Measuring the satisfaction of multimodal travelers for local transit services in different urban contexts

Original
Measuring the satisfaction of multimodal travelers for local transit services in different urban contexts / Diana, Marco. - In: TRANSPORTATION RESEARCH. PART A, POLICY AND PRACTICE. - ISSN 0965-8564. - STAMPA. - 46:1(2012), pp. 1-11. [10.1016/j.tra.2011.09.018]

Availability:
This version is available at: 11583/2486999 since:

Publisher:
Elsevier

Published
DOI:10.1016/j.tra.2011.09.018

Measuring the satisfaction of multimodal travelers for local transit services in different urban contexts

Marco Diana

5th August 2011

Preferred citation: this document may be cited directly referring to the above mentioned final published version:

Diana, M. (2012) Measuring the satisfaction of multimodal travelers for local transit services in different urban contexts, *Transportation Research Part A*, vol. 46(1), pp. 1-11.

MEASURING THE SATISFACTION OF MULTIMODAL TRAVELERS FOR LOCAL TRANSIT SERVICES IN DIFFERENT URBAN CONTEXTS

Marco Diana*

Dipartimento di Idraulica, Trasporti e Infrastrutture Civili

Politecnico di Torino

Corso Duca degli Abruzzi, 24

10129 Torino – ITALY

Phone: +39 011 564 5638 - Fax: +39 011 564 5699

marco.diana@polito.it

* *Corresponding author*

ABSTRACT

The importance of measuring customer satisfaction for a public transport service is apparent, even beyond more immediate marketing purposes. The present paper shows how satisfaction measures can be exploited to gain insights on the relationship between personal attitudes, transit use and urban context. We consider nine satisfaction measures of urban transit services, as expressed by a representative sample of Italian multimodal travelers (i.e. users of both private cars and public transport). We use correlations and correspondence analyses to show if and how each attribute is related to the levels of use of public transport, and how the relationship is affected by the urban context. Then we apply a recently developed method to combine ordinal variables into one score, by adapting it to work with large samples and with satisfaction measures which have a neutral point in the scale (i.e. "neither satisfied nor dissatisfied"). The resulting overall satisfaction levels and frequency of use were not correlated in our sample. We also found the highest satisfaction levels in smaller towns and the lowest ones in metropolitan cities. Since we focus on multimodal travelers, an interpretation paradigm is proposed according to which transit services must be well evaluated by car drivers in smaller towns in order to be considered a real alternative to cars. On the other hand, transit is more competitive on factual elements in larger cities, so that it can still be used by drivers, even if it is not very well evaluated.

KEYWORDS

Multimodality, customer satisfaction, public transport, correspondence analysis, ordinal measures

1. INTRODUCTION

The importance of measuring customer satisfaction for a public transport service is apparent, even beyond the more immediate marketing purposes that one might advocate. Customer satisfaction is in fact one of the key determinants of personal attitudes towards the service itself. In turn, there seems to be a clear link between attitudes and travel choices, particularly concerning short distance and urban trips, where choices are often less deliberate and the importance of factors such as habits and personal opinions is stronger than for long distance trips. Such attitudes-behavior relationship has been extensively studied in past decades also in the public transport domain, leading to different conclusions concerning its nature but little doubt on its relevance (for this debate, see for example Tardiff, 1977; Dobson et al., 1978; Kuppam et al., 1999; Golob, 2001; Parkany et al., 2004). The study of attitudes is therefore one of the keys to better understand the demand for a local transport service, beyond the influence of the characteristics and the performances of the service which is normally considered in transport models.

Once having understood the relevance of attitudes in this context, one needs a method to analyze them. To this effect, the concept of attitude is generally operationalized by (a) identifying a set of relevant attributes related to the service under investigation (for example, reliability, travel time or vehicle cleanliness) and (b) measuring the satisfaction degree of the customers related to those attributes and their relative importance in forming an overall judgment. These two topics have been the object of intensive research in the transportation sector (e.g. Weinstein, 2000; Foote, 2004; Stradling et al., 2007; Beiraõ and Cabral, 2009; Tyrinopoulos and Antoniou, 2008), with strong connections with works dealing with the quality of public transport services. In short, these works have contributed in identifying which are the most relevant aspects that influence customer satisfaction, and how these change for different user groups.

On the basis of the above research findings, transit satisfaction measures are therefore commonly included in marketing surveys, and standard methods are nowadays available to process them (Morpace International Inc and Cambridge Systematics Inc, 1999). The resulting information has a rather intuitive meaning and it can be easily summarized through descriptive statistics, thus giving immediate feedback and support to decision makers. A more advanced use of these data would imply embedding them in a modeling framework, for example in order to assess their relative importance compared to the above mentioned service characteristics in terms of travel costs and times. However, more elaborated quantitative analyses dealing with satisfaction measures are rather tricky to carry out for at least two reasons: (a) the fact that information related to satisfaction is sometimes gathered through semantic scales (like - … - dislike, very satisfied - … - not at all satisfied, etc.), leading to ordinal rather than metric variables, and (b) the fragmentation of the resulting information, since several attributes need to be considered to adequately investigate the general attitudes related to the study object. In particular, most of the above reviewed research seems to overlook the fact that a rating scale does not convey metric information, even if numbers are used to label the points of the scale in the questionnaire, so that specific analytical tools should be used. Even if it is rather common to treat such ordinal variables as ratio-scaled, this is not recommended on theoretical grounds and could be particularly problematic in our application, as we later discuss.

The primary objective of the present paper is to show a new method to overcome the above technical difficulty in better exploiting this kind of data. For this, we quantitatively analyze some satisfaction measures, which are customarily collected by public transport operators and decision makers during their monitoring activities, and are therefore widely available data, in order to gain better insights on the relationship between personal attitudes, transit use and other contextual factors, such as the urban environment that is considered here. This goal will be pursued by keeping into consideration the ordinal nature of the data, thus avoiding potentially misleading outcomes.

The analysis is carried out in two steps. At the beginning, we separately consider each satisfaction measure pertaining to a set of attributes describing “light” urban transit systems (excluding railways and metro) in Italy. We employ correspondence analysis, a multivariate data analysis technique for categorical variables, to show if and how each attribute is related to the levels of use of public transport, and how the relationship is affected by the urban context.

In the second part of the work, we apply a method to aggregate the different satisfaction measures expressed by each survey respondent into one score, thus summarizing the information and better clarifying the relationships between satisfaction levels, public transport use and urban context. This method was specifically conceived to combine ordinal measures (Wittkowski et al., 2004) and it has already been used in several research fields, from epidemiology to economics, including some recent studies in the transportation sector (Diana and Mokhtarian, 2009a; Diana and Mokhtarian, 2009b; Diana et al., 2009). One original contribution of the present study from a methodological point of view consists in proposing a variant in the computation of these scores to decrease the computational burden of the procedure. This is possible by taking advantage of the fact that ordinal scales measuring satisfaction are generally bipolar, i.e. they range from “positive” values, such as “very satisfied”, to “negative” ones, such as “not at all satisfied”, passing through a neutral point “neither satisfied nor dissatisfied”. Outside social sciences, most scales in life sciences and economics do not have this characteristic, since for example they measure the seriousness of clinical symptoms or describe performance rankings.

