Heterogeneity in Perceptions of Service Quality among Groups of Railway Passengers

By: Juan de Oña, Rocío de Oña, Laura Eboli & Gabriella Mazzulla

Juan de Oña, Rocío de Oña, Laura Eboli, & Gabriella Mazzulla (2015) *Heterogeneity in Perceptions of Service Quality among Groups of Railway Passengers*. **International Journal of Sustainable Transportation, 9(8), 612-626** DOI: 10.1080/15568318.2013.849318

Direct access to the published version:
http://dx.doi.org/10.1080/15568318.2013.849318

HETEROGENEITY IN PERCEPTIONS OF SERVICE QUALITY AMONG GROUPS OF RAILWAY PASSENGERS

Juan de Oña
TRYSE Research Group
University of Granada
Department of Civil Engineering
Severo Ochoa s/n
Granada 18071, Spain
ph. +34.958.244979
fax+34.958.246138
jdona@ugr.es

Rocio de Oña
TRYSE Research Group
University of Granada
Department of Civil Engineering
Severo Ochoa s/n
Granada 18071, Spain
ph. +34.958.244950
fax+34.958.246138
rociodona@ugr.es

Laura Eboli**
University of Calabria
Department of Civil Engineering
P. Bucci, cubo 46/B
87036 Rende (CS), Italy
ph. +39.984.496784
fax +39.984.496787
laura.eboli@unical.it

Gabriella Mazzulla
University of Calabria
Department of Civil Engineering
P. Bucci, cubo 46/B
87036 Rende (CS), Italy
ph. +39.984.496782
fax +39.984.496787
gabriella.mazzulla@unical.it

**Corresponding Author

ABSTRACT

The aim of this work is analysing the different perceptions among groups of users on transit service quality. We propose a methodology based on a classification and regression tree approach (CART) allowing the characteristics mostly influencing overall service quality to be identified. The methodology is applied by using data regarding a rail service of the Northern Italy, particularly suburban lines connecting different towns of the hinterland of the city of Milan. Passengers expressed their opinions about service characteristics such as safety, cleanliness, comfort, information, personnel. We found that perceptions about service quality are differentiated among the various groups of users.

1. INTRODUCTION

Rail transport experienced a rapid decline at the second half of the 20th century, mainly due to the developments in road and air transport. At the same time, as privatization of rail operation gains pace, commercial considerations also act to promote the use of rail (Brons et al., 2009). The EU Transport White Paper explicitly stated that "rail transport is literally the strategic sector, on which the success of the efforts to shift the balance (from private to public modes) will depend" (CEC, 2001). To achieve this aim, the rail services should offer adequate standards of quality to the users.

An attempt to improve service quality and, consequently, increase the use of the services could be represented by the examination of the customers' experiences on trains and at stations and the definition of their perceptions and expectations by conducting passenger satisfaction surveys. Developing an understanding of customer service needs is essential for gaining more ridership and a modal shift in favour of the rail (Bruhn and Grund, 2000).

The quality of a rail service is influenced by the quality of several characteristics. These characteristics concern different aspects: accessibility, safety, comfort, service, information, personnel, etc. So, in order to understand users' perceptions of the service, it is fundamental to collect users' perceptions about different characteristics.

Customers' and journey's characteristics influence perceptions of the different aspects describing rail service quality. Identification of different passenger requirements and service expectations for different types of travellers (see for example, MVA, 2007), in addition to research into personal characteristics of users (Evans and Wener, 2007) and passenger interaction (Harris and Baron, 2004), demonstrates the heterogeneity of passengers. Passenger expectations of service aspects also vary across different regions (MVA, 2007). Therefore, the heterogeneity presented in passengers' opinions when they are evaluating service quality is an inherent characteristic of this measure. In fact, Stathopoulos and Marcucci (2014) provided an in-depth discussion about this issue in their research.

To analyse this heterogeneity, one possibility is to stratify the sample of users and then build specific models, or another possibility is to use methodologies that can consider this heterogeneity of perceptions. For example, Bordagaray et al. (2013) used random ordered probit models for analysing service quality on a bus transit service, and for identifying variations in tastes according to users' socioeconomic and journey characteristics. In these models, parameters are not constants estimated, and systematic variations around the mean have been calibrated to account for the effect of these characteristics on the perception.

Regarding to those that have analyzed heterogeneity through segmentation of the sample, not many studies have investigated the differences in users' perceptions based on the travel habits profiles. There is a study which investigates these kinds of heterogeneity (Rail Corporation - NSW Government and City Rail, 2009) showing that people travelling to or from work are more likely to prioritise improvements in the frequency of services and reductions in travel time than those who are travelling for other purposes. There are also interactions between the length of the journey, the purpose of the trip, the frequency of train use and the season of the year; all of which impact on satisfaction ratings (Passenger Focus, 2007). Starting from these considerations about the differences among users, the aim of this paper is to investigate which are the main differences in the attributes that most influence passengers when they are making an overall evaluation about a public rail service using a novel and powerful methodology for the public transport industry. The use of a classification and regression tree approach (CART) (Breiman et

al., 1984) is the methodology suggested to investigate about the perceptions of different groups of users on the services offered by a rail operator of the North of Italy, with the final aim to identify the characteristics mostly influencing the overall service quality perceived by different types of users. The users' profiles are distinguished according to four criteria: the type of user in terms of the purpose of the trip by considering commuters travelling for working, commuters travelling for studying, and people travelling for other purposes; the type of the day of the journey, by distinguishing people travelling during the week, in the days before a holiday, and during holidays; the frequency of the use of the service, which allows us to observe the difference between users travelling daily, users travelling weekly, and people travelling occasionally; the time of the day, according to which travellers were classified as users travelling in the off-peak hours, in the morning peak hours, afternoon peak hours, and evening peak hours.
In the following, there is a literature review of works concerning the analysis of the quality of rail service; the review has the aim to explore how researchers analysed rail service quality and also which factors were investigated. After the literature review, a section about CART methodology is proposed. Before a brief theoretical background of CART, a brief review of the methodologies for analyzing customer satisfaction in public transport is proposed in order to introduce the CART methodology as an alternative way of dealing with transit service quality. Then, the experimental context is presented: the survey conducted for collecting the data is described, the characteristics of the sample of interviewed users and the importance and satisfaction rates collected during the survey are analysed. After the section about the case study, the application of the CART methodology to the experimental data and the results obtained from this application are discussed. The paper ends with a general concluding section.

2. LITERATURE REVIEW

Service quality of rail services was investigated by many researchers. However, most studies about public transport service quality evaluation are based on the analysis of bus services (see for example Cirillo et al., 2011; dell'Olio et al. 2010; de Oña, de Oña, Calvo, 2012; de Oña et al., 2013; Eboli and Mazzulla, 2008, 2010, 2011, 2012a; Joewono and Kubota, 2007; Nurul-Habib et al., 2011). In this work, we focus the review only to studies based on railway services in order to direct the investigation to a specific type of public transport service. In the following, we reported a literature review of some studies conducted for evaluating the quality level of rail service and especially for identifying the characteristics mostly influencing the overall service quality.
Cavana et al. (2007) proposed an extension of the well known SERVQUAL instrument (Parasuraman et al., 1988) to evaluate passenger rail service quality by adding three new transport dimensions (comfort, connection, and convenience). They found, by a regression analysis, that assurance, responsiveness and empathy had significant effects on overall service quality, but also reliability and convenience.
The objective of Nathanail (2008) was to present a framework developed for assisting railway operators in monitoring the quality of services provided to their passengers. This framework is based on the estimation of some indicators concerning itinerary accuracy, system safety, passenger comfort, and so on. Itinerary accuracy and system safety have been attributed the highest grades.
Brons et al. (2009) evaluated how important the 'access-to-the-station' part of a rail journey is to

passengers in their "Overall Satisfaction" with the rail journey, by applying principal component analysis and derived importance techniques; they explained the propensity to use rail through a regression analysis. They found that satisfaction with the quality of the access to the station is an important dimension of the rail journey influencing the "Overall Satisfaction".

Cantwell et al. (2009) proposed a multinomial logit model which revealed that passengers of a rail service would derive a benefit from an improvement in service reliability and a reduction in crowding.

The study of Rahaman and Rahaman (2009) aimed to focus on the railway transportation sector and to develop a model defining the relationship between "Overall Satisfaction" and service quality attributes. They found that overall service satisfaction depends on eight distinct service quality attributes; the quality of security inside the train mostly dominates service satisfaction.

Geetika (2010) identified the factors that determine user satisfaction with the quality of services provided on railway platforms. Determinants identified through a factor analysis are availability and quality of refreshments, effectiveness of information systems, behaviour of railway staff, basic amenities provided on platforms, and safety and security. Refreshments and behavioural factors are considered the most important by passengers.

Prasad and Shekhar (2010) evaluated the quality of a rail service by developing a Service Quality Management model on the basis of SERVQUAL. Three new dimensions (service product, social responsibility and service delivery) are added to the original five SERVQUAL dimensions (i.e. assurance, empathy, reliability, responsiveness and tangibles). Service delivery and social responsibility are identified as the most important and the least important factor, respectively.

