Technical University of Munich
Department of Civil, Geo and Environmental Engineering
Chair of Transportation Systems Engineering

MASTER'S THESIS

Public Transport Users' Satisfaction: Athens vs Munich

Fatemeh Fathisardehaee

Supervised by

Univ.-Prof. Dr. Constantinos Antoniou

August 2018

Abstract

Public transport as the most effective rival of private cars is an important target that saves growing cities from more traffic congestion, environmental damages and energy consumption. Policies are showing more interest in measures that attract more transit use; and therefore, studies on the quality of service and identifying the factors that affect this quality in public transport sectors has been increased in the past years.

With the data available from the surveys conducted in two European cities in 2017, this thesis has aimed to study the indicators of quality of transit service in Athens and Munich. Some research on each of the data sets has been done before but in this work the data from both cities is used to gain more insight about user satisfaction in different areas.

Two statistical methods, exploratory factor analysis and ordered logistic regression modeling, have been applied. Factor analysis as a reduction method examines a large number of indicators to find fewer underlying factors. With ordered logit model the relation between service quality and demographic predictors and overall satisfaction can be investigated.

The results of analysis show that service quality, production of the service, information and transfer quality are the four factors that have been seen in all data sets as indicators of importance. Ticketing system and accessibility are other two important dimensions of the service quality that have been observed in most of the data sets. The results of logit model shows that the factor scores extracted from factor analysis are decisive and affects overall satisfaction. In Athens, age groups and mode have showed some relation with overall satisfaction as well.

The data in Munich seems to lack an adequate level of detail to estimate ordered logit model. This study can be improved by validating the estimated model with data from a new service or more data from the current cities using the same survey questions and also by performing more sophisticated models.

Acknowledgment

I would like to express my profound gratitude to Prof. Dr. Constantinos Antoniou for offering me the topic and providing the data and for his constant supervision. This thesis has benefited a great deal from his immense knowledge in the field of transportation service quality modeling. I am very grateful to him for his invaluable feedback and encouragement and for being such an understanding teacher.

Special thanks to Emmanouil Chaniotakis for his guidance at initial stages of this work. I also would like to thank all the employees and assistants at the chair of Transportation Systems Engineering who supported me unconditionally whenever I asked for help.

And finally, to my parents, my husband, Mehrdad, and all my family and friends, so much love and thanks for your unfailing support. The completion of this thesis would not be possible without the love and care of all the beloved people who helped me not to lose hope and to believe in myself.

Contents

1 Introduction

1.1 The need for improving transit ridership

Riding on cars could provide a faster and more convenient daily trips but ineffective policies and arrangements have lead our cities to suffer from congestion, time loss, fuel consumption, delays and safety issues (Downs, 2005). The environmental damage of transportation, particularly climate change is a serious concern and its negative aspects are getting wider and wider (Bickel et al., 2006). On the other hand the demand for mobility cannot be avoided and there are many drivers and forces behind using more private cars; therefore, new contributions to increase the use of more sustainable and environmentally friendly modes of transport is required (Eboli and Mazzulla, 2013; Jakobsson et al., 2011).

Beside walking and cycling, public transport is considered to be a more sustainable mode of transport than private cars (Redman et al., 2013). Car ownership is one of the most important factors affecting public transport demand (Efthymiou et al., 2017) and that is why we can assume these two transport systems as one another's rivals. The new policies also call for solutions to high tendency of using private cars and a shift to public transport. Public transport agencies also has come up with new strategies to meet this new policy approach (Chowdhury et al., 2018; Downs, 2005; TRL Limited, 2004; van Lierop and El-geneidy, 2017).

1.2 Transit service quality and customer's satisfaction

A shift from private car to public transport happens when the service takes care of users' expectations (Chowdhury et al., 2018). User choice is very much affected by the quality of transit service (Cirillo et al., 2011). Zeithaml et al. (1990) argue that quality is the key factor of user satisfaction and true customers are created through qualified services. They believe service quality is the difference between expectations and perceptions of users. In public transport customer satisfaction is also defined as expectations of the users from the transit service that have been fulfilled in reality (Tyrinopoulos and Antoniou, 2008). The ten dimensions that Zeithaml et al. (1990) explain for service quality and the instruments they introduce to measure it are a guide for policy makers and service providers of any field to understand their customers and build a more attractive service for them.