The application that we developed to demonstrate the above described approach focuses on multimodal travelers, i.e. people who use both cars and transit for their urban trips. In fact, on one hand exclusive car drivers have limited knowledge of public transport, so that it would be difficult to ask for their satisfaction related to a service they do not use. On the other, exclusive transit users are likely to be transit captives, i.e. their levels of use are likely to be not too sensitive to their satisfaction degree. Therefore, both from a marketing and from a policy viewpoint it is more sensible to focus on multimodal users, to try to understand to what extent their satisfaction levels affect their demand for transit under different circumstances. To the best of our knowledge, there is no previous example of work in this area focusing on this group of public transport users.

2. DATASET AND EXPERIMENTAL FRAMEWORK

The data that we use in the following come from the “Aspects of everyday life” survey, which is yearly administered by the Italian National Statistical Institute ISTAT to a stratified sample of about 50,000 inhabitants that is representative of the whole population. The purpose of this annual survey is to investigate habits and opinions of people on a variety of ambits, ranging from public services use to health conditions, quality of life or social inclusion. In the following we use the data from the 2007 wave.

The variables that we consider from this dataset are presented in table 1. Nine attributes are used in the survey to measure satisfaction with urban buses, trolleybuses and tramways, thus excluding “heavy” systems such as metros and railways. In the following we refer to “public transport” or “transit” for the sake of briefness, however the reader should keep in mind this exact definition of the study object. These nine attributes are shown in the upper half of the table and they pertain to key aspects that have an impact on the demand for public transport, as shown for example in those studies mentioned in the introduction. They overall form a reasonably complete list, even if some potentially influential aspects, such as safety and security concerns, were not investigated. Measures of these nine attributes are on a four-point bipolar scale, as shown in the third column of the table. Note that a neutral point “neither satisfied nor dissatisfied” is not among these, so that respondents had to express either a positive or a negative evaluation.

Along with these attributes, we consider the residential environment of the household of the respondent (variable URBAN). Six different categories are used here, with the following meaning. First of all, the Italian law individuates 14 so-called “metropolitan cities” in the whole country (one more has been added in May 2009, after that the survey took place), which correspond to the largest urban agglomerations. The territory of a metropolitan city is constituted by a core municipality at the center of the city, plus other surrounding municipalities. Taken together, core municipality and surrounding municipalities constitute the urban agglomeration. This subdivision corresponds to the first two categories for the variable URBAN. The remaining municipalities that do not belong to a “metropolitan city” were broken down in four classes according to the number of inhabitants, as shown in the third column of the table. However in the present research we exclude observations from households in municipalities below 10,000 (and not being part of a metropolitan city), since public transport services, if any, are likely to be too scarce in those cases. The aim is therefore to focus on those areas were a reasonable service is provided, which could constitute a viable alternative at least for some kind of trips. Note that URBAN is a purely nominal rather than an ordinal variable, because only a partial order can be defined among its categories: municipalities in the category “10-50k” have obviously less inhabitants than those in “>50k”, but municipalities belonging to a metropolitan city have widely different sizes.

We are also interested in checking whether the levels of use of public transport are related to customer satisfaction. For this, we consider the variable TRANS_USE, dropping those subjects who declared never using the service, or that the service is not existing. As we said in the introduction, we also drop those customers for which driving a car is presumably not a viable alternative. This is achieved in our dataset by eliminating those people who declared never driving a car, and also those living in a household without cars.

By applying all these filters to our dataset (kind of municipality, transit frequency of use, car driving frequency and car availability) we retained 4,123 observations from the 48,253 initially available, representing a subsample of about 5.9 out of 58.7 millions of individuals (i.e. the Italian population) when considering the related observation weights. The last column of the table reports the labels that we use in the following to refer to the various categories of the considered variables.

Table 1

3. ANALYSIS OF THE SET OF SATISFACTION MEASURES

3.1. Socioeconomic characterization of the group of respondents

Before starting the analysis on the relationships among satisfaction, levels of use of public transport and urban environment that are developed in the following two subsections, it is important to check whether the socioeconomic and demographic characterization of our sample of multimodal travelers differs according to their urban location. This analysis will be recalled in the following when interpreting the research results.

In general terms, multimodal travelers living in different urban environments show the same kind of differences than the general populations, sometimes attenuated. From figure 1, we see in fact that those living in "CENTRE" are more educated and tend to live alone or without children, whereas the percentage of workers is not radically different across groups. Household motorization is of course greater outside cities, but again the fact of having filtered those without cars makes the differences smaller. Concerning demographics, the mean age of multimodalists living in "CENTRE" and ">50k" is around 47, becoming 42 for "SUBURB" and 39 for "10-50k". Males are the majority in "CENTRE" (56%), "SUBURB" (52.5%) and ">50k" (50.2%), but not in "10-50k" (48.8%).

Figure 1

3.2. Correlations between satisfaction and levels of use, and between levels of use and urban context

The variables that are presented in table 1 are either ordinal (the nine satisfaction measures and TRANS_USE) or nominal (URBAN). Concerning transit satisfaction and frequency of use, it is therefore possible to check for the existence of a relationship by computing Spearman rank-order correlations between each satisfaction measure and TRANS_USE. Three of the resulting nine correlations were not significant at the 5% level (namely, those involving FREQUENT, SPEED and COMFORT), whereas PUNCTUAL, SEAT and CLEAN have a significantly negative correlation with TRANS_USE and CONNECT, CONVEN and COST a positive one. These results seem rather problematic, however it is likely that such large sample size makes significant even rather weak correlations, beyond their practical meaning. In fact, repeating the computation of these correlations on different subsets of this sample made most of them not significant.

We can therefore conclude that no apparent relationship has emerged between satisfaction and levels of use of urban transit for multimodal travelers. We further comment this result in the concluding section, but it is important here to note that the analysis of single satisfaction measures did not give a clear-cutting response, since some significant correlations was actually found in such a large sample, that on the other hand cannot be clearly interpreted. This calls for a method which allows analyzing the overall satisfaction of the customers, in order to have more easily interpretable results. We define and apply such method in the following section 4. By considering a unique satisfaction score, it will also be more straightforward to check if those not so much clear-cutting results are due to the influence of confounding variables, for example URBAN or demographic factors. For this, we analyze at the end of the following section 4 how satisfaction scores vary according to the urban environment and to the above considered socioeconomic variables.