The objective of the research of Saputra (2010) was to find the factors having significant influences on customer satisfaction towards a rail service, by identifying the differences among groups of users. The author identified six factors having significant influences on customer satisfaction for commuter class: information, appearance, service coverage, tangible aspects, safety and security, and cost; seven factors for business class: travel time, information, scheduling, comfort, tangible, safety and security, and service coverage; and also seven factors for executive class: appearance, safety and security, information, comfort, tangible aspects, travel time, and cost. In addition, while the commuter class passengers complain on the safety and security attribute, for business class it is the information provision that influences the desire to complain.

Agunloye and Oduwaye (2011) evaluated the relationships among factors such as arrival time of train, smoothness of ride and cleanliness of the coaches. The study, proposing a statistical analysis of the judgments of satisfaction expressed by the passengers, revealed that the service was ineffective and inadequate, and that only the arrival time of trains at stations has a significant relationship with the passengers' trip frequency.

Chou et al. (2011) evaluated how the incorporation of the Quality-Satisfaction-Loyalty relationship (Chou and Kim, 2009) into a passenger satisfaction index (PSI) calculation can be used to assess the quality of a high-speed rail service. The empirical study concluded that level of access to a station and personal space on train are the top-priority quality indicators that need to be addressed to improve customer satisfaction and profits. To gain further insight into the perceptions by different parties of service indicators, PSI by social-demographic group (e.g., age, gender, education level, occupation, marital status, etc.) was analyzed. As an example, female passengers over 30 years old, passengers with an education level less than that of high school, married passengers, passengers with a monthly income between 40,001 and 70,000

Euros, those with a household size less than three, and passengers not in possession of driving licenses were the groups who were more satisfied with the services.
Irfan et al. (2012) investigated passengers' perceptions about the service quality of a rail transport system using a modified SERVQUAL instrument including eight service quality constructs (empathy, assurance, tangibles, timeliness, responsiveness, information system, food, and safety and security). Results indicated that there is a positive and significant relationship among the service quality constructs. Highest and significant correlation is among tangibles and empathy.
Eboli and Mazzulla (2012b) used a structural equation model (SEM) to explore the impact of the relationship between global customer satisfaction and service quality attributes, such as safety, cleanliness, main and additional services, information, and personnel. The main findings are that service characteristics like punctuality, regularity and frequency of runs, and cleanliness have the highest positive effect on service quality; while also comfort and information have a notable positive effect, personnel and safety have a not very considerable effect.

3. METHODS FOR DETERMINING FACTORS INFLUENCING SERVICE QUALITY

3.1 PRELIMINARY REMARKS

Starting from the studies mentioned in the previous section and by considering also the studies oriented to the evaluation of other types of public transport (e.g. bus services) we can conclude that transit service quality has been measured by different approaches based on transit users' opinions. There are methods based on traditional customer satisfaction surveys, in which users express their opinions by rating the various service characteristics (e.g. Cavana et al., 2007; Eboli and Mazzulla, 2007, 2009; dell'Olio et al., 2010; Hensher et al., 2010; Jen et al., 2011; Nurul-Habib et al., 2011; Pakdil and Aydin, 2007), and methods based on stated preference surveys, in which the importance given by the users to the service attributes is indirectly derived by means of exercises of choice based on the stated preferences (Cirillo et al., 2011; Eboli and Mazzulla, 2008, 2010; Hensher, 2014; Hensher and Prioni, 2002; Hensher et al., 2003; dell'Olio et al., 2011). Most of these models have their own assumptions and pre-defined underlying relationships between dependent and independent variables. If these assumptions are violated, the model could lead to erroneous estimations of the likelihood of quality of service (de Oña, de Oña, Calvo, 2012). In this context, the approach based on a classification and regression tree (CART) can solve these inconveniences, being a non-parametric model with no pre-defined underlying relationship between the target (dependent) variable and the predictors (independent variables). Moreover, looking from the point of view of transport practitioners, CART methodology can provide more useful and informative results than other approaches because of the "If-then" rules generated from the trees. Every terminal node produces a rule which provides a better understanding about the phenomenon being analyzed. Decision rules provide the thresholds that a set of attributes should pass for reaching a good "Overall Satisfaction" with the service, or on the contrary, the thresholds that a set of attributes should pass for not obtaining a poor "Overall Satisfaction" with the service. Practitioners can decide the tool they want to use, according to the available resources. Likewise, this methodology provides results that are simple

and easy for understanding and interpreting by non-statisticians and non-academical users, due to its graphical representation. Also, it should be highlighted that multi-collinearity among attributes and the existence of outliers (characteristics defining passengers opinions' data) are not a problem for this methodology.
A detailed description of CART analysis and its applications can be found in Breiman et al. (1984). In the field of transportation, decision trees have been widely applied in road safety analysis. For example, Abdel-Aty et al. (2005) analyzed the main factors that affect different types of crashes at signalized intersections. Chang and Wang (2006) studied the relationships between crash severity and several characteristics related to drivers, vehicles, roads and the environment. Likewise, the specific influence of driver' characteristics into crash severity of an accident was investigated by Pakgohar et al. (2011). This methodology has also been used for analyzing other aspects of traffic engineering, such as traffic forecasting (Washington and Wolf, 1997), vehicle emissions (Washington et al., 1997; Hallmark et al., 2002) or transportation choice behaviour (Xie et al., 2003).
Some authors have tested decision trees results with other methodologies. For example, Xie et al. (2003) used decision trees, artificial neural networks (ANN) and multinomial logit models for predicting work travel mode choice. The results showed that the three methodologies identified the same attributes as the most significant for the travel choice, and that the two data mining models (decision trees and ANN) offered better performances than the multinomial logit model. Pakgohar et al. (2011) used decision trees and multinomial logistic regression for analyzing crash severity, and they discovered that decision trees obtained more precise predictions and were also simpler and easier to interpret.
Decision trees have been also adopted for analyzing quality of service in industries different from transportation. Huang and Hsueh (2010) analysed the overall service quality satisfaction on a refurbishment industry considering 22 items describing the service. In the field of public transportation, decision trees are relatively new. Wong and Chung (2007) applied this methodology to the sector of air transportation. They investigated the relationships that existed between valuable and non-valuable passengers of an airline service with various service quality attributes. Concerning the analysis of service quality for transit service, this methodology was applied just twice: the first time by de Oña, de Oña, Calvo (2012), who adopted CART for identifying the key factors affecting the quality level of a bus service operating in Granada; and the second time by de Oña, Eboli, Mazzulla (2012), that applied CART to a rail service in the North of Italy.

3.2 CART METHODOLOGY

Decision trees are a data mining technique used to classify and predict a class variable. The CART method can be used for both discrete and continuous target variables. When the value of the target variable is discrete, a classification tree is developed, whereas a regression tree is developed for a continuous target variable. In this study, we used a classification tree because the target variable is discrete (the target is passengers' "Overall Satisfaction" with three levels: Poor, Fair and Good), as we can see in the following.
The development of a CART model begins with all the data concentrated in the root node, which is the node located at the top of the tree (e.g. see Figure 1 in which the 7,333 observations are in the root node). This root node is divided into two child nodes (the internal nodes created under an upper node) on the basis of an independent variable (splitter) that maximizes the "purity" of

the child nodes. In this case, the independent variables used in the model as possible splitters are the 27 different service characteristics (see Table 2). Then, each child node is recursively split until all of them are pure (all the cases are of the same class: Poor, Fair and Good), their "purity" cannot be increased or a stopping criteria has been satisfied. For example, Montella et al. (2012) fixed as stopping criterion the maximum size of the tree (four levels) in order to find shorter and therefore more useful decision rules, with few variables participating in the rule.
. The most famous splitting index is the Gini Index, which measures the impurity of the node. The impurity measure at a node *t*, *I (t)*, may be defined as follows:

$$I(t) = 1 - \sum_{i=1}^{J} \left(\frac{n_i}{n}\right)^2 \quad (1)$$

in which J is the number of classes in the target variable, n_i is the number of cases belonging to the class i, and n is the total number of cases. If a node is "pure", all the observations in the node belong to one class, and the Gini Index or $I(t)$ will be equal to zero. We find maximal impurity at a node, and therefore maximal Gini index, when all the classes are equally likely.
Then, we can define the split criterion based on the Gini Index as the Gini reduction criterion, which measures the "worth" of each split in terms of its contribution toward maximizing the homogeneity through the resulting split. A set of candidate split rules is evaluated and ranked during the tree growth. If a split results in splitting one parent node into B branches, the "worth" of that split may be measured as follows:

$$Worth = I\,(P) - \sum_{b=1}^{B} P(b) * I(b) \quad (2)$$

where $I(P)$ denotes the Impurity of the parent node, $P(b)$ denotes the proportion of observations in the node assigned to a branch b and $I(b)$ denotes the impurity of the node b.
Following this procedure, the maximal tree overfitting the data is created. To decrease its complexity and create simpler trees, pruning is realized according to a cost-complexity algorithm, which is based on removing the branches adding little to the predictive value of the tree. After pruning a branch, if the increase in the misclassification cost is sufficiently lower than the decrease in the complexity cost, the branch will be pruned, and a new tree will be created.
The last step is to select an optimal tree from the pruned trees. Using the misclassification cost on the testing dataset (or an independent dataset), the optimal tree is the one having the least misclassification cost.
One of the methods used for validating the predictive accuracy of the tree model is the k-fold cross validation. Normally, the validation method consists on taking the complete dataset used for the analysis and dividing it into two subsets of data: the training data and the testing data. The training data is used for learning the sample, splitting the nodes and growing the tree, while the testing data is used to extract the precision rate of the model. However, the k-fold cross validation is the most practical validation method in limited data situations (Witten and Frank, 2005). At this, you decide a fixed number of folds (k), or partitions of data. It uses the whole dataset, and randomly divides the sample into k subsets. Sequentially, each subset is used as a testing set, for the tree model generated by the remainder k-1 subsets that have been used for training. That is to use k-1 folds for training and one fold for testing, and repeat the procedure k times so that, at the end, every instance has been used exactly once for testing. Thus, different k models are obtained, in which the accuracy of the classification is calculated. Finally the k accuracy indicators are averaged to yield an overall accuracy value.

