
A Composite Index for Evaluating Transit Service Quality across Different User Profiles

*Juan De Oña, Rocio De Oña, and Francisco Diez-Mesa
University of Granada, Spain*

*Laura Eboli and Gabriella Mazzulla
University of Calabria, Italy*

Abstract

This paper evaluates the quality of service of the metropolitan Metro of Seville (Spain) across different user profiles, as determined through cluster analysis. Service quality evaluation is performed using a composite index that combines the user point of view with the service operator point of view. The combination of these two types of service quality measurement fulfils the need to provide a reliable measurement tool for transit performance. Six user profiles were identified, and it was ascertained that they have different opinions about the service, with heterogeneous gaps between the points of view among the six user profiles.

Key words: *Transit quality; cluster analysis; user profiles; composite index*

Introduction

One way for public transport (PT) to achieve more competitiveness with private vehicles is to improve the service quality (SQ) (De Oña and De Oña 2015; Wen and Lai 2010, Dell’Olio et al. 2010). Evaluating the various aspects of PT could highlight the areas in which it has poor performance to improve service and thereby obtain new users. SQ is a composite concept; it can be evaluated through the perceptions and opinions of the users or through a range of simple disaggregated performance measures collected by the service operators (Federal Transit Administration et al. 1999; Eboli and Mazzulla 2011). Therefore, there are two different agents that measure SQ. First, the service operators provide a quantitative indicator (Objective Indicator) that can be compared with a standard or past performance, but this indicator provides no information in itself regarding how “good” or “bad” a specific result is. Second, the measure of the users (Subjective Indicator) is derived from customer satisfaction surveys (CSS), which provide

qualitative measures of transit SQ related to the perceived discrepancy between the actual and ideal levels of service (Nathanail 2008).

Both indicators are crucial to evaluate the performance of a transit service; if either is not considered, there will be missing information, and the results will not effectively reflect reality. A useful and reliable measurement tool of transit performance could be obtained by combining these two types of measures (Tyrinopoulos and Aifadopoulou 2008; Nathanail 2008; Yeh et al. 2000). The use of a combination methodology allows the assessment of a concrete public transportation service and the solution of cost problems and constraints of service operators. The measurement tool is composed of the Subjective Indicator (S), which shares the crucial aspects that accurately reflect the needs of customers and potential customers, and the Objective Indicator (O), which quantitatively evaluates the performance of the service in comparison with previously established standards of performance in a previous period (Eboli and Mazzulla 2011; Federal Transit Administration et al. 1999; Nathanail 2008).

When S is calculated, it is necessary to consider that the quality of the PT is perceived by many different types of users who have different needs and personal characteristics that require individual attention (Zhou et al. 2004; Paez 2006; Button and Hensher 2001). If these variations are not addressed, it can lead to biased results and conclusions that might not identify some relationships between the data and, thus, might not reflect reality (De Oña et al. 2013; De Oña et al. 2014). In the area of data mining, there are advanced segmentation techniques, such as Cluster Analysis (CA), which enable the reduction of such heterogeneity. This technique has been applied in transport engineering and other fields (e.g., Wen and Lai 2010; Shiftan et al. 2008; Prebensen 2005) with satisfactory results.

The goal of this paper is to apply a methodology that considers different typologies of passengers to evaluate how SQ is perceived and to understand the differences in the SQ gap obtained. An improved formulation of the composite indicator proposed by Eboli and Mazzulla (2011), which combines the subjective data with the objective data, is used and adapted to the specific case study represented by the metropolitan Metro of Seville (Spain). The proposed methodology introduces the use of CA to assemble more homogeneous groups of users and to calculate the composite indicator for different types of users. In this manner, a comparative analysis between the assessments of each obtained group was conducted to extract specific conclusions regarding the aspects that are most critical for each group of users and the main causes and solutions thereof.

The paper is structured as follows: the next section shows the methodology used to evaluate SQ through the composite indicators and define the different groups of users by CA and is followed by a description of the data used for the analysis—specifically, the data used to calculate the objective indicators—and the results of a CSS conducted to calculate the subjective indicators. In addition, this section describes the results of the CA that were applied to stratify the sample and define the different groups of users. The results obtained by applying the composite indicators to the whole sample and to each cluster are explained, and finally, the conclusions are reported.

Methodology

Composite Indicator

The main aim of a composite indicator is to obtain a measure of SQ that combines the service operator point of view and the PT user point of view. The methodology is based on developing an indicator that takes an intermediate value between the quality measurements of service, considering the slope of each one. This process provides a significant quality measurement of the service that is governed by two basic concepts: 1) both indicators have equal importance in evaluating SQ, and 2) an indicator with less heterogeneity or variance has greater repercussions on the composite indicator (Eboli and Mazzulla 2011).

Each attribute is measured by an S and an O, with S calculated by the average of the satisfaction rates expressed by a sample of users with respect to a service attribute (Parasuraman et al. 1985) and O calculated by the average of the estimated value that is assigned to performance indicators about the attribute by service operators or by mystery shopper surveys and compared with standards (Nakanishi 2003).

Many O values were calculated by comparing the value of the predefined parameter (P) with a standard of quality (Q). To obtain these indicators, the criterion proposed by Nathanail (2008) and used by Eboli and Mazzulla (2011) was adopted. A grade of zero is given to the indicator (O) in the event that the parameter is greater than or equal to double the standard, and 10 is given when it is less than the standard. Intermediate grades were calculated according to the following formula (1):

$$O = \left[\frac{2 \cdot Q - P}{Q} \right] * 10 \quad (1)$$

Subsequently, an optimization process based on the variances S and O results in a composite indicator (X) for each attribute. X, similar to indicators S and O, is expressed on a cardinal scale from 0 to 10, where 0 represents the lowest level of quality and 10 is the highest level of quality.

The mathematical formulation proposed by Eboli and Mazzulla (2011) and adopted in this paper is described in the following.

Let \bar{S}_k denote the average rate of satisfaction or user perception about generic service attribute k expressed by a user in a survey according to the specific scale of evaluation. S_k denotes the actual value of the indicator, and the distance between the actual value and the estimated value for indicator k is denoted by ε_k^{PER} , which represents the average error of the perception of the indicator due to heterogeneity in the judgment of different users (2):

$$\bar{S}_k = S_k + \varepsilon_k^{PER} \quad (2)$$

Let \bar{O}_k denote the estimated value of the objective performance indicator of generic service attribute k calculated based on service operator information and converted to the same scale of evaluation adopted for the satisfaction rates. In the same manner, O_k denotes the actual value of the indicator, and the distance between the actual value and the estimated value of indicator k is denoted by ε_k^{OBS} , which represents the

average error in the measure of the indicator caused by measurements that are made by instruments and equipment (e.g., the length of a line path) that are calculated as an average of elements that can have different values that vary in time, space and so on (3):

$$\bar{O}_k = O_k + \varepsilon_k^{OBS} \quad (3)$$

If m is considered to be service attributes adopted to describe the SQ of a transit system, the expressions of the indicators in terms of vectors are (4), (5) and (6):

$$\bar{S} = S + \varepsilon^{PER} \quad (4)$$

$$\bar{O} = O + \varepsilon^{OBS} \quad (5)$$

$$S = O = X \quad (6)$$

where the number of dimensions of all vectors is $[m \times 1]$.

One problem of optimization with a constraint is how to obtain vector X . This constraint consists of maximizing an objective function $Z[X]$ with a constraint that is sum of functions $Z_1(\bar{S}, X)$ and $Z_2(\bar{O}, X)$. Thus, a measure between vectors S and O can be considered vector X , which can be obtained through (7):

$$X^* = \operatorname{argmin}_{X \geq 0} [Z(X)] = \operatorname{argmin}_{X \geq 0} [Z_1(\bar{S}, X) + Z_2(\bar{O}, X)] \quad (7)$$

The functional structure of $Z_1(\bar{S}, X)$ and $Z_2(\bar{O}, X)$ varies with the nature of the information. In this case, the methodology to obtain information is from experimental surveys, so a statistical theory is adopted, specifically the generalized least squares method, which provides an estimation of a parameter vector starting from a system of linear stochastic equations (8).

$$X^{GLS} = \operatorname{argmin}_{X \geq 0} [Z(X)] = \operatorname{argmin}_{X \geq 0} \left[\sum_{k=1}^m \frac{(\bar{S}_k - X_k)^2}{\operatorname{var}(\varepsilon_k^{PER})} + \sum_{k=1}^m \frac{(\bar{O}_k - X_k)^2}{\operatorname{var}(\varepsilon_k^{OBS})} \right] \quad (8)$$

This expression means that the estimate of vector X is vector X^{GLS} , which minimizes the sum of the standard deviations of vectors S and O and sample estimates \bar{S} and \bar{O} . The standard deviations are weighted in inverse proportion to the variances of the errors; this fact indicates that the deviation of the sample estimation from a component of vector X will apply a greater weight with greater variability of the sample values from the mean values.