“A transit performance measure is defined as a quantitative or qualitative factor used to evaluate a particular aspect of a transit service” (Eboli and Mazzulla, 2011) and service quality is one of the specific measures that explains performance of public transportation (Eboli and Mazzulla, 2011); however, measuring the quality is not easy and it is a challenge to be economically analyzed. Number of passengers that ride on a transit or vehicle kilometer cannot give an actual idea on how the quality of transit service is. When the focus is only on these objective approaches an important aspect of service quality is ignored and it is that the quality of service is not homogenous. People may use a specific service but it does not necessarily mean they are satisfied with it. Passengers consider various aspects and attributes of the service to evaluate it and is not always related to the amount of ridership (Cirillo et al., 2011; Roman et al., 2014).

As mentioned above policies tendency is to attract more people from private cars to public transit and as a result the transit providers and agencies modify their strategies to reach this goal. On one hand providers need to improve their service and on the other hand the authorities should observe and supervise the process. Therefore, measuring the quality of service plays an important role for both parties (Garrido et al., 2014; Tyrinopoulos and Aifadopoulou, 2008). If a transit company wants to realize how people get more attracted to their service they need to know passengers’ view is the key point (van Lierop and El-geneidy, 2017). Evaluating the service clarifies which aspects of the service are good enough through customers’ eyes and should be maintained and also which attributes need improvements. It is vital to understand that customers’ satisfaction level shows how well a service performs and it is to be noted users of a service are the judge of its level of quality (Cirillo et al., 2011).

Studying customer satisfaction as a key to success is getting into the focus of many policy makers and service providers but in next step the challenge is to learn about customers’ needs and expectations. Many economic or non-economic factors influence customers’ decision (Biesok and Wyród-wróbel, 2012). Quality concept is in itself complex, fuzzy and abstract and it makes it difficult to measure it. Besides, the huge variety of quality attributes that can be considered, the subjective nature of data collected for studying service quality and also the way surveys are designed and done makes the measurement of public transport service quality more difficult that acquires deeper and more precise insight (Cirillo et al., 2011; de Oña and de Oña, 2014). Thus, how one approaches customers, how questions and indicators are defined and how the collected information is analyzed affect the results.

Service quality can be studied through its related indicators (Eboli and Mazzulla, 2011) but there are different ways to define and obtain these indicators. Listening to customers and learning from their complains is one way of finding related attributes (Berry and Parasuraman, 1997).

Another method which is more objective is to observe how changes in the service improves or reduces the ridership (Tang and Thakuriah, 2012). The next approach which is also the focus of this research is subjective and is to explore users' opinions, expectations and needs through surveys which can be in-person on-board/ off-board or online questionnaire surveys, or telephone interviews.

Figure 1 shows how the need of improving transit ridership leads to studying and learning more about its users. Collecting enough information from riders about service attributes should be analyzed using appropriate procedures or statistical models.

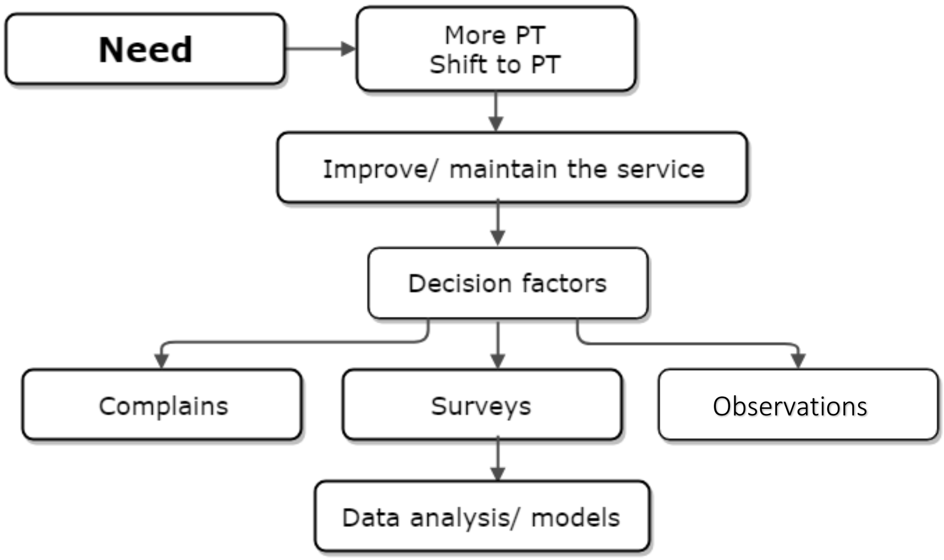


Figure 1 The flowchart of need: from transit ridership improvement to data analysis

1.3 Thesis motivation

Munich is the third largest city in Germany and the capital of the state of Bavaria. With the current population of about 1.5 million inhabitants the city is dealing with growing population and it is predicted that by 2030 a population growth of 200,000 people will occur. The most important goal of the city's transport plan gives priority to the measures that help to reduce traffic or make a shift to walking, cycling and public transport. Between 2002 and 2008, kilometer-cycled had a remarkable improvement but public transport distance travelled has been remained almost the same level of 21% (Munich, 2013).