Compared to the other modeling methods, CART has the advantage that it is presented as easily understandable visual branching images providing effective "If-Then" rules. Every leaf of the decision tree corresponds to a decision rule that extracts very useful information about the data. It is a logic conditional structure starting in the root node with "If", continues with every variable that takes part in the tree growing making an "If" of the rule, and ends in the child nodes with "Then", in which is associated the class of the target variable that shows the highest number of cases in the child node analyzed. The method has some advantages compared to models based on choice probability because choice probability cannot be calculated analytically since the integral does not have a general closed form and the exact maximum likelihood estimation is not possible. In fact, the probability is usually approximated through simulation (Bhat, 2001, 2003; Bastin et al., 2006). Most of these kind of models have their own model assumptions and pre-defined underlying relationships between dependent and independent variables; so, if these assumptions are violated, erroneous estimations could be obtained. Non-parametric models, like CART methodology, with no pre-defined underlying relationship can solve these limitations. Another valuable outcome provided by CART analysis is the value of the standardized importance of independent variables, which reflects the impact of such predictor variables on the model.

4. EXPERIMENTAL CONTEXT

The service quality of a railway service operating in the North of Italy, and specifically in the city of Milan, was investigated in this research work. We analysed the services offered by 9 suburban lines connecting towns of the hinterland of Milan. The lines are used by about 200,000 passengers per day. As many as 7,333 users (sample rate of about 4%) were interviewed in the month of May 2012. The interviews were conducted on board during one week in a time slot between 6.00 a.m. and 10.00 p.m.

The questionnaire was structured into two main sections. In the first section, data concerning general information (e.g. time period of the interview, train, line, station, and operator), socio-economic characteristics (e.g. gender, age, qualification, professional condition, and income), and travel habits (e.g. trip scope and frequency, and type of ticket) were collected. The second section is specific to passengers' perceptions of the used services; users expressed importance and satisfaction rates, on a cardinal scale from 1 to 10, about 27 service quality factors concerning safety, cleanliness, comfort, service, information, personnel and other. Other authors have used the same scale in their studies, such us Huse and Evangelho (2007), Oyewole (2001), Chou and Kim (2009), Chou et al. (2011) and Paquette et al. (2012).

The sample is made up of more of females (54.2%) (see Table 1). Most of the passengers are aged between 16 and 25 (40.5%), and another considerable part is represented by people aged between 26 and 40 (33.1%). The major part of the sampled people are students (38.1%) and employees (37.7%). More than half of the sample had obtained a diploma of upper secondary school, and almost 30% has a university degree. More than one fifth of the sample did not give any kind of information about the income, while about 40.0% has not a fixed income; people stating their monthly income mainly belong to a class between 1,001 and 1,500 Euros. Passengers travel by train mainly for reaching the place of work or study (73.2%). Most of the sample travels by train every day (61.9%), but about 24% of passengers travel occasionally.

People mainly purchase a travel card (74.7%), and about 25% travel using a one-way ticket. Commuters workers (32.3%) and commuters students (27.3%) represents more than a half of the sample. Most part of the sample travel in a weekday (84.6%) and in the off-peak hours (30.7%).

Table 1 Sample characteristics

Table 2 shows the importance and satisfaction rates expressed by the passengers, as well as a coded named for the 27 items used in the survey. They judged most of the attributes (14 out of 27) as very important (showing an average rate of importance around 8 and 9 out of 10); the attributes considered as the most important are the three attributes concerning travel safety. The attributes considered relatively less important are: bicycle transport on board (6.00), parking (7.04), info connection with public transport (7.33), and complaints (7.48). On the other hand, the average satisfaction rates suggest that people are not very satisfied with the service (average rate of satisfaction on the overall service of 5.81), in fact only nine attributes out of 27 have an average rate higher than 6. The service characteristics considered as the most satisfactory safety and personnel; all the other characteristics are judged as not satisfactory.

Table 2 Importance and satisfaction rates

5. APPLICATION OF THE CART

We applied the CART methodology to the rail service described above by using WEKA software (available at: http://www.cs.waikato.ac.nz/ml/weka/). We have considered different groups of users according to four criteria: the type of the user depending on the type of user, the type of the day of the journey, the frequency of the use of the service and the time of the day of the journey. These criteria were selected for investigating on the main differences on passengers' perceptions according to their travel habits profiles, due to not many studies have researched on this issue, and strong reflections could be extracted.

So, in the following the results of different applications will be described. We firstly applied the CART to all the 7,333 passengers of the analysed suburban lines. Afterwards, we applied the CART to the users classified according to the four different criteria.

For all the models, 27 attributes describing the service were used as independent variables. To find out more applicable decision rules, the target variable (overall SQ) and the independent variables were re-coded in a reduced semantic scale. It was a three semantic scale comprising the rates from 1 to 4 as POOR, from 5 to 7 as FAIR, and from 8 to 10 as GOOD[1].

For all the groups, CART used a 10-fold cross-validation of the sample for obtaining the precision ratio (Witten and Frank, 2005) of the categorization of the variable class. All the generated trees obtained high rates, from about 76% to 79%.

The precision rate for the global sample of the suburban service is 78.15% (Table 3). For the global sample, the tree produced 5 levels, 19 nodes and 10 terminal nodes. The variable that

[1] We would like to point out here, that if we had used another recodification of the variables, it would be possible that the trees have changed. However, we believe that the recodification used here is reasonable from the practitioners and transport planners point of view.

splits the root node and obtains the maximum purity of the two child nodes (this split achieves the best “worth”) is FARE (see Figure 1).

Figure 1 CART for suburban lines

On the left branch of the tree there are the passengers having a POOR satisfaction with FARE; it is predicted five terminal nodes (7, 8, 10, 13, and 14), where the “Overall Satisfaction” will be POOR or FAIR. This implies that when a passenger has a POOR satisfaction with FARE, his/her “Overall Satisfaction” with the service probably will never be GOOD. Under Node 1, the tree grows according to different splitters as INFO STATION, REGULARITY, CLEAN SEATS, and SECUR BOARD.

On the right branch, there are the passengers having a satisfaction higher than POOR with FARE. In this case, there are 5 terminal nodes (5, 11, 16, 17 and 18) and their prediction of the “Overall Satisfaction” is FAIR or GOOD. If REGULARITY and FARE are considered with a satisfaction higher than FAIR, the probability to have a GOOD “Overall Satisfaction” increases with regard to the previous parent nodes (Node 12, 55.3% of probability), and two terminal nodes classified as GOOD can be achieved (Node 16 and Node 18, with probabilities of 67.5% and 55.9% respectively).

Table 3 Importance for users of suburban lines and for users classified according to the four criteria (type of user; day of the trip; frequency of the trip; time of the trip)

The main key factors for the global sample are REGULARITY, PUNCTUALITY and CLEAN VEHICLES (see Table 3). These results agree with the study of Eboli and Mazzulla (2012b), in which a Structural Equation Model (SEM) was applied to the same service (survey carried out in 2011) for analysing rail service quality. SEM is a well-known parametric methodology combining regression, factor analysis and analysis of variance (Bollen, 1989) and it has been one of the most popular techniques applied for analyzing service quality in public transportation, due to the possibility of considering both unobserved latent construct and observed indicators, and the relationship among them is revealed. The factors identified in this paper (Eboli and Mazzulla, 2012b) as the most influencing passengers’ “Overall Satisfaction” were those related with Service and Cleanliness, and particularly, regularity and punctuality of the runs, and cleanliness of vehicles and seats. It should be highlighted that both methodologies (decision trees and SEM) have identified the same attributes as the most important for the rail overall service quality, testing the reliability and validity of the approach presented in this paper.

Three different trees were built according to the type of user: “Commuter Workers”, “Commuter Students” and “Others”. The precision achieved in each of these trees is high and similar to the one obtained with the global sample (see Table 3).