However, if the variance of an indicator, e.g., \bar{O}_k , is very low (near 0), the value of X_k would be the same as that of the indicator. This occurs because the weight associated with \bar{O}_k tends to infinity, and the second indicator (\bar{S}_k) would be ignored. The same would occur in the opposite case. This is the most inconvenient aspect of this optimization factor because subjective components usually have higher variances than objective components, for which the variance is sometimes null. Consequently, for many evaluated attributes of service, the composite indicator could tend to be a solely objective indicator, totally ignoring the subjective values. To solve this problem, the optimization function has been slightly modified to avoid the indicator X tending to the indicator with null variance by weighting indicators with the variance of the indicator's errors plus one. The new formulation of the composite indicator is the following (9):

$$X^{GLS} = \operatorname{argmin}_{X \geq 0} [Z(X)] = \operatorname{argmin}_{X \geq 0} \left[\sum_{k=1}^m \frac{(\bar{S}_k - X_k)^2}{1 + \operatorname{var}(\varepsilon_k^{PER})} + \sum_{k=1}^m \frac{(\bar{O}_k - X_k)^2}{1 + \operatorname{var}(\varepsilon_k^{OBS})} \right] \quad (9)$$

There may be cases in which the variance for both indicators (S and O) is zero; thus, they would be weighed equally, and X would be a value halfway between them. Moreover, if S and O have the same value, it would be the ideal situation in which $S = X = O$. However, this is not the usual case. The more homogeneous the judgments expressed by the passengers are, the more reliable the estimated value of \bar{S}_k is. The \bar{O}_k values are generally more reliable than \bar{S}_k because the \bar{O}_k indicators are calculated based on almost accurate measurements effected in different periods but not very variable among the periods (Eboli and Mazzulla 2011). Therefore, following formulation (9), the normal value of X_k is slightly more weighted toward \bar{O}_k than \bar{S}_k .

Cluster Analysis

CA is a technique that is used to segment a group of data (e.g., numbers, things, or events), and it is based on heuristic techniques that attempt to maximize the similarity among items in a group and obtain the maximum differences between items in distinct groups (Fraley and Raftery 1998; De Oña et al. 2013).

To conduct a CA, all methodologies that can be used to achieve clustering segmentation are similarly valid. There is no universal measurement that can compare diverse cluster techniques and classifications because these methodologies are merely exploratory and are used primarily to analyze the groups that are obtained (DeSarbo and Mahajan 1984). However, the Latent Class Clustering (LCC) methodology has significant advantages over the others (Alarcon-del-Amo et al. 2011; de Oña et al. 2013; Hair et al. 2010; Magidson and Vermunt 2002; Vermunt and Magidson 2005):

- It is possible to consider different variables without the need for *a priori* standardization that could influence the results.
- LCC allows classification of probabilities through the use of the membership probabilities of each item, which have previously been classified using the maximum likelihood.
- LCC uses measures that are not based on the distance between data, so the standardization of data has no effect on the final clusters.
- It does not demand a large space in the memory of a computer, allowing the construction of models with large amounts of data.
- The models can usually incorporate independent variables, known as covariates, or grouping variables, that can be used to describe the latent classes rather than defining them.

The formal definition of LCC is as follows: Consider a data sample of N data measured with a set of observed variables, $Y_1 \dots Y_j$, which are considered indicators of a latent variable X and form an LCM with T classes. If each observed value contains a specific number of categories, where Y_i contains l_i categories, within $i=1 \dots j$, then the manifest variables form a multiple contingency table with $\prod_{i=1}^j l_i$ response patterns. If π denotes

the probability, $\pi(X_t)$ represents the probability that a randomly-selected case belongs to latent class t , within $=1, 2 \dots T$. The regular expression of LCMs is given by (10):

$$\pi_{Y_i} = \sum_{t=1}^T \pi_{X_t} * \pi_{Y_i|X_t}, \quad (10)$$

where Y_i is the response pattern vector of case i ; $\pi(X_t)$ is the prior probability of membership in cluster t ; and $\pi Y_i|X_t$ is the conditional probability that a randomly selected case has response pattern $Y_i=(y_1 \dots y_j)$, given its membership in class t of latent variable X . Local independence is the underlying assumption to be verified, so Equation (9) is rewritten (11):

$$\pi_{Y_i} = \sum_{t=1}^T \pi_{X_t} * \prod_{i=1}^j \pi_{Y_{ij}|X(t)} \text{ with } \sum_{i=1}^j \pi_{Y_{ij}|X(t)} = 1 \text{ and } \sum_{t=1}^T \pi_{X_t} = 1 \quad (11)$$

The estimation of the model is based on the nature of the manifest variables because it is assumed that the conditional probabilities may follow different formal functions (Vermunt and Magidson 2005). The method of maximum likelihood is the most widely-used method for estimating the model parameters. Once the model has been estimated, the cases are classified into different classes by using the Bayes rule to calculate the *a posteriori* probability that each subject n comes from class t (the model's estimated values) (12):

$$\pi_{X_t|Y_i} = \frac{\widehat{\pi}_{X_t} * \widehat{\pi}_{Y_i|X_t}}{\widehat{\pi}_{Y_i}} \quad (12)$$

In practice, a set of probabilities is calculated for each response pattern, and the case is assigned to the latent case in which the probability is the highest. Thus, a specific user may belong to different latent cases with specific probabilities of membership (with 100% being the sum total of the membership probabilities). Magidson and Vermunt (2002) and Vermunt and Magidson (2005) provide a detailed explanation of LCC analysis.

The objective in this methodology is to find the optimal number of clusters to align the database with the model. The criteria of selection are based on three information criteria: the Bayesian information criterion (BIC) (Raftery 1986), the Akaike information criterion (AIC) (Akaike 1987), and the consistent Akaike information criterion (CAIC) (Fraley and Raftery 1998). The information criteria and criteria of representativeness and of characterization are used to evaluate the optimal number of clusters because it is important that the obtained groups have remarkable characteristics and be easily characterized. The optimal number of clusters is the one that minimizes the score of these criteria, thus making them more parsimonious and better adapted to the study data (De Oña et al. 2014).

Application

Study Case

The transit system analyzed in this paper is the Metro transit service of Seville, a city located in the south of Spain. The municipality of Seville has a population of approximately 700,000 inhabitants in an area of 140.8 km². The population density is approximately 4,950 inhabitants/km². In 2014, the numbers of private cars and

motorcycles per 1,000 inhabitants were 466 and 131, respectively. The most recent mobility household survey was conducted in 2007, when the analyzed transit system still was not in operation. Nevertheless, in 2007, the modal split showed a predominance of private vehicle (53.9%) against the public transport modes (10.4%) and walking and cycling modes (35.7%).

The analyzed new Metro system entered operation in 2009 and currently consists of a sole line characterized by a length of 18 km (10.08 km underground) and 21 stations that connect 4 of the main municipalities in the metropolitan area of Seville. These four boroughs register a population of approximately 850,000 people. In 2013, the Metro carried more than 13.7 million users. This Metro system coexists with other transit alternatives in the city of Seville, such as a suburban train (5 lines), metropolitan buses (64 lines), urban buses (51 lines), a tram (1 line), and public bicycles (250 facilities and more than 2,500 bicycles for hire), all of which are coordinated by the Transport Consortium of Seville. Moreover, bicycles have significantly increased their importance following the construction of numerous cycle paths (80 km) and the creation of some parking for bicycles. In fact, most Metro stations have parking facilities for bicycles in their vicinity (distances less than 250 m).

An online CSS was addressed to a sample of 3,198 users of Line 1 of Metro de Sevilla in June 2014. Previously, a face-to-face pilot survey was carried out to check the soundness of the questionnaire and perform some modifications, reformulating the way some attributes were introduced, removing inappropriate questions, changing the order of the sections, and so on. The questionnaire adopted for conducting the CSS was then divided into four main sections:

- Part A – attitude of users towards Metro service. In this section, the user rates the different aspects related to their experience with the Metro service. The questions are measured on an 11-numeric scale defined as 0 = totally disagree and 10 = totally agree.
- Part B – perceptions of users about service characteristics. In this section, the user directly rates the different service aspects that they use in Metro Sevilla and provides a global score for the service. This part was developed according to an extensive literature review and the European Norm CEN 320/TC-EN 13816:2002 and contains 37 questions related to various aspects of the Metro service, such as availability of the service, accessibility, information, timeliness, attention to clients, comfort, safety and environmental pollution. The perceived level of quality of each of the 37 attributes was surveyed on an 11-numeric scale from 0 to 10 (0 = poorest quality and 10 = highest quality). Respondents also rated their overall perceived level of quality of the Metro service according to the same scale. In this paper, 18 of these service attributes were used for the analysis by a composite indicator; the other 19 were not considered because there were no elements for the calculation of their respective objective indicators and it was impossible to calculate the composite indicator with only their subjective indicators. User perceptions about the 18 service attributes were adopted to define the subjective indicators for the calculation of the composite index.