Another preliminary analysis investigated the correlation between URBAN and TRANS_USE. Even if TRANS_USE is ordinal, it is possible in this case to assign approximate metric values to the reported frequencies, to have a rough estimation of the mean frequency of use in our sample. Looking at the exact wording of the survey question, we judgmentally assigned the following values to the four categories as an estimate of the number of times that the interviewees use public transport in a year: DAILY = 350, WEEKLY = 100, MONTHLY = 25 and YEARLY = 5.

Not surprisingly, the mean frequency of use of transit for multimodal travelers is the highest one for those living in municipalities at the center of a metropolitan city (about 102 times a year), followed by the dwellers of municipalities above 50,000 (almost 66 times a year), whereas those living in municipalities having from 10,000 to 50,000 inhabitants or in the suburbs of metropolitan cities use less frequently those means (around 64 times a year in both cases). Concerning these latter two values, a non-parametric Kruskal-Wallis test confirmed that the hypothesis of equal frequencies for the two groups cannot be rejected ($p = .73$), whereas these frequencies are significantly different from the “>50k” category ($p < .01$). On a methodological point of view, we note that ordinary t-tests and ANOVA analyses cannot be used in this case, since the involved distributions are severely non-normal.

To sum up, Italian multimodal travelers living in suburbs of metropolitan cities and those living in medium-sized municipalities seem to use public transport with the same frequency, and also their socioeconomic profile is very similar according to the preceding subsection. One would probably expect that transit services in suburbs are in any case “better”, or at least “more appealing” than those in towns below 50,000, at least for trips to/from the centre of the metropolitan city, which should be a non-negligible proportion of the total. More research work is needed to better interpret this finding. One possible explanation is that suburban dwellers have an easy access to a road network serving the metropolis that is better than the one in smaller agglomerations. This could offset the fact of having a better transit service when considering such multimodal travelers.

3.3. Correspondence analysis of satisfaction versus urban context

The variable URBAN is merely nominal rather than ordinal, so that correlations cannot be computed in this case. Therefore in the following we resort on a multivariate analysis statistical technique specifically tailored for categorical variables, namely correspondence analysis. Correspondence analysis is a technique which represents the associations among categorical variables in a single plot or map. All the categories are displayed in it through points which reflect the cell frequencies of the related cross tabulations. This technique is well documented and available through most statistical software packages. The interested reader is referred to Greenacre (2007) for one of the latest available handbooks. Studies which use such technique in the transport research field usually deal with road safety and drivers behaviors, two among the most recent ones being Kim and Yamashita (2008) and Nallet et al. (2008), with other applications in freight industry attitudes (Hensher and Golob, 1999), trip chaining (Golob and Hensher, 2007) and air market segmentation (Wen et al., 2008).

We start by examining the nine 4*4 cross tabulations of each satisfaction measure with URBAN. Related chi-square statistics are highly significant in all cases, suggesting that satisfaction for urban transit and urban context are not independent. Therefore we decided to have a closer look at the data and we computed nine 2*4 cross tabulations by aggregating the two categories related to positive responses (HIGH and MED) and the two related to negative ones (LOW and NO). The results are synthetically presented in table 2, where the percentages of respondents very satisfied or rather satisfied are reported. For all the nine considered attributes, the greatest fraction of "positive" responses was reported for category "10-50k", followed by the ">50k" one. This could be surprising at first sight, since for example the satisfaction expressed with FREQUENT seems inversely correlated with the actual frequency of service, which should be higher for metropolitan cities and lower for smaller agglomerations. However it should be kept in mind that the sample that we consider is not representative of the whole population, and not even of the transit users. This is because we focus on multimodal travelers, who are likely to be a small subset of all the transit users, particularly in smaller towns.

Table 2

It is therefore likely that only those who are very satisfied by the service keep on using it in smaller towns even if they also drive cars, whereas in larger agglomerations also those car drivers less satisfied by public transport use it, because in some circumstances it could be much more convenient for instrumental reasons (congestion, parking problems etc.). This consideration suggests us a different way of looking at the above results. Multimodal travelers are in fact not representative of all travelers, so that the percentages in table 2 do not represent the mean satisfaction level for the service. They can rather be considered as the satisfaction level that needs to be reached in different urban contexts by a transit service to be competitive with cars. In the following we will consistently adopt this interpretation paradigm to comment our findings.

Considering columns 2 and 3 of table 2, two different patterns of interdependence between satisfaction measures and URBAN can be detected. On one hand, the satisfaction for

FREQUENT, COMFORT, CONNECT, CONVEN and COST is higher for those living in the center of a metropolitan city than for those living in the suburbs. The contrary is observed for the remaining four variables PUNCTUAL, SEAT, SPEED and CLEAN. The interpretation of most of these results is rather intuitive, since we expect that services in CENTRE are more frequent and with more convenient schedule than in suburbs, provide better comfort while waiting at bus stops and are cheaper because of transit fare schemes. On the contrary, suburban services are likely to be less crowded and less hindered by congestion. Comparing table 2 with the socioeconomic differences among groups presented in subsection 3.1 seems also to reinforce the above interpretation paradigm. For example, the higher proportion of retired persons among multimodalists living in CENTRE is reflected by higher satisfaction levels for COMFORT, whereas the more active persons in SUBURB are related to higher satisfaction levels for PUNCTUAL. It seems hence reasonable to distinguish these two sets of variables in subsequent analyses by defining two groups “A” and “B”, where the former group contains those attributes that are better rated in metropolitan city centers and the latter one those that are better rated in suburbs. Therefore, the labels of the response categories of the above listed five variables will be indicated in the correspondence analysis plot with an “A” subscript, whereas the other four will have a “B” subscript.

Following these naming conventions, it is finally possible to present the plot from the correspondence analysis of the nine satisfaction measures against the urban context in figure 2, which was obtained through the SAS software by entering the corresponding 36*4 contingency table. The figure clearly shows that higher satisfaction levels of multimodal travelers for public transport are more closely associated with smaller municipalities, and moderate satisfaction with larger ones. Concerning the lower part of the plot, the total absence of satisfaction is clearly associated with suburbs of metropolitan cities for the above defined group A of attributes, whereas for group B it is more in between suburbs and centers. In a sort of symmetric design, low satisfaction levels are clearly associated with metro centers for group B and are in between suburbs and centers for group A.