Figure 2 CART for users classified according to the type of user (“Commuter Workers”)

The variable splitting the root node is different in each one of the three cases. For “Commuter Workers” the splitter is FARE, as well as in the tree built using the total sample (see Figure 2); for “Commuter Students” the root node is divided into two child nodes based on the PUNCTUALITY, while the tree built for “Others” starts growing with WIND AND DOORS. As an example, in Figure 2, the tree derived for the “Commuter Workers” group is shown. This tree

generated 13 nodes and 7 terminal nodes. When FARE, INFO STATIONS, REGULARITY and CLEAN SEATS are perceived with a POOR satisfaction, the probability of having a POOR "Overall Satisfaction" is quite high (80.5% at Node 11). On the other hand, when FARE and COURTESY BOARD have a satisfaction higher than FAIR, also the "Overall Satisfaction" probably will be GOOD (Node 10). This could be explained because they frequently travel, so trip Fare takes a considerable amount of money of their monthly income, producing good satisfaction when it is good. Also, due to their frequency of use, they are who most suffer impolite or incompetence staff.
The most important attributes are also different among the three categories of users (Table 3). For "Commuter Workers" and "Others" the most important attribute is REGULARITY. For "Commuter Students", however, the most important attribute is FARE. For the category "Others" the attribute FREQUENCY is very important; however, this is not a relevant attribute for the other two groups, maybe because they know well the timetable of the service and worry about other aspects of the service.
The trees obtained by classifying the users according to the different type of day of the journey have high and similar precisions around 77% (Table 3). Figure 3 shows that FARE is the splitter for the root node of "Weekdays".

Figure 3 CART for users classified according to the day of the trip ("Weekdays")

It should be remarked that the group "Weekdays" includes most of the sample (84.6%), and the variable splitting the root node is also FARE in the tree built with the global sample (see Figure 1). This splitter divides the tree into two branches. On the left, there are the passengers with a POOR satisfaction with FARE, and 7 terminal nodes (7, 10, 14, 15, 16, 19 and 20) were produced predicting a POOR or FAIR "Overall Satisfaction". On the right branch, when FARE, REGULARITY and INFO BOARD have a satisfaction higher than FAIR, also the "Overall Satisfaction" will probably be GOOD. In this case, people travelling on weekdays comprise most part of the sample that travel for working or studies purposes, with a high frequency of use of the system. For this reason, good FARE, REGULARITY and INFO BOARD become essential for a good "Overall Satisfaction". INFO BOARD could be considered as not important for frequent passengers, but those passengers more than sporadic ones, are who really suffer from unexpected changes, delays, etc.
The tree built for the "Days before holiday" begins growing with COURTESY STATIONS. In this tree a small part of the sample has a GOOD "Overall Satisfaction" (only 8%) and no node predicts a GOOD satisfaction. For the group "Holidays", the tree starts splitting with FARE. This attribute has been also identified as the most important attribute (see Table 3).
Table 3 shows that the most important variables are different for each group. For "Weekdays", many variables have a great impact on the prediction of the "Overall Satisfaction", but the most important is REGULARITY, followed by PUNCTUALITY and COURTESY BOARD; people travelling in the weekdays give more importance, as expected, to the peculiar aspects of a transit service. For the group "Days before holiday" and also for "Holidays", few variables have importance. As an example, COURTESY STATIONS, COURTESY BOARD, and SAFETY have a strong impact for "Days before Holiday" group, which includes people not travelling every day for which more qualitative aspect could be important. Finally, for the group "Holidays" the most important variable is the FARE.

All the trees obtained by classifying the users according to the different frequency of the journey have high precision rates, higher than 77% (see Table 3). The variable splitting the root node is different in each case. For users travelling daily, the splitter is FARE, for users travelling weekly is REGULARITY, and for users who travel occasionally, the variable achieving the best classifications is COURTESY BOARD.

Figure 4 CART for users classified according to the frequency of the trip (“Daily”)

For the “Daily” group, which is the largest group (61.9% of the users), the tree generated 5 levels, 25 nodes and 13 terminal nodes (Figure 4). When FARE, and also INFO BOARD have a POOR satisfaction, the probability of having a POOR “Overall Satisfaction” increases with regard to the previous parent nodes (Node 3, 60.7% of probability). If COURTESY BOARD also has a POOR satisfaction, terminal Node 7 is created with a prediction of the “Overall Satisfaction” as POOR (80.6% of probability). Also if COURTESY BOARD is higher than POOR, but REGULARITY and WIND AND DOORS are POOR, the prediction of the “Overall Satisfaction” will be POOR (Node 19, 72.2% of probability). On the other hand, when FARE and COURTESY BOARD are higher than FAIR, the probability of having a GOOD “Overall Satisfaction” is higher with regard to the previous parent nodes (Node 12, 57.6% of probability).

For the “Weekly” group, a small tree was created with only 3 terminal nodes: one node predicts a POOR “Overall Satisfaction”, while the other two a FAIR “Overall Satisfaction”. This tree only splits according to the REGULARITY and COMPLAINTS. When these two variables are POOR, the “Overall Satisfaction” also will be POOR. The tree built for passengers occasionally travelling is different from the other two groups. In this case there are 5 terminal nodes. When passengers judge COURTESY BOARD as POOR, the probability of having a POOR “Overall Satisfaction” is high (55.8 % of probability); on the other hand, to increase the probability that users have a GOOD satisfaction with the service, not only COURTESY BOARD must be higher than POOR, but also FARE and CLEAN VEHICLES should be higher than FAIR.

For users travelling daily, the most important variables are linked to both more qualitative and less qualitative service aspects (COURTESY BOARD, REGULARITY and INFO BOARD) (see Table 3). For users travelling weekly the most important variable is INFO STATIONS, followed by WIND AND DOORS and REGULARITY; for this kind of users the aspects linked to the information provided to users are more relevant because they have less knowledge of the service than the habitual users and need information for travelling. However, passengers travelling occasionally only focus on the COURTESY BOARD, COURTESY STATIONS and FARE; this group of users prefers to make sure that they have a good treatment from the personnel while travelling and they are less interested in other aspects of the service.

Finally, four different trees were built according to the time of the trip: “Off-peak hour”, “Morning peak hour”, “Afternoon peak hour”, and “Evening peak hour”. The precision rates of the trees have high and similar values, over 75.9% (Table 3).

For “Off-peak hour” and “Afternoon peak hour” groups, the first splitter is FARE. For “Morning peak hour” group the first splitter is AIR CONDIT, and for the “Evening peak hour” group is COURTESY BOARD.

Figure 5 CART for users classified according to the time of the trip (“Afternoon peak hour”)

The tree built for “Off-peak hour” divides on the one side the passengers that have a POOR satisfaction with FARE and “Overall Satisfaction” will not be GOOD, and on the other side those with a FAIR or GOOD satisfaction with FARE, and “Overall Satisfaction” will be FAIR or higher. From the tree for “Morning Peak hour”, if AIR CONDIT is higher than POOR, and FARE and BICYCLE BOARD have a satisfaction higher than FAIR, the probability of having a GOOD “Overall Satisfaction” is high (78.1%). The tree built for the “Evening peak hour” group only created two terminal nodes; when COURTESY BOARD is considered as POOR, also the “Overall Satisfaction” will be POOR, with a probability of 64.8%.

The “Afternoon peak hour” group produces a tree with 17 nodes and 9 terminal nodes (see Figure 5). The root node is split based on FARE. When FARE is POOR and also INFO BOARD is POOR, the probability of having a POOR “Overall Satisfaction” increases with regard to the previous parent nodes (Node 3, 59.6% of probability), and depending on other variables as splitters (as COURTESY STATIONS, FREQUENCY, and CLEAN VEHICLES) the terminal nodes are created. On the other side, the probability of having GOOD “Overall Satisfaction” increases with regard to the previous parent nodes (Node 10, 59.4% of probability) when FARE is considered higher than POOR, and PUNCTUALITY and PRICE INTEGR have a satisfaction higher than FAIR.

COURTESY BOARD is the most important variable for “Morning peak hour” and “Evening peak hour” groups. FARE has a high importance in the “Morning peak hour”. In the “Off-peak hour” REGULARITY is the most important because in that period of the day the number of runs is not very high and people need a more regular service; finally, in the “Afternoon peak hour” the FREQUENCY is in the first position.

6. **CONCLUSIONS**

In this paper, we demonstrated that CART methodology can be very useful for analysing transit service quality and for verifying the main differences in users’ perceptions about the services. The main advantage of this methodology is the high usefulness of its results for transport practitioners: it provides more informative outcomes than other most popular techniques widely applied for analysing service quality in public transportation; the model results are simple and easily understable (due to the graphical representation); they provide powerful “If-then” rules, that could help transport planners to formulate adequate strategies to manage their budget and resources limitations; the most important variables for predicting the “Overall Satisfaction” are easily found, etc.

For this reason, we applied CART to different groups of users of a suburban rail service of the North of Italy, classified according to the type of users, the day of travel, the frequency of use, and the time of travelling during the day. The service aspects mostly influencing the overall service quality at each of these groups and for the overall sample, were identified and the overall quality of the service was predicted with a high precision rate (from about 76% to 79%).

The factors retained as most important by the overall sample were the factors linked to regularity and punctuality of the runs. However, also some more qualitative aspects have a high influence for the users (e.g. cleanliness of vehicle and the seats, windows and doors working, information on board, complaints, courtesy and competence on board). Similar results were found by Eboli and Mazzulla (2012b), in which, by applying a structural equation model for analyzing service quality in the same railway service, similar variables were identified as having the highest influence over the “Overall Satisfaction”. Likewise, the trustworthiness of decision trees has been proved by different authors (e.g. Xie et al., 2003), who highlighted that decision trees had the ability to identify the most important independent variables of the model (as did other most popular methodologies), and that they achieved good performances of the model (high precision rates).

