ones we used could

**Table 1** Variables derived from the two datasets

	Objective evaluation of the travel amount	Subjective estimation of the travel amount	Desired change in the travel amount
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
Method applied to the original measures	Summing the original measures	Combining the original measures according to ranks	Combining the original measures according to ranks
Original mode-specific measures in the French dataset	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” (−5)... “I’d like to travel the same” (0)... “I’d like to travel much more” (+5)
Original mode-specific measures in the US dataset	Metric variables: estimated number of weekly hours spent in that mode	“I feel I travel none” (1)... “I feel I travel a lot” (5)	“I’d like to travel much less” (−2)... “I’d like to travel about the same amount” (0) ... “I’d like to travel much more” (+2)

easily be added to a background or other module accompanying such a survey, to implement our study on a larger scale.

### Cluster analysis

On the basis of the methodological framework sketched in the Introduction, we define groups of travelers from the three OM variables listed in the first column of Table 1. Then, in order to enrich the interpretation of the results, we assess the group means for each of the other six variables, related to SM and RDM. In doing that we focus on the general balance between car and transit within the actual level of use of different modes, as well as respondents' perceptions and desired direction and magnitude of change in their actual "modal portfolio".

Clusters interpretation in the two samples

Tables 2 and 3 present the resulting cluster centroids for the French and the US datasets, respectively. Solutions with four clusters are provided here since they are the most informative for our research problem. We label the resulting segments on the basis of the three objective mobility variables around which the clusters were formed, whereas the other values 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 US OM measures are also not directly

**Table 2** Cluster solution and group means for the 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

The first three rows report the group means of the total 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 3** Cluster solution and group means for the US 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	1,242	−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

The first three rows report the group means of the total estimated weekly hours spent using the considered modes, the remainder the group means of the scores obtained combining SM and RDM ordinal measures

comparable, given the different number of considered modes and the different measurement units for OM in the two samples, as shown in Table 1.

OM measures in Table 2 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 the previous section) 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 need 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.

In Tables 2 and 3, 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 detailed by Diana et al. (2009). According to this method, the individual scores can theoretically range from  $-(n - 1)$  to  $(n - 1)$  if we have  $n$  observations in the sample, lower scores indicating lower SM and RDM levels of that individual in comparison with the whole sample. Thus, in interpreting those scores one must remember that their scale differs with the sample size, and only their relative values matter. It is also 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 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 reported in Diana and Mokhtarian (2007) under the name of “A-type analysis”, where the same cluster definition is adopted but with a different method to combine SM and RDM measures, which hence leads to a different cluster interpretation. That 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 what is reported in the following, since here the focus of the analysis is on the above-mentioned horizontal comparisons.

Table 2 shows a strong correlation by mode between OM and SM measures, confirming previous findings (Collantes and Mokhtarian 2007). Different perceptions concerning mobility levels on the basis of the actual mode use, a relationship we postulated in the introduction, was not observed in our clusters. RDM group means show instead 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 balance.

Repeating the same kind of analysis on the US 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 US 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 the preceding section. Another interesting disparity between the two samples is that the US analysis did not distill a group that intensively uses both means.

SM measures have slightly lower correlations with OM measures compared to the French sample. This is probably due to the coarser SM scale used by the US sample, and to behavioral differences as well. In particular, the US 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, where group 1 also travels more than group 3, but subjective mobility is congruent with objective mobility. This could be ascribed 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. Mode-specific measurement errors in our OM measures (for example, an overestimation of public transport speeds when computing the hours spent traveling by that mode, which would in turn lower OM\_GLOB for the US group 3) could also play a role in weakening these correlations.

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 US 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, lower OM-SM and OM-RDM correlations in the US datasets could be indications that the latent demand for travel is greater in this case, a result that is probably linked to the more diverse



socioeconomic composition of the sample. This could also be due to the lower levels of use of transit, so that the desired modal balances of the two samples are probably closer to each other than is the actual use of different means.