Figure 2

Our correspondence analysis map prompts for a number of considerations, on the basis of the above defined interpretation paradigm that can help us discerning what are the most relevant aspects of the service. Services in SUBURB are not expected to be always satisfactory in terms of frequency, scheduling convenience or comfort at bus stops in order to be used. Lower satisfaction levels on these specific aspects are probably also due to a comparative assessment of the nearby CENTRE services, since we recall that households in those two groups both live in metropolitan cities. It is also likely that some trips from metropolitan cities suburbs to centers are so inconvenient with cars that transit can still be used, even if it is totally disappointing on attributes belonging to group A. On the other hand, if transit is not at all satisfactory for attributes of group B, this could have a greater impact for multimodal travelers living in SUBURB. Instead, for those living in CENTRE, the service must be not totally disappointing in order to be used, but satisfaction levels are generally low. In this case, cars are still relatively weak competitors of transit but other modes like walking or cycling, which are not so easy to use when living in SUBURB, could probably be considered if public transport is too badly evaluated.

Outside metropolitan cities, transit services in larger municipalities need to be considered at least moderately satisfactory on most of the attributes, in order to be used by car drivers. Finally, in smaller towns high satisfaction levels must be reached in all attributes of group A, and moderate satisfaction for many aspects in group B, in order to attract car drivers.

4. COMBINING SATISFACTION MEASURES INTO ONE SCORE

4.1. A new compounding method for bipolar scales

The preceding section has shown how is it possible to deeply investigate the relationship between satisfaction ratings of several attributes pertaining to public transport services and different factors, such as the urban context, through correspondence analysis. Yet it would be desirable to implement an aggregation method for these different satisfaction measures, to have a more synthetic representation of the *overall* satisfaction level for the service. Beyond this, correspondence analysis is a technique tailored for categorical data, since it does not exploit the information related to the partial order that can be defined among the attribute ratings.

Since the variables representing satisfaction levels are not metric, using arithmetic operators (sums, means etc.) to combine such respondents' ratings is not the best solution, even if this is often done for simplicity in widely used methods, such as quadrant analysis or impact scores (Morpace International Inc and Cambridge Systematics Inc, 1999). Numerical values assigned to each point in the scale can in fact be arbitrary, such as in our case where adjectives like “very” or “rather” were used. Even if numbers are explicitly recalled in the questionnaire (e.g. asking “how much would you rate from x to y…”), the respondents tend not to consider the points of the scale as equally spaced, which is a precondition to perform arithmetic operations. Within a modeling framework, ordered logit and ordered probit regression could be used in this case to find actual distances between contiguous points (see for example Redmond and Mokhtarian, 2001). One should also consider that asking for several numeric ratings increases the cognitive burden of the respondent, thus inflating measurement errors.

The use of parametric statistical methods such as analysis of variance with ordinal data is well attested in the transport sector as in other applicative field, even if for example ANOVA is based on differences between means, which are affected by monotonic transformations of the data. However, aggregating scores to build a synthetic indicator is probably less robust to measurement assumptions violations than driving a multivariate statistical analysis. To overcome these problems, in the following we use a specialized method (Wittkowski et al., 2004), that makes it possible to combine a set of ordinal measures by defining a partial order among the observations. Measures need not to be expressed through numbers; any kind of label of the points of the scale is acceptable, including the adjectives of quantity of our satisfaction measures.

Applying this method in our case, we can say that respondent A is overall more satisfied by public transport than respondent B if A has expressed a greater satisfaction degree than B for at least one public transport attribute and a smaller degree in none. All the pairs of subjects in our sample are then compared in this way to assign a score to each respondent A, given by the number of respondents who were overall less satisfied than A, minus the number of respondents who were more satisfied than A. Such scores are called μ-scores and range from $-(n-1)$ to $+(n-1)$ if we have n respondents in our dataset. However some of these pairs of observations cannot be compared, whenever one respondent has given higher scores to at least one attribute and lower scores to another. Therefore the above range is reduced, particularly when we have many measures to combine and these are highly correlated. We note in passing that one could think about performing an aggregation of satisfaction measures through a parametric method such as factor analysis. However, previous studies have shown that this could lead to results that are not so easy to interpret when dealing with ordinal variables from semantic scales (Diana et al., 2009).

One practical limitation of this method is that it can be rather burdensome for large samples. It is in fact easy to see from the above description that the needed computational resources, in terms of memory and CPU time, increase linearly with the number of variables to be combined, but moreover with the squared number of pairs of observations, equaling to (0.5*n*(n-1)), where n is the number of respondents. When using the open source *muStat* package of R to compute μ-scores (http://cran.r-project.org) as we did in the following application, an ordinary personal computer might no more be sufficient when the sample size is more than 2,000 with about 10 measures per observation. Datasets of similar dimensions are frequently available from larger transport surveys.

Another limitation is given by the fact that this method was conceived for ordinal measures which do not have a neutral point, whereas as we said in the introduction we are dealing here with bipolar measures. Unipolar and bipolar measures do not differ on a statistical viewpoint, and therefore μ-scores can legitimately be computed in both cases. However, the above defined partial order might become questionable in the latter case when coming to the interpretation of the results. For example, suppose that three respondents rate three of the attributes on the satisfaction scale of our survey, respectively giving the responses (HIGH, HIGH, MED), (NO, LOW, MED) and (NO, LOW, HIGH). One would say that the first respondent overall seems more satisfied than the third one, since s/he gave all positive responses, against only one positive response and two negative ones given by the last respondent. However the resulting μ-scores are 1, -2 and 1, the scores of the first and third respondent being thus the same. Note that the interpretation of this result would instead be meaningful in a scale without neutral points, since the third respondent was the only one to give the highest score on the third attribute.

We propose in the following a modification in the computation of μ-scores in order to address both the above concerns. Concerning the peculiarity of bipolar scales, the basic idea is to constrain the μ-score of respondent A to be greater than that of respondent B, if A gave more responses falling in the positive side of the scale (i.e. counting both “HIGH” and “MED” responses in our case) than B. This can be done by stratifying the observations by number of positive responses, computing μ-scores within each group and shifting them according to their group. The resulting aggregation measure, named β-score in the following, would hopefully be more fit for bipolar scales. β-scores will therefore be computed through the following steps:

1) cluster the respondents by number of positive responses given, thus forming k+1 groups, if k is the number of measures per observation (in our case, we would have 10 groups ranging from k = 0 to k = 9);

2) sort the groups by ascending number of positive responses with index j = 0, …, k;

3) for every group j, compute μ-scores by considering only the respondents who are within this group;

4) compute the range Δ_j of μ-scores within each group j, given by the difference between the maximum (M_j) and the minimum (m_j) μ-score of the group;

5) the β-score of a generic respondent i that belongs to group j can then be computed as follows:

$$\beta_i = \sum_{l=0}^{j}(\Delta_l + 1) - M_j + \mu_i - 1 \ .$$

We note in step 3 that μ-scores are now computed only within several smaller subsets, thus dramatically reducing the needed computer resources.

This method is not a panacea; in particular, some groups could contain very few observations when the sample size is small. In turn, the computation of μ-scores for groups with few observations but several ordinal measures to combine might be difficult. However β-scores as above defined seem appropriate in our case, where the sample size is adequate. In the following we then present some results on the overall relationship between satisfaction degree for public transit, urban context and frequency of use, by employing both μ- and β-scores to compare their behaviors.