- Part C – general information on the trip. In this section, the users score their travel.
- Part D – socioeconomic characteristics. This part has 11 questions related to age, gender, income, level of study, and labor situation. The different typologies of questions used are categorical answer, comment box, and multiple choice.

Service Quality Attributes and Objective Indicators

The majority of the objective indicators related to service attributes were calculated by adopting the criteria of Nathanail (2008), Cascetta and Carteni (2014), and Eboli and Mazzulla (2012). In the following, we propose a detailed description of the calculation of the indicators, providing an objective measure of the 18 analyzed SQ attributes. These attributes concern six different service aspects (Table 1).

TABLE 1.
Description of Attributes of Service Quality

Service Aspects	Attribute
Availability	A1. Time of performance
	A2. Number of trains per day (frequency of service)
	A3. Proximity of stop to origin and/or destination
	A4. Regularity of service (absence of interruptions caused by breakdown or incidents)
Accessibility	A5. Easy connection with other transportation modes such as bike rental, bikes, buses, etc.
	A6. Performance of lifts and escalators
	A7. Ease for persons with disabilities to access Metro
	A8. Performance of validating tickets machines
Time	A9. Punctuality
	A10. Speed of trip
Customer Care	A11. Performance of customer service (offices, website, contact by phone, etc.)
Comfort and Amenities	A12. Cleanliness of stations
	A13. Cleanliness of vehicle
	A14. Lighting in stations
	A15. Lighting on vehicle
	A16. Availability of Internet and phone service in stations and on vehicle
Safety and Security	A17. Sense of security against theft and aggression in stations and on vehicles
	A18. Sense of security against slipping, falling, accidents at vehicle doors and escalators

Availability

This aspect is described by four attributes that represent the availability of the service in time and space.

“Time of performance” is the number of hours during a day that Metro service is provided. It is calculated as the average value of the number of hours per day in all considered months. It is compared with 20 h, which is the target reference schedule for other Metro services in Spain.

“Frequency” measures how often the transit service is provided. It is calculated with respect to the frequency interval per half hour. It is provided by the Metro Seville

operator and is compared with two target references: 5 min for peak hours and 7.5 min for the rest of the day. These target references are contractual conditions for Metro Seville.

“Proximity to origin and/or destination of users” measures the average distance (or time) between origins, destinations, and Metro stations. It is evaluated based on two average times: the time between origin and Metro station, and the time between Metro station and destination. The target reference for both quantities is 7.5 min (Nakanishi 2003).

“Regularity of service (absence of interruptions caused by breakdown or incidents)” is the evenness of the intervals between transit vehicles. It is calculated with respect to the number of kilometers per breakdown. This ratio was obtained for each month, and the target reference is the minimum percentage in the interval of the month that was considered. This target reference is a contractual condition for Metro Seville.

Accessibility

This aspect is explained by four attributes that represent the facility's access to stations and connection with others transports.

“Easy connection with other transportation modes such as bike rental, bikes, and buses” was evaluated with respect to the number of connections (C). It considers nine connections: taxi, urban bus, interurban bus, bicycle, tramway, surface and underground parking for private cars, train, and parking for bicycles. The final value is obtained through a logarithm equation (Eq. 13): if the number is 0, 0.5 is obtained, because walking will always be possible. The number but 1 is obtained if all of the connections are available at this station. Thus, the target reference is nine connections.

$$y = 0.5 + 0.228 * \ln(C) \quad (13)$$

“Performance of lifts and escalators” measures the functionality of the lifts and escalators in all stations of Metro Seville. This is evaluated with respect to the average effective performance of lifts and escalators for each month. Through comparison between the target time of performance and the real time of performance, an average is obtained. The target reference is the considered contractual target time.

“Ease for persons with disabilities to access Metro” measures the functionality of the lifts, which is the only way for a person with disabilities to access the station. Therefore, it is evaluated based on the comparison between the target time of the lift performance and real time. In the same way, “Performance of validating ticket machines,” which measures the functioning of the validating machines, is evaluated based on the comparison between the target time of the validating machine performance and real time. In both cases, the target reference is the considered contractual target time.

Time

This aspect is described by two attributes that represent the wait time between one train and the next and the speed of traveling by metro.

“Punctuality” measures the number of runs that arrive/depart on time. It is evaluated with respect to the percentage of regular trains. This concept is defined by Metro Seville as the time between a train's departure from a station and that of the preceding train if

lower than its scheduled headway plus 10% (i to $i+10\% \times i$). Metro Sevilla use a definition of regularity indicator similar to that of a bus operator, Lisbon Carris (Trompet et al. 2011). The Metro Sevilla operator provides different tables with the percentage for each day at each station. Thus, an average of the values for all months and stations is considered, and this is adopted as the value for this indicator.

“Speed of the trip” measures the average speed that is provided per day. This is evaluated based on the ratio (Eq. 14) between the total commercial kilometers (L) traveled and the performance effective hours (T_e). It is weighted by the number of trains (N_i) and the ratio between its intervals of time performed (T_{ei}) and the sum of the total target hours (T_t).

$$V_c = \frac{L}{T_e} * \left(\sum_{i=1}^N N_i * \frac{T_{ei}}{T_t} \right) \quad (14)$$

The average commercial speed in three other Spanish undergrounds is taken as the target reference.

Customer Care

This aspect consists of one attribute that represents the speed with which employees answer the passenger applications/suggestions.

“Performance of customer service (office, website, contact by phone, dealing with complaints, etc.)” measures the performance of the customer service system through the average answer times for the total number of complaints and suggestions. The average answer time from the previous six months is compared with the target reference, which is the average answer time for the previous three years. The methodology used to establish the target reference is similar to that used by Nathanail (2008), who examined safety aspects.

Comfort and Amenities

This aspect is described by five attributes that represent the cleanliness and lighting of vehicles and stations and smartphone (phone and Internet) coverage.

“Cleanliness of vehicle” and “Cleanliness of station” are evaluated based on a rate provided by an inspector from Administration. This ratio measures the level of dirt in the Metro interior and exterior, and its range is from 0 to 3, where 0 = total filth and 3 = perfect cleanliness and, hence, the target reference. These attributes are contractual conditions for Metro Seville.

“Lighting on vehicle” and “Lighting in stations” were evaluated based on a rate provided by an inspector from Administration. This ratio measures the visual clarity in the Metro interior and exterior, and its range is from 0 to 1, where 0 = total darkness and 1 = perfect clarity and, hence, the target reference. These attributes are contractual conditions of Metro Seville.

“Availability of Internet and phone service in stations and on vehicle” is evaluated based on the rate between the line length where telephonic coverage is available and the total length of the line. Compared with other metros in Spain, where there is phone coverage and 3G along the total length of the line, the target reference is established as the total length of Metro Seville.

Safety and Security

This aspect is explained by two attributes that represent safety and security issues.

It is necessary to differentiate between two typologies in this section (Carr and Spring 1993; Eboli and Mazzulla 2008). Safety refers to involvement in an accident of transport—in this case, a Metro accident—and is measured by the “Sense of security against slipping, falling, and accidents at vehicle doors and escalators.” It is calculated based on the average number of Metro accidents during the first half of the previous year and is compared with a target reference, the average number of Metro accidents during the first halves of the previous three years (Nathanail 2008).

In contrast, security refers to victimhood of a crime or a robbery and is measured by the “Sense of security against theft and aggression in stations and on vehicles.” It is calculated based on the mean number of complaints registered during the first half of the previous year, which is compared with a target reference, the average number of Metro complaints registered during the previous three years (Eboli and Mazzulla 2011).

Sample Characteristics

The general characteristics of the collected sample are shown in Table 2. It is made up of more females (53.30%) than males (46.70%). The majority of respondents were ages 18–25 (41.70%), and the next largest groups were 26–40 (28.90%) and 41–65 (25.60%). There is an underrepresentation of the groups younger than age 18 and older than age 65 (2.80% and 1.00%, respectively). The main reasons for traveling are studies (38.80%) and work (35.50%), with leisure and other reasons showing a similar percentage (15.30% and 10.30%, respectively). Most users travel daily (52.10%). Users generally have a high school diploma (41.90%) or are university graduates (48.50%), but there is also a small group who have only secondary obligatory education (8.40%). Most of the sample has a low household monthly family income (less than 1,800 euros). The sample of users is fairly equally distributed between those who have a private vehicle available to make the trip and those who do not (54.78% and 45.22%, respectively). The users in the sample are sufficiently satisfied with the overall service (average rate of 7.6).