By analysing different groups of users, we found that passengers have different preferences based on their belonging to a specific group; each group considers as important different service aspects. The variable that splits the root node in most of the cases is FARE. This is the splitter variable for the case in which we considered the whole sample, but it is also the splitter variable of at least one group regarding each criterion of classification, usually for the largest groups inside of each criterion because it is the closest to the whole sample. The trees help to identify the variables mostly influencing service quality and to predict the level of "Overall Satisfaction" starting from the satisfaction of the users with each one of the service characteristics.On the other hand, there is a set of variables (CLEAN TOILET, CROWDING, LOCALIZAT and PARKING) that have not been used as splitters for building the trees across the different segments of users, and also they have not been identified as key factors for their "Overall Satisfaction". So, it can be considered that these variables have a low influence when a passenger is evaluating the railway service.

Another interesting finding is that the factors identified by the models as the most important for the travellers (REGULARITY, PUNCTUALITY and CLEAN VEH), in the case of the global sample (see Table 3), are not the same as those stated by the travellers during the survey (SECUR BOARD and SECUR STATION) (see Table 2). Such as factors related with safety were stated by users as the most important, while in the model they were not identified as key factors, neither for the whole sample nor across segments of passengers. Likewise, while users stated CLEAN TOILET as highly important, it was one of the variables less influent across the models, despite of receiving the lowest satisfaction rate. On the other hand, COURTESY BOARD was not stated by users with one of the highest importance rates, but it was derived as one of the three most relevant variables in quite a few segments of passengers. This proves that the importance of the variables stated by the users can fail to identify the real key factors of the passengers' "Overall Satisfaction", as other previous studies have also pointed out (Weinstein, 2000; de Oña, de Oña, Calvo, 2012).

Therefore, the findings arising from the application of the CART methodology could be very useful for service operators and policy makers to identify the strategy to be adopted for the improvement of the service by considering the different market segments. Once they know the preferences and needs of specific groups of passengers whose loyalty are seeking, they will be able to perform more successful interventions.

However, not all are advantages. CART methodology presents some limitations (Chang and Wang, 2006): CART always produces binary trees; unlike other parametric models, CART does not provide a confidence interval or probability level to the splitters and predictions in the model; in addition, tree models could be unstable, and the structure and accuracy of the models generated could change depending on the strategy followed for stratifying the sample in the training and testing subsets. In addition, while other methodologies have the ability to identify variations in the effect of some variables over users' perceptions (i.e. the random ordered probit model used by Bordagaray et al., 2013), for CART methodology this heterogeneity has to be studied by segmentation of the sample.

In future research, the results of this paper could be validated and compared with the outcomes obtained by using other algorithms for building decision trees (such as C5.0 algotrithm), in order to identify the most important attributes and more powerful decision rules when the variables inducing them are coincidential in both algorithms. Moreover, service quality could be predicted by regression trees, forecasting a value of the overall quality and not a classification of this quality.

REFERENCES

Abdel-Aty M, Keller J, Brady PA. 2005. Analysis of types of crashes at signalized intersections by using complete crash data and tree-based regression. Transportation Research Record 1908:37-45.

Agunloye OO, Oduwaye L. 2011. Factors influencing the quality of rail transport services in metropolitan Lagos. Journal of Geography and Regional Planning 4(2):98-103.

Bastin F, Cirillo C, Toint PL. 2006. Application of an adaptive Monte Carlo algorithm to mixed logit estimation. Transportation Research Part B 40:577-593.

Bhat CR. 2001. Quasi-random maximum simulated likelihood estimation of the mixed multinomial logit model. Transportation Research Part B 35:677-693.

Bhat CR. 2003. Simulation Estimation of Mixed Discrete Choice Models Using Randomized and Scrambled Halton Sequences. Transportation Research Part B 37:837-855.

Bordagaray M, Dell'olio L, Ibeas A, Cecín P. 2013 Modelling user perception of bus transit quality considering user and service heterogeneity. Transportmetrica, DOI: 10.1080/23249935.2013.823579

Breiman L, Friedman JH, Stone CJ, Olshen RA. 1984. Classification and Regression Trees. Boca Raton, Fla: Chapman & Hall/CRC.

Brons M, Givoni M, Rietveld P.2009. Access to railway stations and its potential in increasing rail use. Transportation Research Part A 43:136-149.

Bruhn M,Grund MA. 2000. Theory, development and implementation of national customer satisfaction indices: the Swiss Index of Customer Satisfaction (SWICS). Total Quality Management 11:1017-1028.

Cantwell M., Caulfield B., O'Mahony M. 2009. Examining the Factors that Impact Public Transport Commuting Satisfaction. Journal of Public Transportation 12(2):1-21.

Cavana RY, Corbett LM, Lo YL. 2007. Developing zones of tolerance for managing passenger rail service quality. International Journal of Quality and Reliability Management 24(1):7-31.

CEC, Commission of the European Communities. 2001. White Paper – European transport policy for 2010: time to decide. Commission of the European Communities, 370, Brussels.

Chang L, Chen W. 2005. Data mining of tree-based models to analyze freeway accident frequency. Journal of Safety Research 36(4):365-375.

Chang L, Wang H. 2006. Analysis of traffic injury severity: An application of non-parametric classification tree techniques. Accident Analysis and Prevention 38(5):1019-1027.

Chen W, Jovanis PP. 2000. Method for identifying factors contributing to driver-injury severity in traffic crashes. Transportation Research Record 1717:1-9.

Chou J-S, Kim C b, Kuo Y-C, Oua N-C. 2011. Deploying effective service strategy in the operations stage of high-speed rail. Transportation Research Part E 47:507-519.

Cirillo C, Eboli L, Mazzulla G. 2011. On the asymmetric user perception of transit service quality. International Journal of Sustainable Transportation 5:216-32.

Council FM, Stewart JR. 1996. Severity indexes for roadside objects. Transportation Research Record 1528:87-96.

de Oña J, de Oña R, Calvo FJ. 2012. A classification tree approach to identify key factors of transit service quality. Expert System with Applications 39:11164-11171.

de Oña J, de Oña R, Eboli L, Mazzulla G. 2013. Perceived service quality in bus transit service: A structural equation approach. Transport Policy 29:219-226

de Oña R, Eboli L, Mazzulla G. 2012. Key factors affecting rail service quality. a decision tree

approach. XIX Conference SIDT, Padua, 18-19 October, 2012.
dell'Olio L, Ibeas A, Cecín, P. 2010. Modelling user perception of bus transit quality. Transport Policy 17(6):388-397.
dell'Olio L, Ibeas A, Cecín P. 2011. The quality of service desired by public transport users. Transport Policy 18(1):217-227.
Eboli L, Mazzulla G. 2007. Service quality attributes affecting customer satisfaction for bus transit. Journal of Public Transportation 10(3):21-34.
Eboli L, Mazzulla G. 2008. A Stated Preference Experiment for Measuring Service Quality in Public Transport. Transportation Planning and Technology 31(5):509-523.
Eboli L, Mazzulla G. 2010. How to capture the passengers' point of view on a transit service through rating and choice options. Transport Reviews 30:435-450.
Eboli L, Mazzulla G. 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, Mazzulla G. 2012a. Performance indicators for an objective measure of public transport service quality. European Transport 51:1-21.
Eboli L, Mazzulla G. 2012b. Structural Equation Modelling for Analysing Passengers' Perceptions about Railway Services. Procedia-Social and Behavioural Science 54:96-106.
Evans GW, Wener RE. 2007. Crowding and personal space invasion on the train: Please don't make me sit in the middle. Journal of Environmental Psychology 27:90-94.
Geetika SN. 2010. Determinants of Customer Satisfaction on Service Quality: A Study of Railway Platforms in India. Journal of Public Transportation 13(1):97-113.
Hallmark SR, Guensler R, Fomunung I. 2002. Characterizing On-Road Variables That Affect Passenger Vehicle Modal Operation. Transportation Research Part D 7(2):81-98.
Harris K, Baron S. 2004. Consumer-to-consumer conversations in service settings. Journal of Service Research 6:287-303.
Hensher DA. 2014. The Relationship Between Bus Contract Costs, User Perceived Service Quality and Performance Assessment. International Journal of Sustainable Transportation 8(1): 5-27
Hensher DA, Prioni P. 2002. A Service Quality Index for Area-wide Contract Performance Assessment. Journal of Transportation Economics and Policy 36(1):93-113.
Hensher DA, Stopher P, Bullock P. 2003. Service quality–developing a service quality index in the provision of commercial bus contracts. Transportation Research Part A 37:499-517.
Hensher DA, Mulley C, Yahya N. 2010. Passenger experience with quality-enhanced bus service: the tyne and wear 'superoute' services. Transportation 37:239-256.
Huang C, Hsueh S. 2010. Customer behavior and decision making in the refurbishment industry - A data mining approach. Journal of Civil Engineering and Managing 16(1):75-84.
Huse C, Evangelho F. 2007. Investigating business traveller heterogeneity: Low-cost vs full-service airline users? Transportation Research Part E 43:259–268
Indian Railways - A Study of South Central Railways. International Journal of Business and Management 5(9):139-146.
Irfan SM, Kee DMH, Shahbaz S. 2012. Service Quality and Rail Transport in Pakistan: A Passenger Perspective. World Applied Sciences Journal 18(3):361-369.
Jen W, Tu R, Lu T. 2011. Managing passenger behavioral intention: an integrated framework for service quality, satisfaction, perceived value, and switching barriers. Transportation 38:321-342.
Joewono TB, Kubota H. 2007. User perception of private paratransit operation in Indonesia.