4.2. Application of the method and results

According to the above discussion, our sample size of 4123 is too big to compute μ-scores with an ordinary desktop computer, as we would like to do to compare them with β-scores. Therefore we randomly selected a subset of 1620 cases that we considered for further analysis. The computation of β-scores involves in our case the definition of ten groups, since we considered nine satisfaction measures (see point 1 of the above methodology). The number of observations within each group ranges between 116 and 211, so that the computation of within-group μ-scores is perfectly feasible according to the above point 3.

We are preliminary interested in understanding if β-scores are a viable alternative to μ-scores. One possible inconvenient of β-scores is that their distribution has some peaks, due to the fact that they are computed from within-group μ-scores. In fact, the pattern of responses concerning satisfaction measures inside each group is necessarily less variegated, also considering that the scales that were used have only four points and that the correlation among our measures is generally high. μ-scores are instead more dispersed when computed across the whole sample. Both μ- and β-scores distributions severely depart from normality in our sample, and on the other hand the two measures were highly correlated (r = .966). Moreover, the statistical results that are reported in the following did not change considering either μ- or β-scores. To sum up, β-scores were proven to be a good alternative in our case, and their computation took few seconds, against about 8 minutes for obtaining μ-scores for all the 1620 cases on a computer with a 2.4 MHz CPU and 512 Mb of RAM.

Coming back to the research question concerning the relationship between transit frequency of use and satisfaction degree for the service, we noticed in subsection 3.2 that some correlations were occasionally significant. Formally testing the hypothesis of equal means of these satisfaction scores across different levels of use of transit always confirmed that this hypothesis cannot be rejected (Kruskal-Wallis $p > .05$). This gives clear statistical support concerning the independence of satisfaction and frequency of use of urban transit, against the more ambiguous indications that could be inferred by separately considering each satisfaction attribute.

Also the study of the relationship between transit services satisfaction levels and urban context is much more straightforward when using our compounded satisfaction scores rather than single measures. The pattern of such relationship, when considering either μ- or β-scores, reproduces what was observed in subsection 3.3 for measures belonging to the above defined group A. Therefore, highest satisfaction levels are recorded for multimodal travelers of smaller towns, followed by those living in larger ones and in centers of metropolitan cities. The least satisfied are those living in suburbs of those cities. However, the difference in global satisfaction level between these latter two groups is not significant (Kruskal-Wallis $p > .45$). This result is quite interesting, and can probably be at least partly explained by considering that many metropolitan cities have a unique transit operator that serves both the city center and the surrounding municipalities, so that judgments are rather blurred. We had already noticed similar interactions between the CENTRE and the SUBURB categories when commenting the results for single satisfaction measures. Again, we note that the use of an aggregated measure allows the researcher to more clearly understand how satisfaction for transit changes according to the urban context.

Another related interesting application of β-scores is to assess how many multimodal travelers are overall very satisfied, rather satisfied, little satisfied and not satisfied by urban public transport services. Travelers overall very satisfied could logically be defined as those respondents whose β-score is within the β-score of a respondent who rated HIGH all satisfaction measures and the β-score of a respondent who rated MED all the measures, and the same goes for the other groups. However if the bipolar scale of the questionnaire does not include the neutral point, as in our case, we need an additional threshold value T to discriminate those satisfied from those unsatisfied. One could look at β-scores of people who gave positive and negative evaluations on the same number of items, but such scores are not unique (and in any case the number k of satisfaction measures would be odd in our dataset). This is not a flaw of the method, since the considered attributes do not have the same importance in determining the overall satisfaction for the service. Recalling the previously introduced notation, when k is odd the most straightforward solution could be to compute the needed threshold T as the midpoint between (1) the highest β-score among all the respondents who gave more positive than negative responses and (2) the lowest β-score among all those who gave more negative than positive responses:

$$T = \sum_{l=0}^{(k+1)/2} \left(\Delta_l + 1 \right) - 0.5 \ .$$

Therefore, it has been possible to classify each observation in one of the four levels of global satisfaction. The results of this computation, keeping into account the different observation weights of our stratified sample, are reported in figure 3. The values here shown further substantiate the discussion on the relationship between satisfaction level and urban context developed in the preceding paragraph. Nearly 70% of multimodal travelers living in smaller municipalities outside metropolitan areas can be considered to be overall satisfied by public transport, whereas this figure falls to about 60% for larger municipalities and to little more than 40% in metropolitan areas. In particular, it is insightful to note that category “10-50k” has the highest proportion of very satisfied people and the lowest proportion of not at all satisfied ones. This reinforces the previous finding related to the higher satisfaction degree of multimodal travelers belonging to this group.

Figure 3

It is finally interesting to check how this synthetic satisfaction score is correlated with some key socioeconomic and demographic variables, much like the group analysis by urban context that was presented in subsection 3.1. The results are shown in figure 4. It is worth noting for example the higher proportion of satisfied persons among retired and lower proportion among students, a difference that is probably due to their respective mobility patterns (off-peak trips without time pressure for retired, the contrary for students). Such discrimination is somewhat blurred for workers, who are more or less equally present in all the four satisfaction groups, probably due to the fact that they have higher car availability and less economic constraints compared to students, i.e. they can more readily switch to another option if they are not satisfied with transit. However we recall that in absolute terms transit captives are not included in our sample, according to the objectives of the present research and to the subsequent selection of the group of individuals under investigation. In fact, we see little influence of car ownership on satisfaction levels for transit, a result that would probably not be confirmed when considering the general population. Concerning demographics, there seems to be a positive correlation between satisfaction score and age (the mean age of those unsatisfied is 43.9 and it raises up to 50.0 for those highly satisfied) whereas males are the majority in HIGH, MED e NO groups (58.0%, 53.2% and 60.2% respectively) but not in LOW (49.2%)

Figure 4

5. CONCLUSIONS

In the present paper we analyzed the satisfaction degree of a group of multimodal travelers concerning urban public transport services in the area where they live. On a methodological point of view, we used correspondence analysis and then we defined and applied a new method to combine several ordinal measures which improves the state of the art when measures are expressed through bipolar scales and large datasets need to be examined. This allowed us to sum up the information from the single measures of satisfaction pertaining to different attributes of the system and to draw some more coherent and insightful conclusions. The computational method that has been showcased in this paper was much more efficient compared to previous ones, while being better tailored for bipolar semantic scales. This new method could be of interest also for researchers outside the transport sector.