TABLE 2. Distribution of Complete Sample

Variable	Characteristic
Trip purpose	Work (35.5%), studies (38.9%), leisure (15.3%), other (10.3%)
Frequency of use	>4 days/week (52.0%), 3–4 days/week (17.9%), 1–2 days/week (13.7%), occasionally (16.4%)
Gender	Man (46.7%), woman (53.3%)
Availability of:	Driver license (75.0%), access to private car (54.7%), access to motorcycle (6.7%), access to bicycle (43.2%) none (12.0%)
Age	<18 (2.8%), 18–25 (41.6%), 26–40 (28.9%), 41–65 (25.6%), >66 (1.0%), no response (0.1%)
Level of studies completed	None or secondary school (9.0%), high school or professional education (42.0%), bachelor's or higher (48.5%), no response (0.6%)
Household monthly income	<1,201 (28.8%), 1,201–1,800 (21.0%), 1,801–2,400 (16.5%), >2,401 (16.0%), no response (17.7%)
Satisfaction with Overall SQ	Average (7.6), standard deviation (1.5)

Stratification of Sample by Cluster Analysis

The methodology applied to obtain the optimal number of clusters consists of four steps.

Step 1: Selection of variables. In total, 17 variables were included in the CA. (The used variables and their sample distribution for each cluster are reported in Table 5.)

Step 2: Identification of optimal solution. This is an iterative process in which the number of clusters was selected and refined by the variables. The Latent GOLD software program was used to obtain the optimal number of clusters. This step was necessary because not all variables were valid for use in the CA. Therefore, the objective was to achieve homogeneity of variables in all clusters to be compared between them.

First, all variables were input to the software, and 1–10 clusters were simulated. Using the Wald Test, it was possible to determine which variable should be deleted or changed into a covariable. A covariable can describe or predict (instead of defining or measuring) the latent class and reduce the classification error (Vermunt and Magidson 2005). When no variable was rejected by all clusters, the optimal number of clusters was selected. For this objective, the information criteria (BIC, AIC, CAIC), representativeness, and simplicity of the structure criteria were used. Although the information criteria (when the variation of BIC, AIC, and CAIC in % is less than 1%, it is the optimal number of clusters (De Oña et al. 2013)) noted that the optimum was five clusters, by following the representativeness and simplicity of the structure criteria, six clusters were selected. The main reason for this conclusion was that the representativeness of the six clusters was better than that of the five clusters because more heterogeneous profiles of users among clusters and, in turn, greater homogeneity within clusters were obtained. Furthermore, the model provided six additional clusters of user profiles that were sufficiently differentiable from the standpoint of characterization. Moreover, the selection of six clusters improved the information indicators that were considered, and the complexity of the model did not substantially increase (Table 3).

TABLE 3.
Variation of Information Criteria

N° Clusters	BIC (%)	AIC (%)	CAIC (%)
3	1.85%	2.11%	1.81%
4	0.83%	1.09%	0.79%
5	0.33%	0.59%	0.29%
6	0.25%	0.51%	0.20%
7	0.16%	0.43%	0.12%

Step 3: Development of the cluster model. Once the number of clusters was chosen, cluster depuration was performed. In this case, the variables in this specific group were deleted or changed, regardless of whether they were in other groups. The process was similar to that of Step 2. The adjustment parameters obtained for this cluster model are shown in Table 4. It was observed that the information criteria (BIC, AIC, and CAIC) improved significantly from the model obtained in Step 2 of the CA (Table 3). Finally, the model calibration of six clusters showed an entropy value of 0.88, indicating good separation of the clusters and interpretability (De Oña et al. 2013; Depaire et al. 2008).

Additionally, it is important to emphasize the classification of error, which did not exceed 0.082 and was within the range reported in the other studies (Vermunt and Magidson 2002; Reyna and Brussino 2011). Additionally, the correlation coefficient showed a value of 0.85, near the value of a perfect setting 1, which indicated a good model fit (Rondán-Catalan et al. 2007).

TABLE 4.
Model Parameters

Model CA with Six Clusters	
Number of cases	3198
Number of parameters (Npar)	352
Statistics of Cluster Classification	
Error classification	0.082
Entropy	0.883
R	0.854

Step 4: Cluster characterization. To characterize each cluster, a ratio based on the frequency distribution (in %) of the categories of all the variables considered in the analysis is used. This ratio compares, in percentages, the distribution of frequency of a category of a variable in a cluster with the distribution of the frequency of that category in the total sample. Following this indicator, the categories of variables that are overrepresented in each group can be identified. If the ratio is more than 1.5 for a category, the category is considered to be highly over-represented; if it is between 1.4 and 1.5, it is moderately over-represented; if it is between 1.3 to 1.4, it is slightly over-represented; and if it is less than 1.3, nothing can be said about this category because it follows a similar distribution to the total sample.

The results are described in the following, and the characteristics of the identified groups of users are summarized in Table 5:

- Cluster 1 (CL1) can be named “High-income users with predisposition to use private cars” and comprises 26.46% of the whole sample. It is predominantly made of users from ages 26–65 whose employment status is employed or less predominantly retired and who hold a university degree. Users belonging to CL1 predominantly show the availability of a personal vehicle and driver license and make sporadic trips for work or other reasons, with Metro stations not near the origin or destination of their trip. Moreover, a notable proportion of these users jointly use a car and the Metro service for their trip, and they mainly consider the car an alternative to the Metro. Finally, CL1 users show a greater level of agreement in considering the lack of parking, traffic congestion, and, less frequently, the unavailability of their vehicle as the main reasons to use the Metro. This cluster is characterized by a predominant household monthly income of more than 2,401€.
- Cluster 2 (CL2) represents 18.26% of the complete sample and comprises “High-income users with predisposition to use the Metro.” This cluster contains a higher proportion of users over age 26 whose status is employed or retired; university-level studies predominate. These users notably show the availability of private

cars and bicycles; they mainly make sporadic and short trips for work, other reasons, and leisure. Additionally, they consider walking, tram, and motorcycle as alternatives to the Metro more frequently than the general trend of the sample. A notably higher proportion of CL2 users consider speed, comfort, and fare as the main reasons for using the Metro; thus, we can consider this group of users as those who have a predisposition to use the Metro and are attracted by service characteristics. Moreover, this group shows the highest average perception of the overall SQ. Additionally, in this case, a monthly household income level of more than 2,401€ is more predominant than in the complete sample.

- Cluster 3 (CL3) comprises 16.73% of the sample and is named “Captive young students.” This cluster predominantly comprises users younger than age 26 who are students seeking a high school or professional education degree. This cluster shows a higher proportion of captive users (who have no alternative mode of transportation to the Metro for their trip) and users with no driver license or personal vehicle available; more than 20% of CL3 users live in large families and habitually use the Metro to reach places of study.
- Cluster 4 (CL4) comprises 14.92% of the whole sample and predominantly includes “Captive university students” who are mainly users ages 18–25 and are students seeking a high school or professional education degree. This cluster contains a distinctively higher proportion of users who used the Metro because they are captive to the Metro service, unable to drive (no private vehicle available, although they have a driver license), and who habitually take long trips for studies. Moreover, users in the CL4 cluster show the lowest assessment of the overall SQ.
- Cluster 5 (CL5) represents 12.32% of the sample and can be named “Non-captive university students.” This cluster mainly consists of users aged ages 18–25 years and are students seeking a high school or professional education degree and travel daily for studying. In contrast to CL4, this cluster did not show a proportion of captive users higher than the general trend; almost 90% of users had access to a private car. Furthermore, compared with the complete sample, these users more frequently stated that they were using the Metro service due to traffic congestion, lack of parking, or unavailability of a private vehicle; a relevant part of these users reach the Metro station by car.
- Cluster 6 (CL6) is the last cluster and comprises 11.41% of the whole sample. They are “Users with low income and high predisposition to use the PT” who more frequently declared that urban and metropolitan buses and trams are transport alternatives to the Metro service and that they use the Metro service mainly due to a lack of a driver license and the unavailability of a private vehicle. Approximately 40% of the users have no transport means available. However, they show a relatively higher average perception of the overall SQ. CL6 contains a high proportion of users over age 26 who are without education. There is a higher proportion of employed or, less prominently, retired users, and the percentage of women (65%) is considerable compared with the other clusters. They mainly travel for work or other reasons. CL6 shows a predominant monthly household income level less than €1,200.