Journal of Public Transportation 10(4):99-118.
Kashani AT, Mohaymany AS. 2011. Analysis of the traffic injury severity on two-lane, two-way rural roads based on classification tree models. Safety Science 49:1314-1320.
Kuhnert PM, Do K, McClure R. 2000. Combining non-parametric models with logistic regression: An application to motor vehicle injury data. Computation Statistics and Data Analysis 34(3):371-386.
Lao Y, Liu L. 2009. Performance evaluation of bus lines with data envelopment analysis and geographic information systems. Computers, Environment and Urban Systems 33:247-255.
Magazzù D, Comelli M, Marinoni A. 2006. Are car drivers holding a motorcycle licence less responsible for motorcycle - Car crash occurrence?: A non-parametric approach. Accident Analysis and Prevention 38(2):365-370.
Montella A, Aria M, D'Ambrosio A, Mauriello F. 2012. Analysis of powered two-wheeler crashes in Italy by classification trees and rules discovery. Accident Analysis and Prevention 49:58-72.
MVA Consultancy for Passenger Focus. 2007. Passengers' priorities for improvements in rail services: Summary of research conducted by MVA Consultancy for Passenger Focus, accessed 22 July 2010.
Nathanail E. 2008. Measuring the Quality of Service for Passengers on the Hellenic Railways. Transportation Research Part A 42:48-66.
Nurul-Habib KM, Kattan L, Islaam T. 2011. Model of personal attitudes towards transit service quality. Journal of Advanced Transportation 45:271-285.
Oyewole P. 2001. Consumer's socio-demographic characteristics and satisfaction with services in the airline industry. Services Marketing Quarterly 23(2):61–80
Pakdil F, Aydin Ö. 2007. Expectations and perceptions in airline services: An analysis using weighted SERVQUAL scores. Journal of Air Transport Management 13:229-237.
Pakgohar A, Tabrizi RS, Khalili M, Esmaeili A. (2011) The role of human factor in incidence and severity of road crashes based on the CART and LR regression: a data mining approach. Procedia Computer Science, 3, 764−769.
Paquette J, Bellavance F, Cordeau J, Laporte G. 2012. Measuring quality of service in dial-a-ride operations: The case of a canadian city. Transportation 39(3):539-564.
Pande A, Abdel-Aty M, Das A. 2010. A classification tree based modeling approach for segment related crashes on multilane highways. Journal of Safety Research 41(5):391-397.
Parasuraman A, Zeithaml VA, Berry LL. 1988.SERVQUAL: A Multiple-Item Scale for Measuring Consumer Perceptions of Service Quality. Journal of Retailing 64(1):12-40.
Passenger Focus. 2007. The Pennine Class 185 experience: What do passengers think? May 2007.
Prasad MD, Sheklar BR. 2010. Impact of Service Quality Management (SQM) Practices on Indian Railways. A Study of South Central Railways. International Journal of Business and Management 5(9):139-146.
Qin X, Han J. 2008. Variable selection issues in tree-based Regression Models. Transportation Research Record 2061:30-38.
Rahaman KR, Rahaman A. 2009. Service quality attributes affecting the satisfaction of railway passengers of selective route in south western part of Bangladesh. Theoretical and Empirical Researches in Urban Management 3(12):115-125.
Rail Corporation-NSW Government and CityRail. 2009. Annual progress report into delivering City Rail's customer charter 2009, accessed on 12/8/2010.

Saputra AD. 2010. Analysis of train passenger responses on provided service. Karlstads Universitet.

Sheth C, Triantis K, Teodorovic´ D. 2007. Performance evaluation of bus routes: A provider and passenger perspective. Transportation Reseaarch Part E 43:453-478.

Sohn SY, Shin H. 2001. Pattern recognition for road traffic accident severity in Korea. Ergonomics 44(1):107-117.

Stathopoulos A, Marcucci E. 2014. De Gustibus Disputandum Est: Non- Linearity in Public Transportation Service Quality Evaluation. International Journal of Sustainable Transportation 8(1): 47-68

Washington S, Wolf J, Guensler R. 1997. Binary Recursive Partitioning Method for Modeling Hot-Stabilized Emissions from Motor Vehicles. Transportation Research Record 1587:96-105.

Washington S, Wolf J. 1997. Hierarchical Tree-Based Versus Ordinary Least Squares Linear Regression Models: Theory and Example Applied to Trip Generation. Transportation Research Record 1581:82-88.

Weinstein A. 2000. Customer satisfaction among transit riders. How customer rank the relative importance of various service attributes. Transportation Research Record 1735:123-132.

Witten IH, Frank E. 2005. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufman, Amsterdam

Wong J, Chung P. 2007. Managing valuable Taiwanese airline passengers using knowledge discovery in database techniques. Journal of Air Transportation Management 13(6):362-370.

Xie C, Lu JY, Parkany E. (2003). Work travel mode choice modeling with data mining: decision trees and neural networks. Transportation Research Record 1854, 50–61.

Yan X, Radwan E. 2006. Analyses of rear-end crashes based on classification tree models. Traffic Injury Prevention 7(3):276-282.

Yan X, Richards S, Su X. 2010. Using hierarchical tree-based regression model to predict train-vehicle crashes at passive highway-rail grade crossings. Accident Analysis and Prevention 42(1):64-74.

Table 1 Sample characteristics

Characteristics	*Statistics*
1.Gender	Male (45.8%), female (54.2%)
2.Age	16-25 (40.5%), 26-40 (33.1%), 41-65 (22.4%), > 65 year-olds (4.0%)
3.Professional condition	employee (37.7%), manager (1.6%), entrepreneur (1.3%), freelancer (5.2%), self-employed worker (4.9%), unemployed (4.1%), student (38.1%), housewife (2.1%), pensioner (4.3%), other (0.7%)
4. Monthly income level	<= 1,000 (9.6%), 1,001-1,500 (14.5%), 1,501-2,000 (7.4%), 2,001-3,000 (3.8%), 3,001-4,000 (1.3%), > 4,000 Euros (1.5%), no fixed income (39.5%), no answer (22.5%)
5.Qualification	University (28.9%), upper secondary school (55.6%), lower secondary school (14.0%), primary school (1.6%)
6.Scope of journey	Work (40.4%), studying (32.8%), bureaucratic activities (3.6%), personal activities (20.6%), tourism (2.6%)
7.Frequency of journey	Daily (61.9%), weekly (14.5%), occasionally (23.6%)
8.Type of ticket	One-way ticket (25.3%), travel card (74.7%)
9. Type of User	Commuter Workers (32.3%),Commuter Students (27.3%), Others (40.4%)
10. Day of the trip	Weekdays (84.6%), days before holidays (8.2%), holidays (7.2%)
11. Time of the trip	Off-peak hour (30.7%), morning peak hour (18.8%), afternoon peak hour (29.4%), evening peak hour (21.1%)

Table 2 Importance and satisfaction rates

Service aspect	*Service quality attribute*	CODE	Importance rate	Satisfaction rate
Safety	1. Travel Safety	SAFETY	8.98	7.43
	2. Personal Security on Board	SECUR. BOARD	9.01	6.76
	3. Personal Security at Station	SECUR. STATION	9.00	6.48
Cleanliness	4. Cleanliness of Vehicles	CLEAN VEH	8.43	5.32
	5. Cleanliness of Seats	CLEAN SEATS	8.48	5.22
	6. Cleanliness of Toilet Facilities	CLEAN TOILET	8.26	4.44
	7. Cleanliness of Stations	CLEAN STATION	7.96	5.53
	8. Maintenance of Stations	MAINTEN STATION	7.82	5.49
Comfort	9. Crowding on Board	CROWDING	8.01	5.33
	10. Air-conditioning on Board	AIR CONDIT	8.09	5.41
	11. Windows and Doors Working	WIND AND DOORS	7.95	5.74
Service	12. Fare/Service Ratio	FARE	8.39	5.17
	13. Frequency of Runs	FREQUENCY	8.41	6.12
	14. Punctuality of Runs	PUNCTUALITY	8.69	5.52
	15. Regularity of Runs	REGULARITY	8.53	5.80
	16. Price Integration with PT	PRICE INTEGR	7.54	5.95
	17. Localization of Stations	LOCALIZAT	7.84	6.65
Other	18. Parking	PARKING	7.04	5.49
	19. Bicycle Transport on Board	BICYCLE BOARD	6.00	6.03
	20. Facilities for Disabled	FAC.DISABLED	7.91	5.25
Information	21. Information at Stations	INFO STATION	8.04	5.72
	22. Information on Board	INFO BOARD	8.00	5.45
	23. Complaints	COMPLAINTS	7.48	5.17
	24. Info Connections with PT	INFO CONNECT	7.33	5.26
Personnel	25. Courtesy and Competence on Board	COURTESY BOARD	7.92	6.67
	26. Ticket Inspection	TICKET INSP	7.68	6.20
	27. Courtesy and Competence in Station	COURTESY STATION	7.91	6.38
	Overall service			5.81