In Munich 522 million people were transported by public transit system in 2011. More than 60 percent of residents of the city use underground, tram and bus several times a week, and 35 percent of them use the system on a daily basis (Hall and Gerike, 2014).

The capital and largest city of Greece, Athens, is one of the biggest economic centers in southern Europe. In 2011, the urban area of Athens has been the home for more than three million inhabitants (Eurostat).

In Athens, private car has the biggest share in modal split and its magnitude is more than double of the public transport share. In a time span of 10 years before 2010 in Athens the private car use has increased by 7% while the transit ridership has declined. About 43% of the vehicles in the country are concentrated in Athens (The National Archives, 2011), and it is not a surprise that in the last few years a lot of transport policies have been implemented to address car traffic issues in the capital city. The available public transport modes in the Athens are metro, bus, trolley bus and tram (Efthymiou et al., 2014).

These statistics show that there are potential transit riders and there is room for much more improvements of the public transport services. Studying quality of service and exploring passengers' perceptions and expectation is a great contribution in having more sustainable mobility in cities.

Efthymiou et al. (2017) have tried to explore user satisfaction and quality of service in Athens and Zuo (Zuo, 2017) has used the same survey with some changes and conducted it in the city of Munich. The data of the surveys in Athens and Munich both in 2017 was available at the chair of Transportation Systems Engineering in Technical University of Munich and the opportunity to explore these data sets further and to build some models that make the comparison between the two cities possible has been a big motivation to do this thesis.

Some questions that this work has tried to find an answer to are listed as:

- ❑ What aspects of public transport service influence users' satisfaction and which ones are not decisive for them?
- ❑ How satisfaction indicators differ between transit riders of two European cities and what similarities are there?

- ❑ Does it make sense to have similar standards of transit service quality in different areas?
- ❑ Can a model that measures transit performance be generalized to different cities?

1.4 Expected contributions

It is expected that this study and the findings of it contribute both methodologically and practically as follow:

- The modeling results and output can be used for further model developments such as hybrid models that can deal with relationships between different models simultaneously and build a more powerful unite model.
- The output of models can also give an idea to choose input for more complicated models such as a Structural Equation Model in which the most relevant factors should be chosen by researcher.
- The subjective indicators studied in this research can also influence hand books and standards to modify their objective indicators and thresholds to be more user centered.
- Public transport operators and policy makers could use the findings of applied statistical methods in their future plans and measures as it gives an insight for understanding users' perception and expectations better and therefore to increase their ridership.
- With such information they also can allocate their resources in a much more effective way. If budget and resources are spent on features of the service that are not that decisive for customers they could be better allocated to improve the quality of service regarding those attributes that are more important to users. They can save resources or improve weak points in this way. If an indicator is below standards but people feel satisfied about it does not need more allocation of resources.

1.5 Approach

The focus of this thesis is public transport users' satisfaction of service quality indicators in cities of Athens and Munich. The idea is to use an extraction method to investigate the underlying influential aspects of service quality and also to use a modelling framework that can estimate the impact of different quality attributes on the overall satisfaction. These methods should have the ability to deal with ordered data.

To figure out which explanatory method and choice model is suitable, how they are formulated and how to implement them the following steps have been adopted:

- Review of the previous studies on customer satisfaction and service quality;
- Review of the literature on modeling framework that has been used with similar purposes and data type;
- Exploration of data in hand;
- Specification of model choice set, variables and predictors and structure;
- Estimation of model parameters and analysis of results;
- Interpretation of results to answer defined research questions.

Following chapters cover all the steps mentioned here and a very detailed methodology of selected method and model is given in chapter 5.

1.6 Framework of the work

This thesis goes over the state-of-the-art and relevant literature in Chapter 2. Chapter 3 explains how data sets have been prepared to be used in this study and then gives more insight about data. Model development and analytical methods that have been used in this thesis and the detailed methodologies are explained in Chapter 4. Chapter 5 includes the results of analysis and modeling from previous chapter and the related discussion and interpretation of the output. The report is concluded with further recommendations in Chapter 6.