### Socioeconomic characteristics of the US clusters

Characterizing the traveler groups that we found by looking at some key socioeconomic variables is useful to assess the added value of our methodology compared to other market segmentation studies that do not consider multimodality measures. The French sample is, however, not representative of a general population and we have a small number of observations for three out of the four clusters, so that such analysis is not so insightful. In contrast, characterizing the clusters of the US dataset is an effective way to understand how personal traits relate to multimodality attitudes and behaviors. Concerning US clusters, it is particularly interesting to study the differences between the first two clusters on one hand (heavily versus rather 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, since very different situations are probably represented there (e.g., retired, unemployed, poor or sick persons, partners of affluent persons that do not work, parents caring for their children at home...). On the other hand, they offer less insight concerning multimodality behaviors, because of the low mobility levels which make them less easily observable with an acceptable relative measurement error.

We performed Kruskal-Wallis tests to ascertain that the group differences that we found were not simply due to random variation. The corresponding *p*-values are always well below .01, so that we can safely reject the null hypothesis of no difference among the group populations. We used a non-parametric test because three out of these four socioeconomic variables are ordinal rather than continuous, and (more importantly) are not always approximately normally-distributed, so that a standard analysis of variance (ANOVA) would not necessarily be appropriate.

The socioeconomic variables that we consider, together with the Kruskal-Wallis test statistics and their corresponding *p*-values assuming a Chi-square distribution with three degrees of freedom, are the following:

- CAR\_NO represents the number of available cars in the household (108.3,  $p < .001$ ).
- CAR\_AVAIL is the percentage of time a personal vehicle is available to the respondent. Respondents could choose among six pre-defined values (143.8,  $p < .001$ ).
- EDUCAT represents the educational level of the respondent. Six different levels were indicated in the questionnaire (11.7,  $p = .008$ ).
- INCOME indicates the income of the household. Six income brackets were specified (86.3,  $p < .001$ ).

Categories for each of the above variables are listed in the first column of Table 4. The remainder of the table shows the number of cases for each category by group and for the whole sample and the corresponding row percentages. Concerning CAR\_NO, respondents who indicated more than four cars have been grouped into one category.

Concerning CAR\_NO and CAR\_AVAIL, 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-

**Table 4** Contingency table of the considered socioeconomic variables and within-group row percentages

Variable	Category	Group 1 Heavily car-oriented	Group 2 Rather car-oriented	Group 3 More transit-oriented	Group 4 Light travellers	Overall sample
CAR_NO	0	0	3	35	67	105
	1	50	206	82	315	653
	2	79	316	58	303	756
	3	31	111	22	92	256
	4	14	41	5	29	89
CAR_AVAIL	More than 4	10	18	1	10	39
	0%	0	3	32	58	93
	20%	1	1	18	26	46
	40%	1	7	8	11	27
	60%	1	4	1	11	17
EDUCAT	80%	7	25	10	54	96
	100%	174	654	136	651	1,615
	Up to high school	0	2	1	12	15
	High school diploma	11	37	5	73	126
	College/tech school	47	194	36	229	506
INCOME	4-year degree	66	216	89	232	603
	Graduate school	18	77	29	87	211
	Graduate degree	43	168	45	185	441
	Less than \$ 15,000	4	15	12	40	71
	\$ 15,000–\$ 34,999	18	54	24	132	228
Number of cases per group	\$ 35,000–\$ 54,999	29	115	49	196	389
	\$ 55,000–\$ 74,999	35	125	29	143	332
	\$ 75,000–\$ 94,999	21	130	37	106	294
	\$ 95,000 or more	70	231	47	155	503
		185	696	205	818	1,904

The numbers of cases per category do not always sum up to these values due to non-response, in particular for the INCOME variable. Row percentages are computed excluding missing data

oriented cluster. It would be interesting to explore whether the greater symbolic value of cars for males (Steg 2005) can explain part of this difference. Another interesting extension of this study related to modal availability could focus on public transport accessibility for the sample, which surely plays a role in shaping attitudes and the degree of familiarity with the different means. However, such a study would require more detailed data analyses as well as geocoding, which was not available for the datasets we used.

Considering EDUCAT, 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. On the other hand, these latter have a lower income than the average; 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 Table 4. To sum up, the use of a multimodality-based clustering technique can give results not easily reproduced by segmentations based on socioeconomic characteristics, since patterns of modal usage are generated from both instrumental and affective factors. In fact, cluster analyses must be based on a limited number of variables in order to give interpretable results, so that a socioeconomic segmentation based on all the relevant factors (including land use, activity patterns etc..) would be in any case rather difficult to manage. In such situations, our method of looking just at multimodality patterns can give results otherwise not easily obtainable in a different way.