TABLE 5.
Distribution of Clusters
according to 17 Variables
Adopted in Cluster Analysis

Variable	CL1 (842*)	CL2 (584*)	CL3 (534*)	CL4 (479*)	CL5 (394*)	CL6 (365*)
V1. Average overall SQ (Overall SQ)	7.5	8.3	7.6	7.1	7.4	7.7
(Standard deviation)	(1.6)	(1.1)	(1.5)	(1.5)	(1.5)	(1.4)
V2. I use the metro service because of						
Fare	7.5	14.7	10.7	7.1	13.2	9.0
Comfort	54.7	62.7	46.0	37.1	49.2	42.5
Quickness	59.2	84.2	65.8	58.3	66.0	68.5
Frequency	22.8	28.6	32.1	27.0	36.3	33.2
Ecologic reasons	21.2	19.0	14.2	8.2	18.5	14.2
I do not have a driver license	0.1	0.0	63.2	1.3	0.0	32.9
I do not own private car	1.8	3.4	46.5	53.5	2.0	52.6
It is my unique alternative	3.2	1.0	32.9	30.2	7.1	14.5
Lack of parking	57.5	38.9	6.2	6.3	59.6	5.2
Traffic jam	40.3	27.6	8.0	7.3	48.0	6.8
I can use my own private car for any reason	8.1	3.4	0.7	10.1	10.9	2.7
V3. Trip purpose						
Work	61.9	56.8	2.8	4.8	4.1	62.5
Studies	5.3	4.3	77.7	83.9	88.1	2.7
Leisure	18.8	21.1	15.4	8.6	5.1	17.8
Other	14.0	17.8	4.1	2.7	2.8	17.0
V4. Arrival transport from origin to station						
On foot	43.5	80.0	69.3	70.0	53.3	68.5
Other	5.5	2.2	6.4	10.3	1.8	7.9
Bus	5.1	1.0	13.8	16.6	9.4	18.9
Car	45.9	16.8	10.5	3.1	35.5	4.7
V5. Length from origin to station						
< 10 min	45.4	91.1	57.6	55.3	63.5	54.0
10–15 min	26.7	8.0	17.9	18.9	18.8	21.1
> 15 min	27.9	0.9	24.5	25.8	17.8	24.9
V6. Arrival transport from station to destination						
On foot	83.6	93.8	85.0	86.6	90.6	77.5
Other	4.4	1.9	2.6	4.6	0.0	7.7
Bus	5.3	1.5	8.4	7.1	4.3	11.5
Car	6.6	2.7	3.9	1.7	5.1	3.3
V7. Length from station to destination						
< 5 min	26.2	70.9	39.4	37.1	61.7	32.3
5–10 min	34.2	26.0	31.8	36.1	20.8	26.8
> 10 min	39.6	3.1	28.8	26.8	17.5	40.8

Variable	CL1 (842*)	CL2 (584*)	CL3 (534*)	CL4 (479*)	CL5 (394*)	CL6 (365*)
V8. Total length						
< 25 min	6.9	87.3	33.3	28.5	34.0	27.1
25–40 min	57.7	12.7	36.6	37.5	41.9	40.5
> 40 min	35.5	0.0	30.1	34.0	24.1	32.3
V9. Type of ticket						
Day ticket	8.9	10.3	12.0	9.4	7.1	10.4
Bonometro	39.1	41.1	24.9	22.9	25.1	26.8
Consortium card	52.0	48.6	63.2	67.7	67.8	62.7
V10. Frequency of use						
>4 days/week	43.5	42.3	59.4	59.9	66.0	51.2
3–4 days/week	17.2	15.2	17.2	21.1	18.3	20.5
1–2days/week	17.6	17.5	11.6	10.9	6.3	13.2
Occasionally	21.7	25.0	11.8	8.1	9.4	15.1
V11. If you could not use your own car, what alternative do you use?						
On foot	1.8	6.7	5.0	2.1	3.8	3.0
Bicycle	5.2	8.6	10.3	11.9	3.0	7.9
Urban Bus (Tussam)	12.9	27.2	40.2	43.6	18.3	40.8
Interurban Bus	9.0	5.8	17.9	18.0	14.7	32.3
Private Car	58.5	38.9	12.7	13.0	49.7	4.1
Motorcycle	1.9	3.6	1.1	2.1	1.0	0.5
Tram	0.4	2.4	0.6	1.0	0.5	2.5
Combination of vehicles	9.1	5.5	10.5	6.9	8.1	7.1
Other	1.2	1.4	1.7	1.3	0.8	1.6
V12. Gender						
Man	48.8	50.2	43.6	48.2	49.7	35.1
Woman	51.2	49.8	56.4	51.8	50.3	64.9
V13. Availability of						
Driver license	98.8	95.5	0.2	99.6	97.5	39.7
Access to private car	93.8	83.4	2.6	20.0	87.8	4.4
Access to motorcycle	7.2	11.0	3.7	4.6	8.4	4.1
Access to bicycle	38.2	39.6	53.6	45.5	52.0	32.6
None	0.0	0.0	43.8	0.8	0.0	39.7
V14. Age						
<18	0.0	0.0	16.5	0.2	0.0	0.3
18–25	6.3	3.1	78.1	89.1	92.6	13.7
26–40	44.9	47.1	5.2	10.6	6.6	45.5
41–65	47.5	47.3	0.0	0.0	0.8	38.1
>66	1.3	2.2	0.0	0.0	0.0	2.5
No response	0.0	0.3	0.2	0.0	0.0	0.0

Variable	CL1 (842*)	CL2 (584*)	CL3 (534*)	CL4 (479*)	CL5 (394*)	CL6 (365*)
V15. Level of studies completed						
None or secondary School	6.7	10.3	14.6	1.3	1.3	22.7
High school or professional education	26.1	23.3	66.1	63.3	61.2	24.4
Bachelor's or higher	66.6	65.6	18.4	35.5	37.3	52.3
No response	0.6	0.9	0.9	0.0	0.3	0.5
V16. Employment situation						
Employee	80.4	76.0	0.2	3.6	2.3	67.9
Student	0.5	0.9	97.0	90.6	92.6	0.8
Retired	3.6	6.3	0.2	0.0	0.0	3.8
Other	15.5	16.8	2.6	5.9	5.1	27.4
V17. People who live at your home						
1–2 people	33.9	39.0	9.3	9.4	3.8	39.7
3–4 people	54.7	52.7	68.2	70.6	72.1	48.8
More than 4 people	11.4	8.2	22.4	19.9	24.1	11.5