2 Literature Review

2.1 Standards and guidelines available in Europe and US

Public transport service quality is measured by criteria which are qualitative and quantitative. One reference that has been widely used in most of public transport systems is EN 13816:2002[1] (Ngoc et al., 2017). It has eight categories for service quality indicators which are availability, accessibility, information, time, customer care, comfort, security and environmental impact. The advantage of this European Standard is that it can be applied not only in European countries but also in any country including developing countries.

The TRB Handbook for Measuring Customer Satisfaction and Service Quality (TRB, 1999) in light of the eight categories determined by EN 13816:2002 recommends some quality attributes. These attributes are defined both from customer's point of view and operator's point of view but they are more disaggregate quantitative performance measures. Later TRB Guidebook for Developing a Transit Performance-Measurement System (TRB, 2003) introduces 400 indicators grouped in 31 criteria.

The Transportation Research Board (TRB) has more manuals and guidebooks related to performance attributes, standards and thresholds such as the TRB Transit Capacity and Quality of Service Manual (TRB, 2004). There are also other research associations that have provided guidance, benchmarks and recommendations in the field of public transport service quality such as A Practical Guide (TRL Limited, 2004). Eboli and Mazzulla (2013) argue that unlike the United States, there is a lack of regulations and univocal procedures for service quality measurement in Europe. However, an attempt has been made by European Committee for standardization.

The Hellenic Institute of Transport (HIT) in Greece with a lot of background in the field of transportation gives advice and counsel to transport organizations (Tyrinopoulos and Aifadopoulou, 2008). To meet the growing enquiry for assessment of public transport service quality in this European country, HIT has introduced a methodology that has been used by the major public transport organizations such as the Urban Transport Organization in Athens (Tyrinopoulos and Aifadopoulou, 2008). Tyrinopoulos and Aifadopoulou (2008)

[1] EN 13816:2002. Transportation - Logistics and Services - Public Passenger Transport - Service Quality Definition, Targeting and Measurement. European Committee for Standardization.

explain this methodology and provide an overview of the main elements of this approach and the indicators used for quality assessment. Something notable in this work is that they focus both on subjective and objective attributes to assess the level of transit service quality. Two statistical methods, factor analysis with ordered logit model and principle component analysis with multinomial logit model, have been introduced to be used for data analysis.

2.2 Service quality attributes

Many studies have focused on understanding customers' expectations and have analyzed their level of satisfaction. Some studies have tried to figure out which attributes of the service is more important to users (Eboli and Mazzulla, 2011). Cirillo et al. (2011) note that the most investigated aspect of transit service in literature is reliability of the service. When a service is unreliable it is very likely to lose passengers. Their results also validate that reliability of public transport plays a crucial role on users' evaluation of the service. Another interesting point in their work is how different users would have various perceptions of the service. For example as their respondents have been more of young people, information provision has been one of the least important aspects; which would not be the case for older ones who need more help and support.

Dell' Olio et al. (2010) investigate bus users' perception of the quality of the public transport in the city of Santander in Spain. They take a deeper look on how important and relevant the attribute of information provision is. They claim that user groups show different perception of the service but the overall quality of the bus service should be improved. Reliability, travel time, information provision and driver kindness are some of most important attributes to passengers.

Eboli and Mazzulla (2013) indicate that objective indicators provide a reliable and realistic measure of level of service quality and it will be less biased comparing to subjective indicators that are only base of passenger's perception; however, they suggest to benefit from both types of attributes together. Service availability, service reliability, comfort, cleanliness, safety and security, information, customer care and environmental impacts are the criteria which they have found in their investigations as main aspects that characterize a service.

While some studies only have focused on objective attributes, others have tried to explore both objective and subjective indicators. Strong subjectivity would bring bias, moreover, it does not consider non users' opinion and perception (de Oña and de Oña, 2014). Eboli and Mazzulla (2011) argue that because of the heterogeneity of users' judgments and also the risk that users could have distorted opinions, it is not recommended to only rely on subjective measures. To be able to fully describe and assess the level of the service of transit it is recommended to consider both objective and subjective indicators. The most critical attributes considering both measure in Eboli and Mazzulla's (2011) study have been availability of schedule/maps at stop, availability of shelter and benches at stop, and service frequency.