On more applicative grounds, focusing on an ad-hoc choice-based sample (multimodal persons), rather than to a representative sample of the whole population, lead us to a different interpretation paradigm of the quantitative results. According to this latter, the main findings of the present study that are of relevance for transport researchers and practitioners can be summed up as follows:

- The overall satisfaction levels for the service and its frequency of use seem not correlated for multimodal travelers.
- Frequency of use is related to the urban context. Beyond the well-known figures related to the general population, transit use also for multimodal travelers is highest at the center of metropolitan cities, lower in towns above 50,000 inhabitants and the lowest one in smaller towns and in the suburbs of metropolitan cities. The frequency of use in those latter two cases is practically the same.
- Satisfaction levels are highest in smaller towns and lowest in metropolitan cities, without significant distinction between city center and suburbs. Given the fact that our sample of multimodal travelers is neither representative of the general population nor of transit users, this means that transit services must be evaluated quite satisfactory by car drivers, at least in some aspects and especially in smaller town, in order to be really attractive.

The finding that satisfaction and levels of use are not correlated in our sample of multimodal travelers seem to contrast with other studies, in which such correlation was observed when considering the general population. However we believe that such contrast is only apparent. Beyond the differences between multimodal travelers and general population, we mention here that according to Diana (2010), a transport service must be competitive on more instrumental factors, such as cost and travel times, in order to attract customers who already know it, as the respondents in our sample. On the other hand, that research pointed out that people who are less familiar with a mode (such as the general public, for which the levels of use of transit are rather low) can be more influenced in their choices by more subjective evaluations, such as the satisfaction degree related to the service, that was studied here.

It must finally be acknowledged that more factors should have been considered to draw more conclusive results. In particular, no information was available in the dataset concerning specific mobility patterns when using public transport, beyond trip frequencies. Trip travel times and purposes could actually have an influence on both transit satisfaction and levels of use. The socioeconomic characterization of the sample that has been carried out could only indirectly investigate such aspect, for example by looking at the differences between active and non-active persons that are likely to have different mobility patterns when using public transport.

One interesting extension of the present work would be to perform this analysis to representative samples of public transport users and of the general population and check how our results would be affected. Dropping μ-scores, which were still considered here for a comparative assessment, and exclusively using β-scores would in fact allow the researcher to easily work with much larger datasets. By using such methodology, it could be possible for transit managers, policy makers and stakeholders to gain better insights on public transport services by better exploiting those satisfaction data that are often used only to compile descriptive statistics, despite the financial effort that is generally necessary to gather them.

REFERENCES

Beirão, G. and Cabral, J.S. (2009) Measuring dissatisfaction with public transport service. Presented at the *88th Transportation Research Board Annual Meeting*, Washington, D.C., 17pp.

Diana, M. (2010) From mode choice to modal diversion: a new behavioural paradigm and an application to the study of the demand for innovative transport services. *Technological Forecasting and Social Change*, 77(3), pp. 429-441.

Diana, M., Song, T. and Wittkowski, K.M. (2009) Studying travel-related individual assessments and desires by combining hierarchically structured ordinal variables. *Transportation*, 36(2), pp. 187-206.

Diana, M. and Mokhtarian, P.L. (2009) Desire to change one's multimodality and its relationship to the use of different transport means. *Transportation Research F*, 12(2), pp. 107-119.

Diana, M. and Mokhtarian, P.L. (2009) Grouping travelers on the basis of their different car and transit levels of use. *Transportation*, 36(4), pp. 455-467.

Dobson, R., Dunbar, F., Smith, C.J., Reibstein, D. and Lovelock, C. (1978) Structural models for the analysis of traveler attitude-behavior relationships. *Transportation*, 7(4), pp. 351-363.

Foote, P.J. (2004) Making buses better in Chicago - Strategic implementation of customer-derived performance measures from 1995 to 2001. *Transportation Research Record*, No. 1884, pp. 18-26.

Golob, T.F. (2001) Joint models of attitudes and behavior in evaluation of the San Diego I-15 congestion pricing project. *Transportation Research A*, 35(6), pp. 495-514.

Golob, T.F. and Hensher, D.A. (2007) The trip chaining activity of Sydney residents: A cross-section assessment by age group with a focus on seniors. *Journal of Transport Geography*, 15(4), pp. 298-312.

Greenacre, M.J. (2007) *Correspondence Analysis in Practice – Second Edition*. Chapman & Hall, Boca Baton.

Hensher, D.A. and Golob, T.F. (1999) Searching for policy priorities in the formulation of a freight transport strategy: A canonical correlation analysis of freight industry attitudes. *Transportation Research E*, 35(4), pp. 241-263.

Kim, K. and Yamashita, E. (2008) Corresponding characteristics and circumstances of collision-involved pedestrians in Hawaii. *Transportation Research Record*, No. 2073, pp. 18-24.

Kuppam, A., Pendyala, R. and Rahman, S. (1999) Analysis of the role of traveler attitudes and perceptions in explaining mode-choice behavior. *Transportation Research Record*, No. 1264, pp. 68-76.

Morpace International Inc and Cambridge Systematics Inc (1999) *A Handbook for Measuring Customer Satisfaction and Service Quality*. Transit Cooperative Research Program Report 47. Transportation Research Board, National Academy Press, Washington D.C..

Nallet, N., Bernard, M. and Chiron, M. (2008) Individuals taking a French driving licence points recovery course: Their attitudes towards violations. *Accident Analysis & Prevention*, 40(6), pp. 1836-1843.

Parkany, E., Gallagher, R. and Viveiros, P. (2004) Are attitudes important in travel choice? *Transportation Research Record*, No. 1894, pp. 127-139.

Redmond, L.S. and Mokhtarian, P.L. (2001) The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation*, 28(2), pp. 179-205.

Stradling, S.G., Anable, J. and Carreno, M. (2007) Performance, importance and user disgruntlement: A six-step method for measuring satisfaction with travel modes. *Transportation Research A*, 41(1), pp. 98-106.

Tardiff, T.J. (1977) Causal inferences involving transportation attitudes and behavior. *Transportation Research*, 11(6), pp. 397-404.

Tyrinopoulos, Y. and Antoniou, C. (2008) Public transit user satisfaction: Variability and policy implications. *Transport Policy*, 15(4), pp. 260-272.

Vredin Johansson, M., Heldt, T., and Johansson, P. (2006) The effects of attitudes and personality traits on mode choice. *Transportation Research A*, 40(6), pp. 507-525.

Weinstein, A. (2000) Customer satisfaction among transit riders - How customers rank the relative importance of various service attributes. *Transportation Research Record*, No. 1735, pp. 123-132.

Wen, C.-H., Lai, S.-C. and Yeh, W.-Y. (2008) Segmentation and positioning analysis for international air travel market: Taipei-to-Tokyo route. *Transportation Research Record*, No. 2052, pp. 46-53.

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.