## Policy implications and conclusions

We believe that the findings of this study are rather insightful concerning some relevant policy questions. For example, concerning the issue of the existence of individual travel time budgets (Mokhtarian and Chen 2004), the OM-SM-RDM patterns of values across groups in Tables 2 and 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 “modal consumptions” by decreasing the use of this mode more than the average, and increasing the use of the alternative mode. Subsequent research should be aimed at clarifying if this is exclusively due to situational variables (for example, limited accessibility to public transport for strong car users and limited availability of a private means for transit users), or if attitudes and self-related factors also play a role. In more general terms, it seems in any case 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 US groups 3 and 4 are the same for cars, but the corresponding RDM levels are quite different.

Turning our attention to objective and relative desired global mobility levels, other interesting patterns emerge. Restricting our analysis to multimodality behaviors and setting aside the underlying possible explanations that we mentioned in the preceding paragraph, we note for example that the US group 3 desires to travel dramatically less than the US group 1, even if the two OM\_GLOB values are comparable. The decisive difference is that 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 modal balance than with one's global mobility levels and that (2) when the modal balance 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 US dataset (although the northern Californians comprising the US sample would in turn be expected to be more environmentally aware and proactive than the country as a whole).

The analysis that we presented shows the importance of considering multimodality behaviors to understand how the use of different transport modes impacts desires and ultimately transport trends and scenarios, in addition to the more conventionally-studied socioeconomic impacts. The general policy target of maximizing modal diversion to transit and non motorized modes can be probably best reached by adapting strategies on the basis of actual modal balances, thus completing the indications that come from car-use reduction theories where the stress is on the levels of use of cars. In other words, modal diversion strategies are more effective if they combine car use reduction targets with alternative modes promotion, thus better supporting the aforementioned observed tendency to equilibrate modal baskets. This finding can be related to the well-known "carrot and stick" approach that has been long recommended over more partial interventions. On the other hand, differences in the ideal mobility levels need to be taken into account and call for policy actions possibly differentiated according to the different segments that we found.

Future research efforts in this direction will be aimed at better defining which policy interventions are best suited for each market segment. It would also be interesting to broaden the range of modes, including nonmotorized means and perhaps separately considering different kinds of public transport. This would require a method to synthetically express the levels of use of different transport means with a single index, in order to keep the number of variables to be considered in cluster analysis at an acceptable level (Diana and Mokhtarian 2007). The authors have started applying such a method in a related paper (Diana and Mokhtarian 2009), which is a follow-up to the present one. Finally, seminal work is trying to assess the relative importance of multimodality behaviors and of other factors, such as attitudes and performances of competing modes, in determining modal choices (Diana 2009). This extension could be useful among other things to adjust cross elasticities between modes by taking into account a wider range of factors, beyond travel costs and times.

**Acknowledgments** Comments from five anonymous referees helped improve an earlier version of this paper.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

## References

- Anable, J.: 'Complacent car addicts' or 'aspiring environmentalists'? Identifying travel behaviour segments using attitude theory. *Transp. Policy* **12**(1), 65–78 (2005)
- Anable, J., Gatersleben, B.: All work and no play? The role of instrumental and affective factors in work and leisure journeys by different travel modes. *Transp. Res. A* **39**(2–3), 163–181 (2005)