Results and Discussion

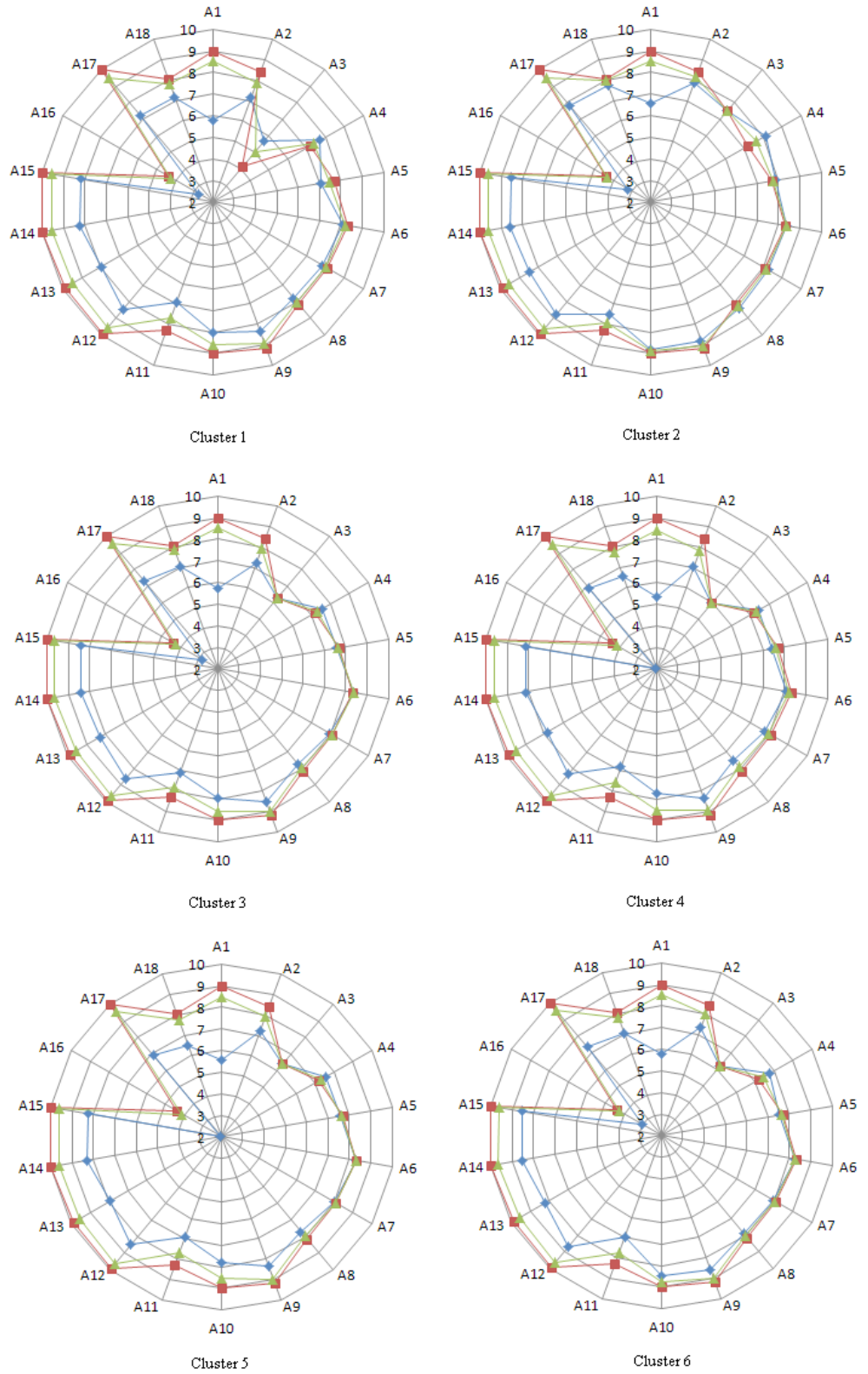
Composite Indicator for Each Cluster

The composite indicator was calculated for each cluster, and some interesting differences were found among the clusters. It must be emphasized that in this analysis, only subjective indicators change across clusters, whereas the objective indicators are always the same for all cases, with the only exception of the third attribute “Proximity of stop to origin and/or destination,” which changes due to its measurement methodology. Table 6 shows the results for each cluster and the results for the overall sample, which are compared and discussed with each cluster. Figure 1 shows the results for each cluster.

TABLE 6. Calculation of Composite Indicators of Clusters (CL1, CL2, CL3, CL4, CL5, CL6)

Attribute	Overall Sample						CL1		CL2		CL3		CL4		CL5		CL6		CL1	CL2	CL3	CL4	CL5	CL6
	\bar{X}_k	Var	\bar{O}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	\bar{S}_k	Var	X_k	X_k	X_k	X_k	X_k	X_k
A1	5.82	8.17	8.95	0.52	5.77	8.23	6.58	6.36	5.70	8.53	5.32	8.00	5.57	8.42	5.78	9.08	8.50	8.55	8.50	8.55	8.51	8.43	8.48	8.54
A2	7.29	4.42	8.43	3.11	7.16	4.45	7.85	3.12	7.19	4.92	7.01	4.23	7.22	4.65	7.31	5.08	7.88	8.14	7.88	8.14	7.92	7.80	7.92	7.98
A3	6.30	8.22	2.60-6.75	9.28-13.97	5.67	8.39	7.51	5.64	6.27	8.79	5.97	8.18	6.41	8.11	6.18	8.00	4.21	7.24	4.21	7.24	5.38	5.03	5.85	4.92
A4	7.72	4.17	7.16	2.56	7.73	3.84	8.15	3.12	7.54	4.85	7.47	4.17	7.56	4.98	7.79	4.32	7.40	7.62	7.40	7.62	7.31	7.29	7.31	7.41
A5	7.46	4.13	7.67	1.70	7.09	4.44	7.84	3.34	7.56	3.85	7.43	4.24	7.56	4.08	7.55	4.48	7.48	7.74	7.48	7.74	7.63	7.59	7.63	7.63
A6	8.24	3.09	8.32	3.04	8.07	3.21	8.40	2.83	8.39	2.56	8.10	3.52	8.33	3.31	8.27	3.05	8.20	8.36	8.20	8.36	8.35	8.22	8.32	8.29
A7	7.96	3.80	8.11	2.13	7.84	3.69	8.27	3.00	7.99	3.89	7.74	4.11	8.01	4.30	7.96	4.03	8.00	8.18	8.00	8.18	8.06	7.97	8.07	8.05
A8	7.86	4.06	8.14	2.34	7.80	3.86	8.39	2.80	7.76	4.14	7.50	4.87	7.71	4.87	7.92	3.90	8.01	8.26	8.01	8.26	7.99	7.91	7.99	8.05
A9	8.49	2.39	9.18	0.32	8.35	2.41	8.81	1.50	8.50	3.09	8.38	2.04	8.36	3.06	8.57	2.20	8.95	9.05	8.95	9.05	9.02	8.94	8.98	9.00
A10	8.14	3.13	8.98	1.83	8.01	3.25	8.80	1.71	7.99	3.60	7.74	3.53	7.82	3.50	8.49	2.32	8.59	8.89	8.59	8.89	8.60	8.50	8.53	8.75
A11	7.05	4.58	8.30	3.51	6.90	4.96	7.54	3.55	7.10	5.16	6.82	4.00	6.90	4.45	6.96	4.93	7.70	7.92	7.70	7.92	7.79	7.60	7.66	7.72
A12	8.55	2.28	9.93	0.01	8.44	2.07	8.77	1.69	8.60	2.50	8.33	3.05	8.50	2.58	8.72	1.90	9.56	9.61	9.56	9.61	9.63	9.61	9.61	9.62
A13	8.11	3.04	9.90	0.03	7.97	2.99	8.46	2.09	8.27	3.04	7.84	3.53	7.93	3.64	8.22	2.99	9.50	9.54	9.50	9.54	9.57	9.51	9.54	9.55
A14	8.36	2.31	10.00	0.00	8.22	2.19	8.60	1.80	8.45	2.65	8.14	2.56	8.33	2.48	8.50	2.17	9.58	9.63	9.58	9.63	9.67	9.59	9.63	9.64
A15	8.33	2.34	10.00	0.00	8.21	2.15	8.53	1.95	8.41	2.55	8.12	2.84	8.25	2.71	8.51	1.87	9.57	9.63	9.57	9.63	9.65	9.61	9.63	9.61
A16	2.69	8.67	4.40	0.00	2.75	8.06	3.23	10.13	2.80	9.27	2.08	7.34	2.04	6.67	3.04	9.30	4.24	4.30	4.24	4.30	4.26	4.15	4.13	4.28
A17	7.27	4.80	10.00	0.21	7.24	4.54	7.82	3.12	7.31	5.38	6.84	5.08	6.91	5.98	7.35	4.75	9.50	9.50	9.50	9.50	9.57	9.47	9.54	9.54
A18	7.07	5.26	8.05	0.83	7.15	4.65	7.77	3.13	7.06	5.95	6.55	5.72	6.51	6.83	7.04	5.31	7.83	7.97	7.83	7.97	7.85	7.73	7.76	7.83
Mean	7.37	4.38	8.23	1.75	7.24	4.30	7.85	3.38	7.38	4.71	7.08	4.50	7.22	4.70	7.45	4.43	8.04	8.34	8.04	8.34	8.15	8.05	8.14	8.13

FIGURE 1.
Value of S (blue), O (red), and X (green) for all clusters



The main characteristic of CL1 (“High-income users with predisposition to use a private car”) is that the values are lower than in the complete sample, with all subjective values showing differences of less than 1 point, and the variance is slightly lower as well. Specifically, the “Proximity of stop to origin and/or destination” shows the worst subjective (5.67) and objective (2.60) evaluation and lower variance than the complete sample (9.28). These results might indicate that the origins/destinations are commonly farther than a 7.5-min walk from the Metro service. Furthermore, compared with the complete sample, in this cluster, more users prefer using their cars, but they must sporadically use the Metro due to limitations such as traffic jams and a lack of parking. Thus, despite the inconvenient distance from/to the origin/destination, the use of Metro by this cluster is an obligation, so they are critical.

In this cluster, aspects related to accessibility present worse subjective values than the complete sample. In the case of composite indicators, “Proximity of stop to origin and/or destination” is beneath the satisfactory threshold (5.00). Therefore, transit operators should consider this attribute to be a critical aspect to improve the perceived service quality of users in this cluster.