Figure 1 CART for suburban lines

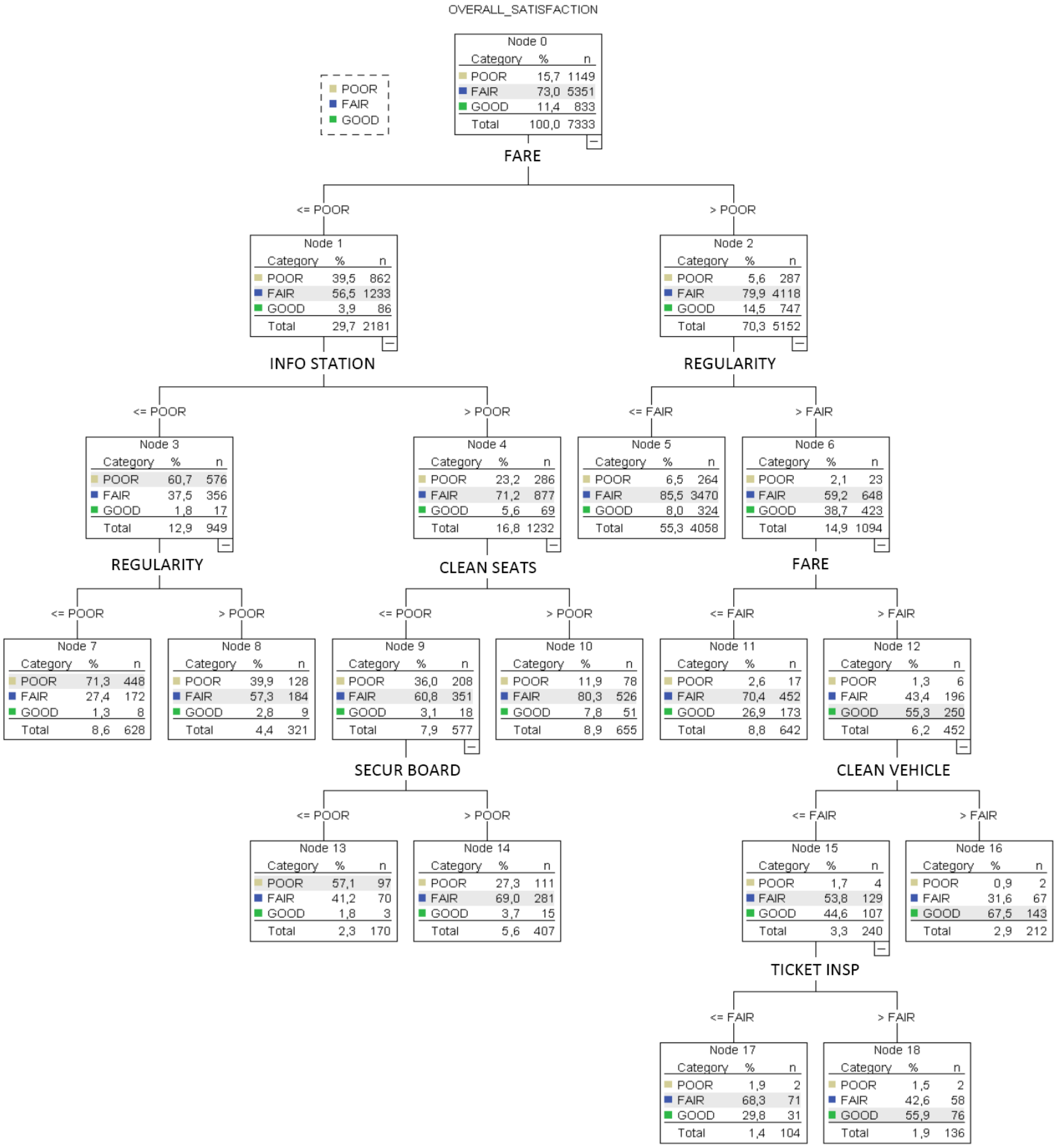

Table 3 Importance for users of suburban lines and for users classified according to the four criteria (type of user; day of the trip; frequency of the trip; time of the trip)

Criteria of Classification	*Category of User*	*Independent Variable*	*Normalized Importance*
SUBURBAN LINES (n. obs. 7,333; prec. rate 78.15%)		Regularity of Runs	100.0%
		Punctuality of Runs	93.7%
		Cleanliness of Vehicles	86.6%
TYPE OF USER (n. obs. 7,333)	COMMUTER WORKERS (n. obs. 2,371; prec. rate 78.87%)	Regularity of Runs	100.0%
		Information at Stations	93.7%
		Information on Board	93.4%
	COMMUTER STUDENTS (n. obs. 2,000; prec. rate 77.60%)	Fare/Service Ratio	100.0%
		Courtesy and Competence on Board	84.7%
		Windows and Doors Working	78.6%
	OTHERS (n. obs. 2,962; prec. rate 77.28%)	Regularity of Runs	100.0%
		Complaints	81.6%
		Frequency of Runs	80.4%
DAY OF THE TRIP (n. obs. 7,333)	WEEKDAYS (n. obs. 6,207; prec. rate 76.83%)	Regularity of Runs	100.0%
		Punctuality of Runs	96.3%
		Courtesy and Competence on Board	85.9%
	DAYS BEFORE HOLIDAY (n. obs. 600; prec. rate 77.50%)	Courtesy and Competence in Station	100.0%
		Courtesy and Competence on Board	96.6%
		Travel Safety	80.2%
	HOLIDAYS (n. obs. 526; prec. rate 76.43%)	Fare/Service Ratio	100.0%
		Personal Security at Station	66.5%
		Frequency of Runs	60.1%
FREQUENCY OF THE TRIP (n. obs. 7,333)	DAILY (n. obs. 4,542; prec. rate 77.41%)	Courtesy and Competence on Board	100.0%
		Regularity of Runs	99.3%
		Information on Board	94.5%
	WEEKLY (n. obs. 1,064; prec. rate 77.63%)	Information at Stations	100.0%
		Windows and Doors Working	92.4%
		Regularity of Runs	87.2%
	OCCASIONALY (n. obs. 1,727; prec. rate 77.53%)	Courtesy and Competence on Board	100.0%
		Fare/Service Ratio	93.2%
		Courtesy and Competence in Station	77.9%
TIME OF THE TRIP (n. obs. 7,333)	OFF-PEAK HOUR (n. obs. 2,250; prec. rate 76.22%)	Regularity of Runs	100.0%
		Cleanliness of Vehicles	99.8%
		Windows and Doors Working	99.6%
	MORNING PEAK HOUR (n. obs. 1,382; prec. rate 76.19%)	Courtesy and Competence on Board	100.0%
		Fare/Service Ratio	99.1%
		Courtesy and Competence in Station	94.4%
	AFTERNOON PEAK HOUR (n. obs. 2,154; prec. rate 75.90%)	Frequency of Runs	100.0%
		Punctuality of Runs	95.7%
		Regularity of Runs	95.3%
	EVENING PEAK HOUR (n. obs. 1,547; prec. rate 77.96%)	Courtesy and Competence on Board	100.0%
		Courtesy and Competence in Station	81.7%

Figure 2 CART for users classified according to the type of user ("Commuter Workers")