Table 1. List of considered variables

Variable	Question	Response categories*	Label*
Degree of satisfaction with the following aspects related to buses, trolleybuses and trams within the municipality:			
FREQUENT	Service frequency		
PUNCTUAL	Punctuality		
SEAT	Possibility of finding sitting places	Very satisfied	HIGH
SPEED	Speed of the service	Rather satisfied	MED
CLEAN	Cleanliness of the vehicles	Little satisfied	LOW
COMFORT	Comfort while waiting at bus stops	Not at all satisfied	NO
CONNECT	Connectivity with other municipalities		
CONVEN	Convenience of the schedules		
COST	Cost of the ticket		
URBAN	Municipality where the household is located	Metropolitan city center	CENTRE
		Metropolitan city suburb	SUBURB
		Other municipality above 50,000	>50k
		Other municipality from 10,000 to 50,000	10-50k
		(Other municipality from 2,000 to 10,000)	-
		(Other municipality below 2,000)	-
TRANS_USE	Frequency of use of urban public transport	Everyday	DAILY
		Sometimes a week	WEEKLY
		Sometimes a month	MONTHLY
		Sometimes a year	YEARLY
		(Never)	-
		(Service not existing)	-

* Observations pertaining to categories reported in brackets were not considered in this work, the corresponding label was therefore not defined

Table 2. Percentage of satisfied multimodal travelers according to the urban context

Attribute / Urban context	CENTRE	SUBURB	>50k	10-50k
FREQUENT	48.14%	45.43%	65.89%	70.57%
PUNCTUAL	41.43%	49.14%	65.27%	71.35%
SEAT	33.00%	46.42%	59.06%	72.14%
SPEED	51.20%	57.78%	69.54%	79.63%
CLEAN	33.40%	38.27%	57.41%	60.32%
COMFORT	34.40%	29.63%	44.31%	48.36%
CONNECT	56.57%	42.72%	59.96%	63.60%
CONVEN	52.06%	40.49%	60.23%	62.68%
COST	38.05%	32.10%	41.63%	46.65%
Number of responses	*1506*	*405*	*1451*	*761*

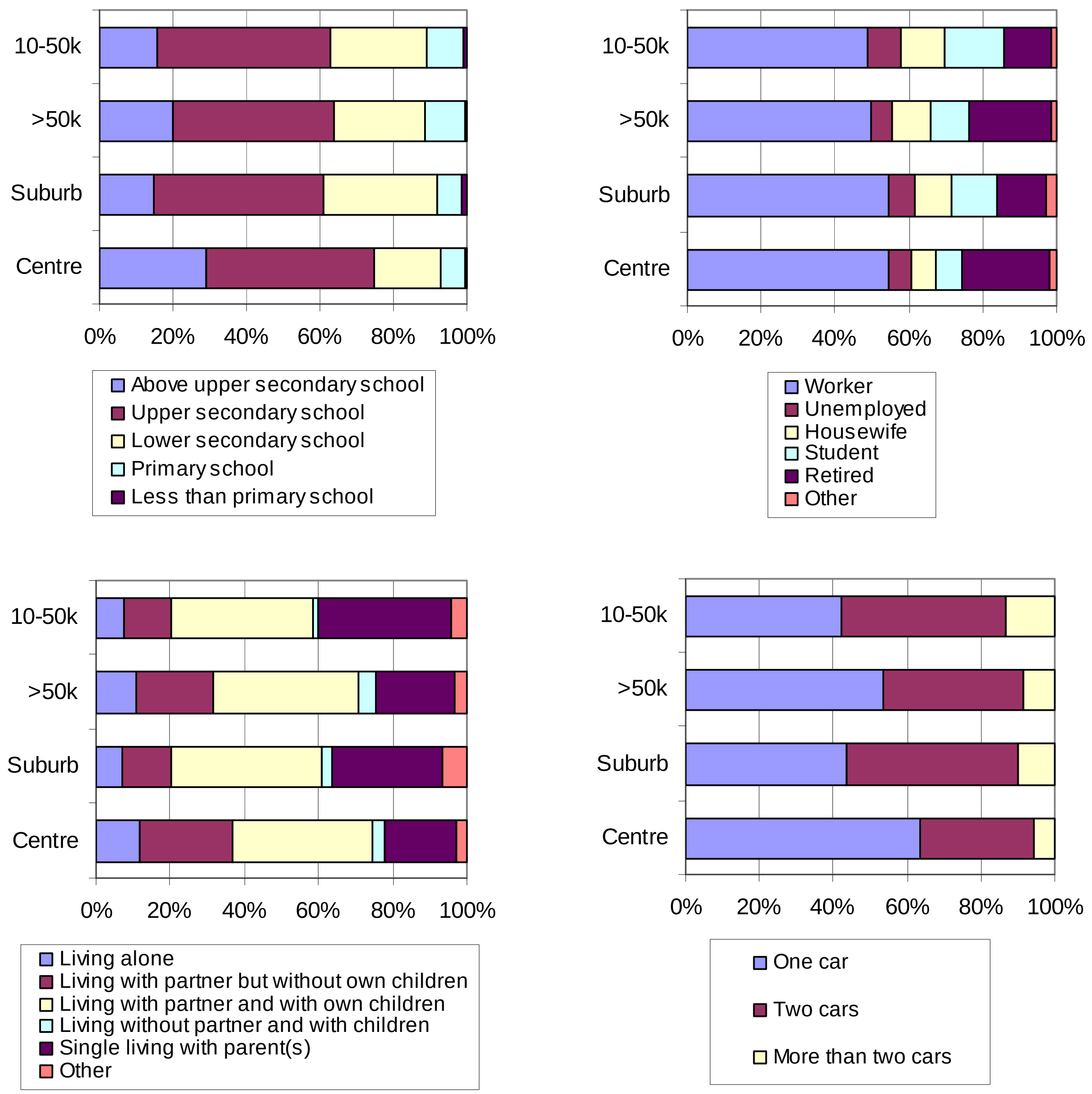

Figure 1. Educational levels, employment status, position in the household and car ownership of multimodal travellers by urban context

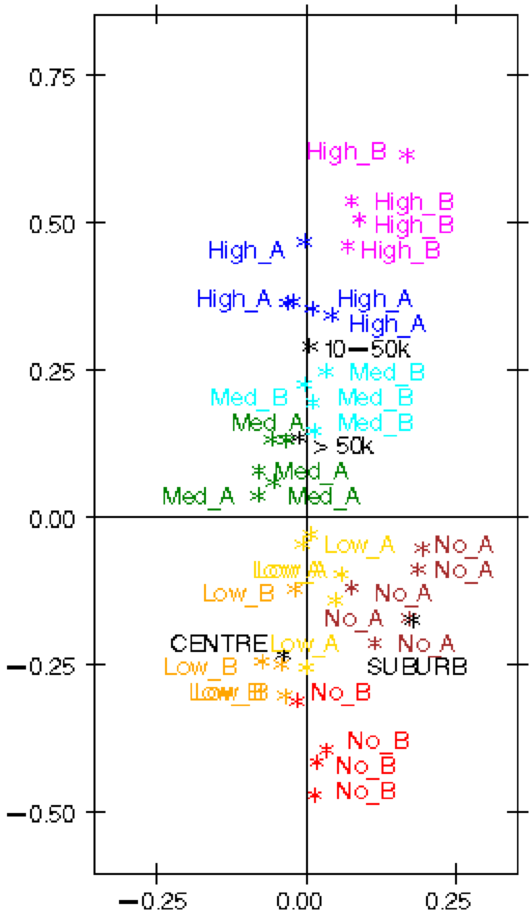

Figure 2. Perceptual map of satisfaction levels with urban transit versus urban context