The main characteristic of CL2 (“High-income users with predisposition to use the metro”) is that it shows the highest level of perceived overall SQ evaluation and the best evaluation of the single service attributes. All subjective values have a higher value (average of 7.85) than the complete sample (7.37) and a lower variance (average of 3.38, one point below the corresponding value of the total sample). This shows that this cluster had homogeneous positive opinions about the service of the Metro. Furthermore, in contrast to CL1, the indicator related to proximity objectively (6.75) and subjectively (7.51) shows a notable improvement, a positive difference of 2.50 and 0.70, respectively, compared with the complete sample. This is the group of users for which stations are nearest their origins/destinations. The second indicator with a notable improvement is “Connections between metro stations with other transports” (almost 0.40 positive differences). This could be justified by arguing that in this cluster, users with private cars but who prefer to use the Metro for reasons of fare are over-represented. This predisposition to use the Metro jointly with the proximity of the stations to their origins/destinations could lead this cluster to perceive a very satisfactory level of SQ of the Metro. In this group, the subjective and objective indicators are very close, and there are some indicators related to accessibility aspects that have a greater subjective value than objective value; however, this difference is negligible. Thus, the perceptions and expectations of service of this group are similar to those provided by the service operator. Finally, with regard to the composite indicators, “Availability of Internet and phone service in stations and on vehicle” is the only indicator that is below the satisfactory level (4.30). This attribute is critical in all clusters and could be improved by installing phone and 3G coverage along a greater extension of the line.

CL3 (“Young and captive students”) has subjective values (average of 7.38) and an objective value (only B3 indicator) similar to those of the complete sample (average of 7.37). However, it has a slightly greater variance (average of 4.71). Therefore, this group has more heterogeneous opinions about the service, and it is more difficult to find agreement on specific aspects. There are possibly more extreme values that, on average,

have the same mean as the complete sample. However, in this cluster, young students whose age is under 18 or who are 18–25 are over-represented. They show the same opinions as the complete sample. They consider the time performance of the Metro to be unsatisfactory, probably because they like to use it during extreme schedules such as on weekends. This aspect is accentuated in this group because captive users are over-represented and because the only way to move across Seville is PT, similar to the Metro service. Furthermore, it is currently highly important for this demographic of user to be connected to the Internet everywhere; therefore, this aspect is also observed to be critical based on the results of the corresponding composite indicator.

The main characteristic of CL4 (“Captive university students”) is the lowest level of perceived overall SQ evaluation and of several service attributes, although this group of users is also satisfied with the service, given that the satisfaction with the overall service is equal to 7.1 and that most attributes are satisfactory for the users. This cluster shows a high level of variance (average of 4.50). As in CL3, users have more heterogeneous opinions about the service, but they are generally low. The objective value of “Proximity of stop to origin and/or destination” is 0.5 (3.73), below the corresponding value observed in the complete sample (4.13). Young students, whose age is 18–25, are over-represented in this cluster. They perceive that the provided service is slightly less satisfactory than the complete sample, and there are two aspects that they consider to have an especially lower level of quality—phone coverage (“Availability of Internet and phone service in stations and on vehicle”) and security (“Sense of security against slipping, falling, and accidents at vehicle doors and escalators”). The first may be a common necessity among young people who must be connected to the Internet at all times. The second is probably because in this group, people with a temporal disability are over-represented and they might not feel secure with the automatic doors or feel that they can fall down in some areas. Therefore, a strategy could be to provide users with a personal assistant service to help them if necessary, and their assessment of the security of the service might thus improve. However, the “Sense of security against slipping, falling, and accidents at vehicle doors and escalators” presents a limitation, similar to “Sense of security against theft and aggression in stations and on vehicle” and “Performance of customer Service,” showing a possible potential disconnect between S and O values. For instance, passengers could be unsatisfied with their sense of security; however, the objective score receives full marks because the actual number of reported crimes has not increased compared with past performance.

Finally, with regard to the composite indicators, “Availability of Internet and phone service in stations and on vehicles” (4.30), which is critical in all clusters, and “Proximity of stop to origin and/or destination” are the indicators that are beneath the satisfactory level. Students who must take long trips are over-represented, so it is important for them that the stations are near their origins/destinations. This aspect is accentuated in this group because captive students are over-represented.

It was worth noting that CL5 (“Non-captive university students”) is similar to both CL4 and the complete sample. Subjective values are slightly greater (average of 7.22), but they also show a greater variance (average of 4.70) than the complete sample (average of 4.38). However, non-captive users are over-represented in this group, although they

regard the use of the Metro as an obligation, probably because they do not use private cars due to traffic jams, lack of parking, and so on. Thus, they express their disapproval with the service. Singularly, the value of "Proximity of stop to origin and/or destination" is 5.85, near the critical level, which shows that users have stations nearer their origins/destinations than CL4 and the complete sample.

CL6 ("Users with low income and high predisposition to use the PT") users have subjective values (average of 7.45) and variance similar to the complete sample (average of 7.37 and 4.38, respectively). In the case of the objective value of "Proximity of stop to origin and/or destination," it presents a lower value than the complete sample. Therefore, this indicates that the stations are farther than for the complete sample. However, in this cluster, users over age 26 are over-represented. They show the same point of view as the complete sample. They regard the time performance of the Metro as unsatisfactory because they like to use the Metro in extreme schedules, possibly because they work at night. This aspect is accentuated in this group because captive people are over-represented, and the only way to move across Seville is by PT.

Conclusions

The exclusive use of either subjectively- or objectively-measured indicators could be insufficient to achieve a comprehensive assessment of the quality of a PT because of its diffuse, complex, and heterogeneous nature. This could provide a biased representation of reality. Therefore, it is necessary to combine all information collected from users and provided by transit agencies (subjective and objective data). The methodology that is proposed is quite adequate to achieve a composite indicator, equally considering both indicators and awarding indicators that present less heterogeneity in their assessments. The subjective data were provided by means of a CSS, and the objective data were calculated by adopting different criteria based on a comparison of the actual performance with standards. This point presents a limitation; indicators that have a standard based on past performance (e.g., "Performance of customer service," "Sense of security against theft and aggression in stations and on vehicles") present a possible potential disconnect between the S and O values. For instance, passengers could be unsatisfied with their sense of security; however, the objective score receives full marks because the actual number of reported crimes has not increased compared with past performance.

Another aspect that should be taken into account when considering opinions on the quality of service is that currently there is a great diversity of users with different needs who depend on such factors as their socioeconomic characteristics and travel patterns. Therefore, a segmentation of the sample was made to identify different groups of users and analyze their opinions about the service.

Because the conventional methods of segmenting categories according to some socioeconomic variables have disadvantages of under-representation or lack of different parameters, it is necessary to use cluster analysis as an effective segmentation procedure to facilitate the achievement of these objectives. Among the different CA techniques, it is noteworthy that there is no universal measure that serves to determine the best

technique because they are exploratory techniques. However, Latent Class Clustering has several advantages over other procedures, the main one of which is that several types of variables that are not segmentation-based measurements of the data can be used.

The combined use of these techniques was applied to the case of Metro Seville, and the benefits were 1) allowing the characterization of the different typologies of users, 2) enabling the study of each attribute for each cluster and, consequently, identifying the attributes that present the lowest and highest values, and 3) explaining and contextualizing for each of them the composite indicators and results, which allows agencies and consortiums of PT to provide different strategies for marketing for specific demographics of users.

Specifically, six different groups of users who have relatively different perceptions of the service were identified. Two groups of users are characterized by a wide presence of adults who travel for work and have a relatively high income level. These two groups differ regarding their opinions about the Metro service and the transportation modes in general: one group comprises people who are more disposed to use private cars, and the other group is more inclined to use Metro. The second cluster clearly shows better opinions of the service.

In addition to the two groups of adults, three groups of students were identified that differ according to the availability of driver licenses and private cars as alternative mode to the Metro. Specifically, there is one group of captive young students and one group of captive university students who use the Metro because they have no other alternatives; this last group expressed the worst opinions about the service. The third cluster of students mainly comprises non-captive university students who use the Metro because they want to use it; the last identified cluster is made up of users with low income and high predisposition to use PT who expressed higher rates of satisfaction regarding several attributes compared with other clusters.

Thus, it can be concluded that users who choose to travel by Metro but who have the option to travel by private car are more satisfied with the service, whereas users who use only the Metro because they do not have other alternatives (captive users) are more critical towards the service. This is a comforting result, which suggests that if transit operators offer services characterized by high levels of quality, public transport can become a real alternative to private cars. This is an important point for solving environmental and social problems resulting from the excessive use of private cars.

From a practical perspective of this research for service operators, the number of cases in which the objective and subjective rankings are not similar could be identified. If the target values for these indicators could be adjusted, then service operators may be able to simply measure the objective indicators, supplemented by less frequent customer surveys to ensure that the two types of measures are still aligned. This would facilitate a purely objective measurement of passenger satisfaction. Moreover, this would allow for more frequent measurement and incorporation into a service operator's performance management program while reducing measurement costs.