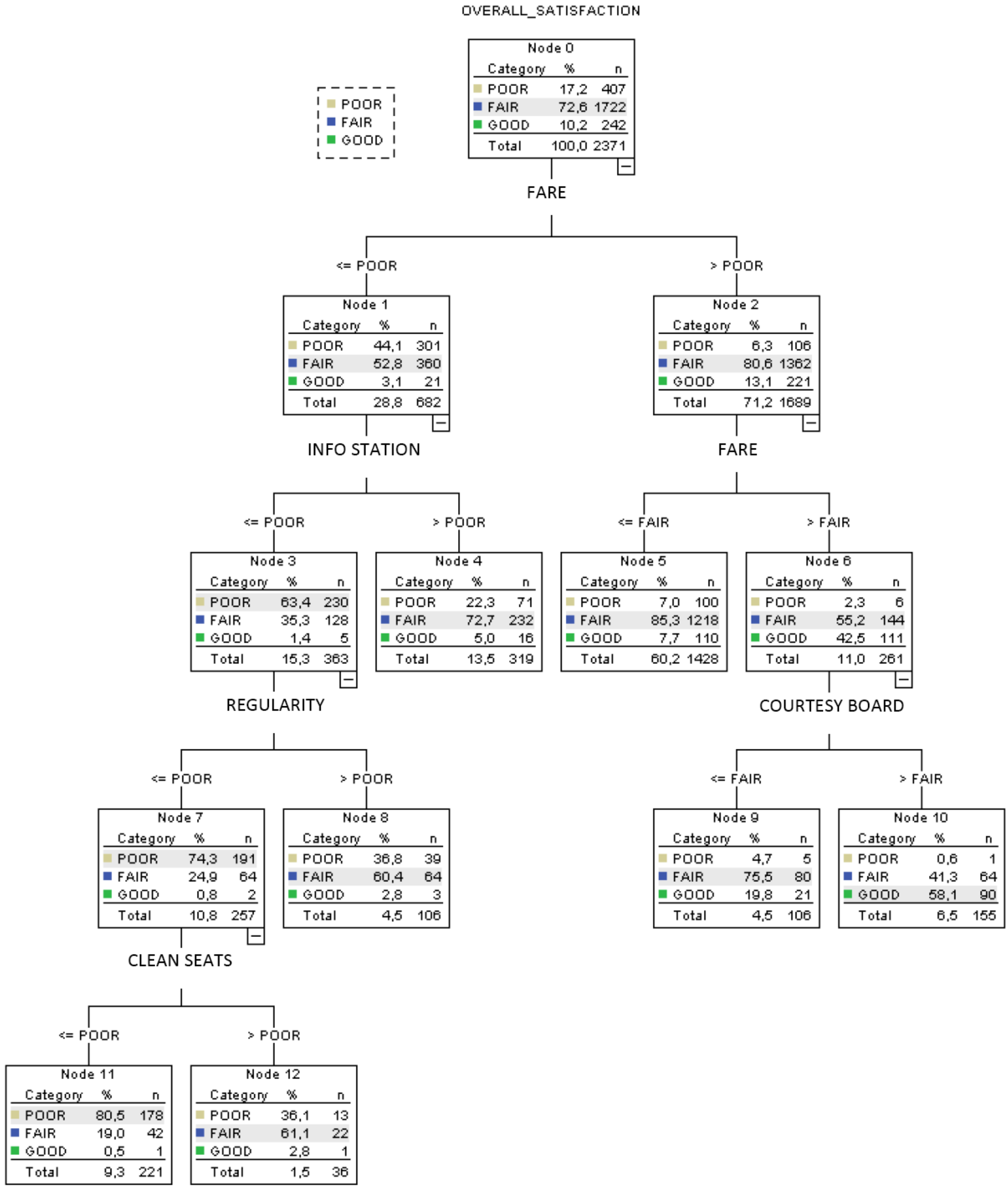

Figure 3 CART for users classified according to the day of the trip ("Weekdays")

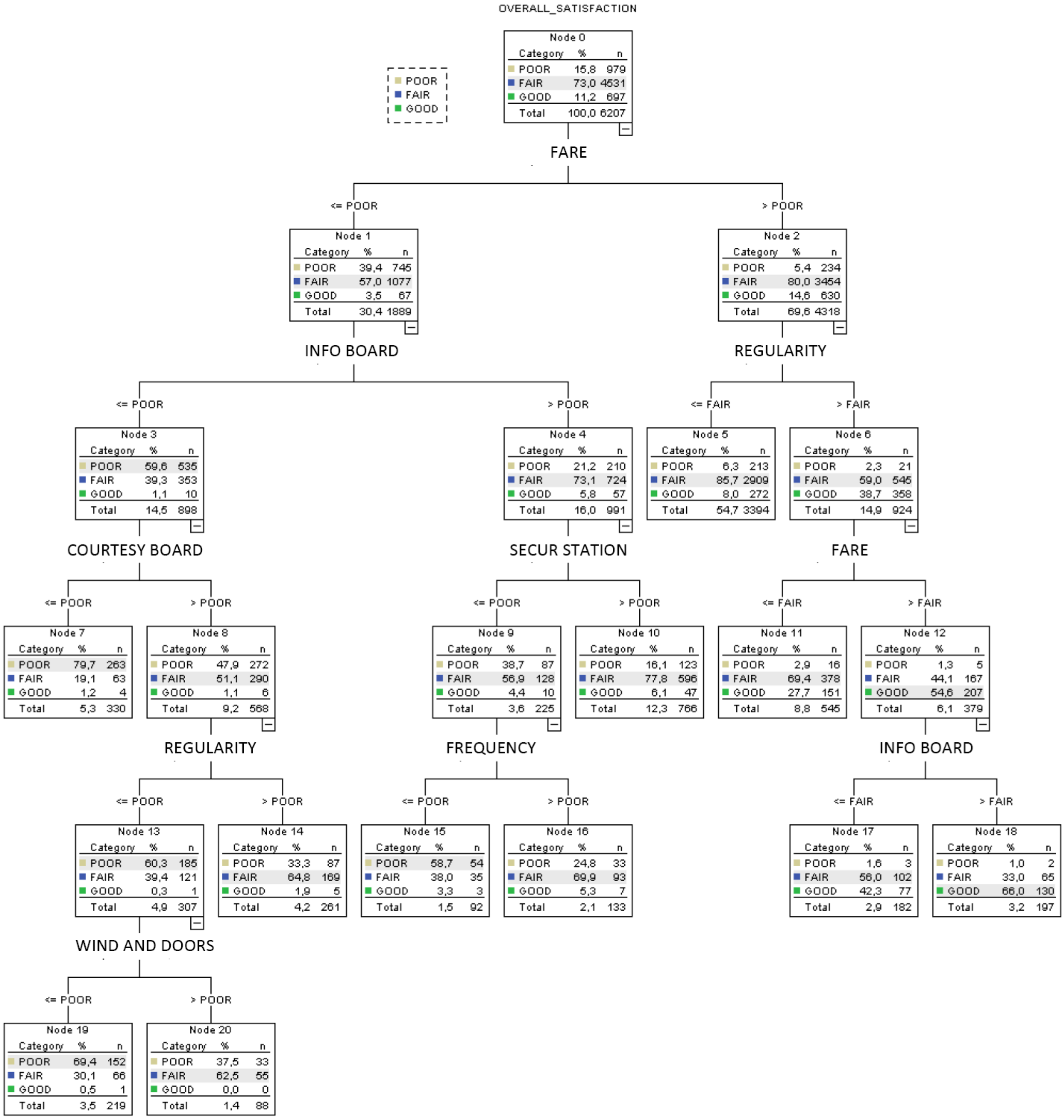

Figure 4 CART for users classified according to the frequency of the trip ("Daily")

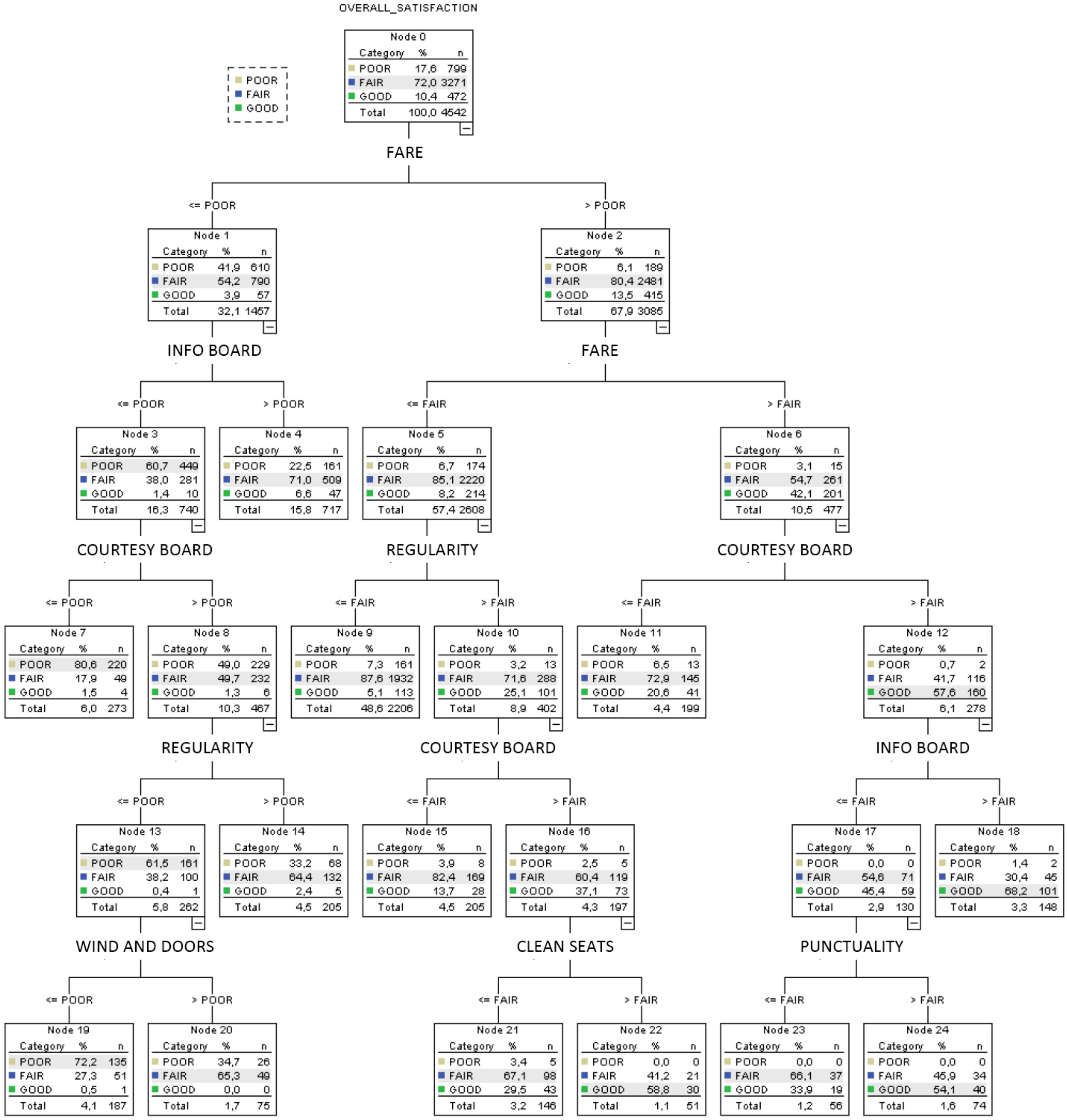

Figure 5 CART for users classified according to the time of the trip ("Afternoon peak hour")