- Beale, J.R., Bonsall, P.W.: Marketing in the bus industry: a psychological interpretation of some attitudinal and behavioural outcomes. *Transp. Res. F* **10**(4), 271–287 (2007)
- Choo, S., Collantes, G.O., Mokhtarian, P.L.: Wanting to travel, more or less: exploring the determinants of the deficit and surfeit of personal travel. *Transportation* **32**(2), 135–164 (2005)
- Collantes, G.O., Mokhtarian, P.L.: Subjective assessments of personal mobility: what makes the difference between a little and a lot? *Transp. Policy* **14**(3), 181–192 (2007)
- Diana, M.: An exploratory web-based attitudinal travel survey administered to INRETS staff. Les collections de l'INRETS—“Outils et méthodes” n°12, 187 pp. Arcueil, France (2005). ISBN 2-85782-623-0
- Diana, M.: From mode choice to modal diversion: a new behavioural paradigm and an application to the study of the demand for innovative transport services. Manuscript available from the author, (2009)
- Diana, M., Mokhtarian, P.L.: The individual desire to change a “modal basket” and its relationship with the actual and perceived use of different transport means, 41 pp. 11th World Congress on Transport Research, Berkeley, USA (2007)
- Diana, M., Mokhtarian, P.L.: Desire to change one's multimodality and its relationship to the use of different transport means. *Transp. Res. F* **12**(2), 107–119 (2009)
- Diana, M., Song, T., Wittkowski, K.M.: Studying travel-related individual assessments and desires by combining hierarchically structured ordinal variables. *Transportation* **36**(2), 187–206 (2009). doi: [10.1007/s11116-009-9186-z](https://doi.org/10.1007/s11116-009-9186-z)
- Dobson, R., Tischer, M.L.: Perceptual market segmentation technique for transportation analysis. *Transp. Res. Rec.* **673**, 145–152 (1978)
- Elgar, A., Bekhor, S.: Car-rider segmentation according to riding status and investment in car mobility. *Transp. Res. Rec.* **1894**, 109–118 (2004)
- Fujii, S., Gärling, T., Kitamura, R.: Changes in drivers' perceptions and use of public transport during a freeway closure. *Environ. Behav.* **33**(6), 796–808 (2001)
- Gensch, D.H., Torres, P.T.: Perceived-difference segmentation model for mass transit marketing. *Transp. Res. Rec.* **765**, 16–22 (1980)
- Guiver, J.W.: Modal talk: discourse analysis of how people talk about bus and car travel. *Transp. Res. A* **41**(3), 233–248 (2007)
- Jensen, M.: Passion and heart in transport—a sociological analysis on transport behavior. *Transp. Policy* **6**(1), 19–33 (1999)
- Mokhtarian, P.L., Chen, C.: TTB or not TTB, that is the question: a review and analysis of the empirical literature on travel time (and money) budgets. *Transp. Res. A* **38**(9–10), 643–675 (2004)
- Mokhtarian, P.L., Salomon, I.: How derived is the demand for travel? Some conceptual and measurement considerations. *Transp. Res. A* **35**(8), 695–719 (2001)
- O'Farrell, P., Markham, J.: Commuter perception of public transport work journeys. *Environ. Plan. A* **6**(1), 79–100 (1974)
- Outwater, M.L., Castleberry, S., Shiftan, Y., Ben-Akiva, M., Zhou, Y.S., Kuppam, A.: Attitudinal market segmentation approach to mode choice and ridership forecasting—structural equation modeling. *Transp. Res. Rec.* **1854**, 32–42 (2003)
- Steg, L.: Car use: lust and must. Instrumental, symbolic and affective motives for car use. *Transp. Res. A* **39**(2–3), 147–162 (2005)
- Tardiff, T.J.: Attitudinal market segmentation for transit design, marketing, and policy analysis. *Transp. Res. Rec.* **735**, 1–7 (1979)
- Transportation Research Board: Preferences, perceptions, and market segments in travel behavior. *Transp. Res. Rec.* **649**, 48 pp. (1977)
- Wittkowski, K.M., Lee, E., Nussbaum, M., Chamian, F.N., Krueger, J.G.: Combining several ordinal measures in clinical studies. *Stat. Med.* **23**(10), 1579–1592 (2004)

## Author Biographies

**Marco Diana** is a permanent researcher working at the Department of Hydraulics, Transport and Civil Infrastructures of Politecnico di Torino, Italy. His research interests include the study of innovative forms of public transport services and the analysis of multimodality behaviors.

**Patricia L. Mokhtarian** is a professor of Civil and Environmental Engineering, chair of the interdisciplinary MS/PhD program in Transportation Technology and Policy, and Associate Director for Education of the Institute of Transportation Studies, at the University of California, Davis. She has specialized in the study of travel behavior for more than 30 years.