Acknowledgments

The authors would like to thank the ERDF of European Union for financial support via project “Mejora de la calidad del TP para fomentar la movilidad sostenible: Metro de Sevilla” of the “Programa Operativo FEDER de Andalucía 2007–2013.” We also thank to Public Works Agency and Regional Ministry of Public Works and Housing of the Regional Government of Andalusia, as well as the experts who participated in the questionnaire development and the team of trained interviewers who collected the data. Support from Spanish Ministry of Economy and Competitiveness (Research Project TRA2015- 66235-R) is also gratefully acknowledged

References

- Akaike, H., 1987. “Factor Analysis and AIC.” *Psychometrika*, 52(3).
- Alarcon-del-Amo, M. D., C. Lorenzo-Romero, and M. A. Gomez-Borja. 2011. “Classifying and Profiling Social Networking Site Users: A Latent Segmentation Approach.” *Cyberpsychology, Behavior, and Social Networking*, 14(9): 547-553.
- Button, K. J., and D. A. Hensher. 2001. “Modal Diversion.” *Handbook of Transport Systems and Traffic Control*, 3: 107-123.
- Carr, K., and G. Spring. 1993. “Public Transport Safety: A Community Right and a Communal Responsibility.” *Crime Prevention Studies*, 1: 147-155.
- Cascetta, E., and A. Carteni, A. 2014. “A Quality-Based Approach to Public Transportation Planning: Theory and a Case Study.” *International Journal of Sustainable Transportation*, 8(1).
- CEN/TC 320. 2002. “Transportation – Logistics and Services.” European Standard EN 13816: Public Passenger Transport – Service Quality Definition, Targeting and Measurement. European Committee for Standardization, Brussels.
- De Oña, J., and R. De Oña. 2015. “Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A Review and Assessment of Methodological Approaches.” *Transportation Science*, 49(3): 605-622.
- De Oña J., R. De Oña, L. Eboli, C. Forciniti, and G. Mazzulla. 2016. “Transit Passengers’ Behavioural Intentions: The Influence of Service Quality and Customer Satisfaction.” *Transportmetrica A: Transport Science* 12(5): 385-412
- De Oña, J., R. De Oña, L Eboli, and G. Mazzulla. 2015. “Heterogeneity in Perceptions of Service Quality among Groups of Railway Passengers.” *International Journal of Sustainable Transportation*, 9(8): 612-626.
- De Oña, J., G. López, R. Mujalli, and F. J. Calvo. 2013. “Analysis of Traffic Accidents on Rural Highways using Latent Class Clustering and Bayesian Networks.” *Accident Analysis & Prevention*, 51: 1-10.
- Dell’Olio L., A. Ibeas, and P. Cecin. 2010. “Modelling User Perception of Bus Transit Quality.” *Transport Policy*, 17(6): 388-397.

- Depaire, B., G. Wets, and K. Vanhoof. 2008. "Traffic Accident Segmentation by Means of Latent Class Clustering." *Accident Analysis and Prevention*, 40(4): 1257-1266.
- DeSarbo, W., and V. Mahajan. 1984. "Constrained Classification: The Use of *a Priori* Information in Cluster Analysis." *Psychometrika*, 49(2): 187-215.
- Eboli, L., and G. Mazzulla. 2008. "A Stated Preference Experiment for Measuring Service Quality in Public Transport." *Transportation Planning Tech.*, 31(5): 509-523.
- Eboli, L., and G. Mazzulla. 2011. "A Methodology for Evaluating Transit Service Quality Based on Subjective and Objective Measures from the Passenger's Point of View." *Transport Policy*, 18: 172-181.
- Eboli, L., and G. Mazzulla. 2012. "Performance Indicators for an Objective Measure of Public Transport Service Quality." *European Transport*, 51: 1-21.
- Fraley, C., and A. E. Raftery. 1998. "How Many Clusters? Which Clustering Method? Answers via Model-Based Cluster Analysis." *The Computer Journal*, 41(8): 578-588.
- Hair, J. F., W. C. Black, B. J. Babin, and R. E. Anderson. 2010. "Multivariate Data Analysis: A Global Perspective." In *P. Education Ed., 7th Edition*. New Jersey: Prentice Hall
- Magidson, J., and J. Vermunt. 2002. "Latent Class Models for Clustering: A Comparison with K-means." *Canadian Journal of Marketing Research*, 20(1): 37-44.
- Nakanishi, Y. 2003. "A Guidebook for Developing a Transit Performance Measurement System." TCRP Report 88, National Academy Press, Transportation Research Board, Washington, DC.
- Nathanail, E. 2008. "Measuring the Quality of Service for Passengers on the Hellenic Railways." *Transportation Research Part A*, 42: 48-66.
- Paez, A. 2006. "Exploring Contextual Variations in Land Use and Transport Analysis using a Probit Model with Geographical Weights." *Journal of Transport Geography*, 14(3): 167-176.
- Parasuraman, A., V. A. Zeithaml, and L. L. Berry. 1985. "A Conceptual Model of Service Quality and Its Implications for Future Research." *Journal of Marketing*, 49(4): 41-50.
- Prebensen, N. K., 2005. "Country as Destination—Norwegian Tourists' Perceptions and Motivation." *Journal of Hospitality & Leisure Marketing*, 12: 63-85.
- Raftery, A. E., 1986. "A Note on Bayes Factors for Log-Linear Contingency Table Models with Vague Prior Information." *Journal of the Royal Statistical Society, Series B*, 48: 249-250.
- Reyna, C., and S. Brussino, S. 2011. "Revisión de los fundamentos del análisis de clases latentes y ejemplo de aplicación en el área de las adicciones" ("Review of the Fundamentals of Latent Class Analysis and Application Example in the Area of Addictions"). *Trastornos adictivos (Addictive Disorders)*, 13(1): 11-19.
- Rondán-Cataluña, F., A. F. Villarejo-Ramos, and M. J. Sánchez-Franco. 2007. "Segmentación por clases latentes en servicios tecnológicos: una investigación sobre criterios y modelos para segmentar" ("Latent Class Segmentation Technology:

Research on Criteria and Models to Segment”). Paper presented at XIX Encuentro de Profesores Universitarios de Marketing Vigo, September.

- Shiftan, Y., M. L. Outwater, and Y. S. Zhou. 2008. “Transit Market Research using Structural Equation Modelling and Attitudinal Market Segmentation.” *Transport Policy*, 15: 186-195.
- Trompet, M., X. Liu, and D. Graham. 2011. “Development of Key Performance Indicators to Compare Regularity of Service between Urban Bus Operators.” *Transportation Research Record*, 2216: 33-41.
- Tyrinopoulos, Y., and G. Aifadopoulou. 2008. “A Complete Methodology for the Quality Control of Passenger Services in the Public Transport Business.” *European Transport*, 38: 1-16.
- Federal Transit Administration, Transit Development Corporation, Morpace International, Transit Cooperative Research Program, and Cambridge Systematics. 1999. *A Handbook for Measuring Customer Satisfaction and Service Quality*. National Academy Press, Transportation Research Board, Washington, DC.
- Vermunt, J. K., and J. Magidson. 2002. “Latent Class Cluster Analysis.” In Hagenaars, J. A., and A. L. McCutcheon, eds., *Applied Latent Class Analysis*. Cambridge: Cambridge University Press, 89-106.
- Vermunt, J. K., and J. Magidson. 2005. *Latent Gold 4.0 User's Guide*. Belmont, MA: Statistical Innovations, Inc.
- Wen, C. H., and S. C. Lai. 2010. “Latent Class Models of International Air Carrier Choice.” *Transportation Research Part E*, 46(2): 211-221.
- Yeh, C. H., H. Deng, and Y. H. Chang. 2000. “Fuzzy Multicriteria Analysis for Performance Evaluation of Bus Companies.” *European Journal of Operations Research*, 126(3): 459-473.
- Zhou, Y., K. Viswanathan, Y. Popuri, and K. E. Proussaloglou. 2004. “Transit District Customers in San Mateo County, California.” *Transportation Research Record*, 1887: 183-192.

About the Authors

JUAN DE OÑA (*jdona@ugr.es*) is with the TRYSE Research Group at the University of Granada (Spain) Department of Civil Engineering.

ROCIO DE OÑA (*rociadona@ugr.es*) is with the TRYSE Research Group at the University of Granada (Spain) Department of Civil Engineering.

FRANCISCO DIEZ-MESA (*franmesa@ugr.es*) is with the TRYSE Research Group at the University of Granada (Spain) Department of Civil Engineering.

LAURA EBOLI (*laura.eboli@unical.it*) is with the University of Calabria (Italy) Department of Civil Engineering.

GABRIELLA MAZZULLA (*gabriella.mazzulla@unical.it*) is with the University of Calabria (Italy) Department of Civil Engineering.