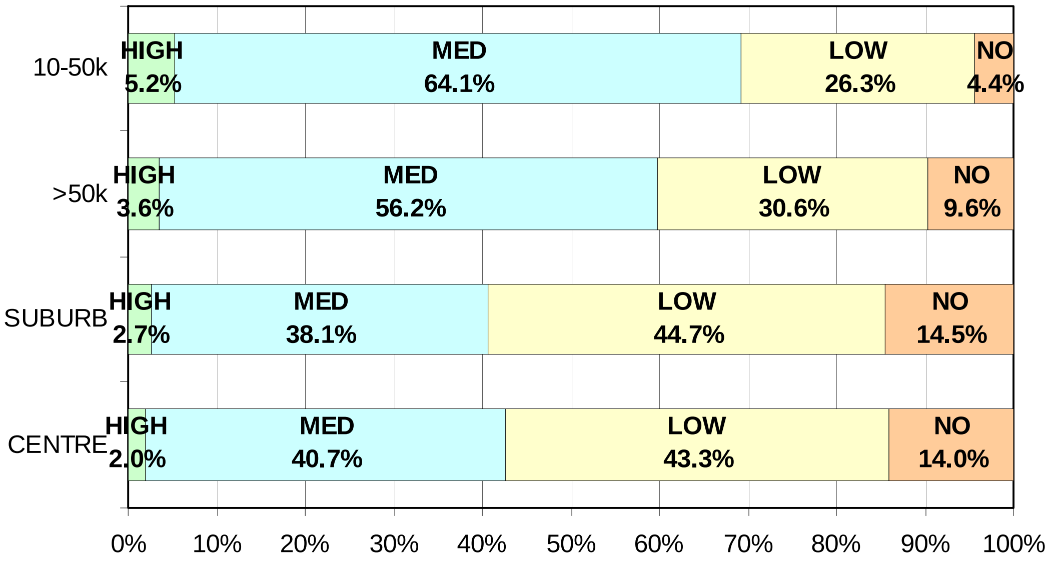

Figure 3. Overall satisfaction level with urban transit of multimodal travellers by urban context

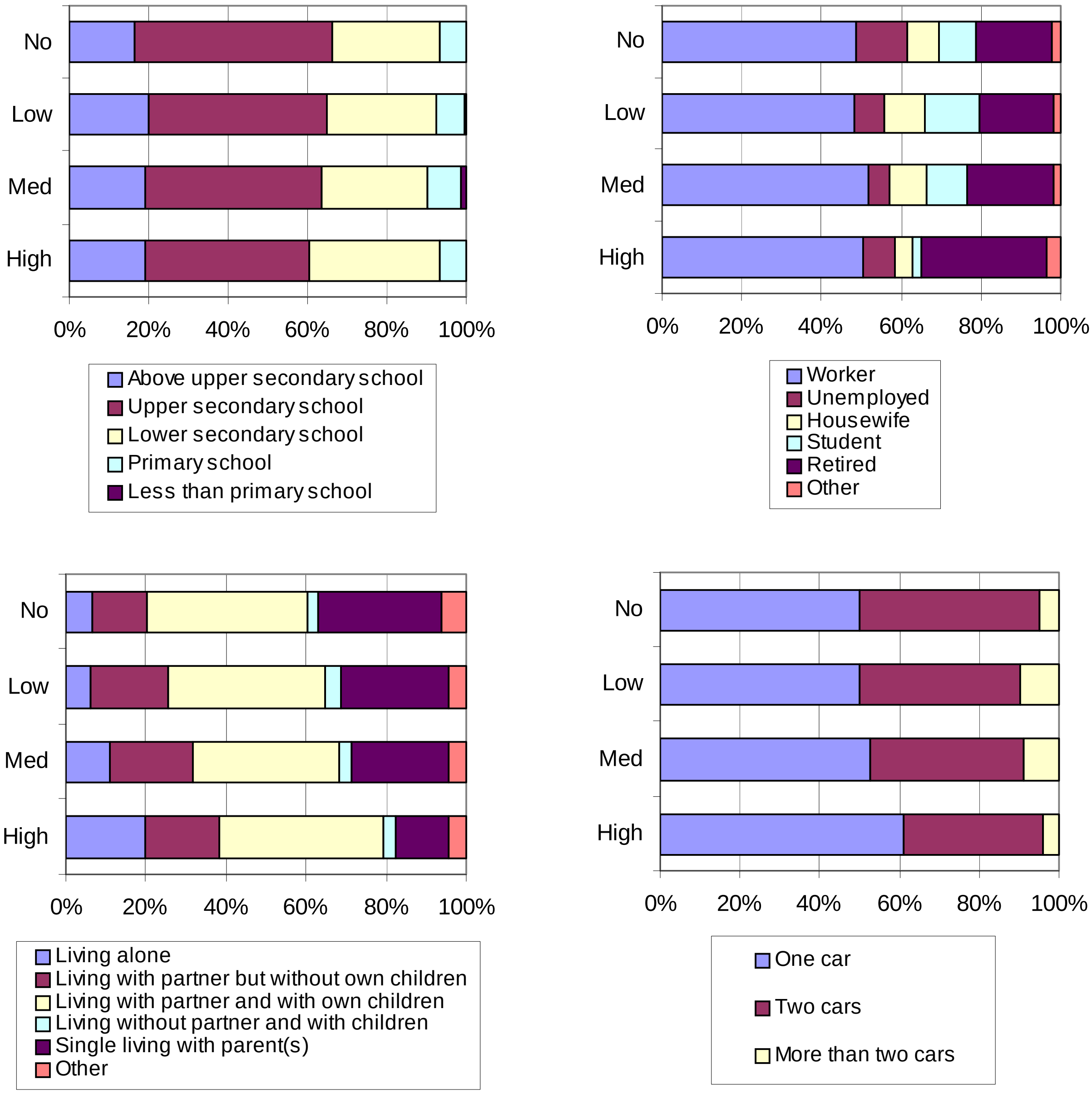

Figure 4. Educational levels, employment status, position in the household and car ownership of multimodal travellers by overall satisfaction level with urban transit