Most of the studies on transit user satisfaction focus on current passengers and riders. What seems to be also important is to learn more about passive and potential users. In a research done by Redman et al. (2013) they have been tried to figure out which quality indicators are critical to users and with some changes and improvements in them the car users would be attracted to public transport. Beirao and Sarsfield Cabral (2007) consider both transit and private car users in their research to find key factors that affect mode choice. Their findings indicate that service quality of transit should meet customers' expectation in order to attract potential users.

The list of indicators for service quality is lengthy and normally they are grouped as dimensions. There is not a specific number of variables that everyone has or can agree on. Some indicators such as frequency of service, punctuality, comfort and cleanliness, safety, availability of information, personnel courtesy, and fare and so on are used repeatedly in various related studies. Beside these frequently used indicators some more attributes can be used for each context-specific service (de Oña and de Oña, 2014).

Garrido et al. (2014) studied service quality on the Granada bus metropolitan transit system in 2007. They concluded that frequency, speed, information and proximity are the most important attributes; although punctuality, safety and courtesy are other relevant attributes that should be taken into account.

2.3 Studies on different areas or data sets

Fellesson and Friman (2008) investigated transit user satisfaction in eight European cities and they concluded that characteristics of the population affects perception of satisfaction; however, there have been many similarities in their findings. They have used factor analysis and the satisfaction dimensions they extracted were present in not all, but most of the cities and they were system (travel/waiting time, supply and reliability items), comfort, staff and safety. They concluded that there has been a uniform way of interpreting transit in these

European cities. (Friman and Fellesson, 2009) is another similar study of these researchers about public transportation service supply and customer satisfaction.

In the US Taylor et al. (2009) did a cross- sectional analysis to study the transit ridership in 265 urbanized areas regarding external variables. They applied a two-stage simultaneous equation regression models to address simultaneity between service supply and consumption. They concluded that the characteristics of an area such as metropolitan economy, reginal geography and population characteristics have an impact on ridership which are out of the control of transit systems.

Tyrinopoulos and Antoniou (2008) analyzed survey data obtained from riders of five transit services in two cities in Greece, Athens and Thessaloniki, to study public transport users' satisfaction on services. Two statistical methods, factor analysis and ordered logit model, have been used for their work.

Buehler and Pucher (2012) investigated the factors affecting transit ridership in the US and Germany and revealed that Germany has higher odds of using public transit as this European country benefits from a more public transport focused policies. These researchers have been used logistic regression models for their analysis.

Chou and Kim (2009) conducted a structural equation model to study the relationship between service quality, satisfaction and loyalty as major factors that influence service-oriented transportation systems. They studied Taiwan High Speed Rail and then did a cross country comparison of ridership on this system and Korea Train eXpress. In their findings they noted that the introduction of mediation variables, such as passenger complains and corporate image into the service have an impact on ridership experience.

2.4 Service quality and customer satisfaction models

There have been various approaches to investigate the transit service quality in the past decade. Mostly researches have been used logit and probit regression models and Structural Equation Models (SEM). De Oña et al. (2015) note that SEM is the most appropriate methodology, because it is an integrated strategic method that can deal with the phenomenon in whole. A hybrid choice model which integrates SEM in discrete choice models has also found a way in assessing service quality (Efthymiou et al., 2014). Less complex methods such as factor analysis and principle component analysis have been applied on user satisfaction data to give a quicker insight on relevant indicators of satisfaction (Tyrinopoulos and Antoniou, 2008).

Other models that have been used for quality of service in public transport and users' satisfaction are neutral networks (for instance: Garrido et al. (2014) and Shen and Li (2014)), classical tree approach as in (de Oña et al., 2012), Decision trees and If-then method (de Oña and de Oña, 2014), nested choice methods (Hensher et al., 2003) and Analytic Hierarchy Method (AHP) and cluster analysis (Chowdhury et al., 2018).

Some of the most popular methods is explained in this section. Ordered logit models and factor analysis are used in this study; therefore, they are explained in more details.

2.4.1 Structural Equation Models

A Structural Equation Model as a more advanced method in studying service quality is also used for Public Transport users' satisfaction analysis. This method uses a combination of analyses like Regression, Factor and Variance Analysis to consider both unobserved latent variables and observed indicators at the same time when studying a phenomenon. This method consists of a latent variable model and a measurement model (de Oña and de Oña, 2014). This approach is used by many researchers like de Oña et al. (2015), Ye and Titheridge (2017), Stuart et al., (2000), Golob (2003), Chou and Kim (2009) and Eboli and Mazzulla (2012). Efthymiou and Antoniou (2017) and Efthymiou et al. (2014) have used hybrid choice models which is a combination of classical choice models with structural equation models for latent variables.

2.4.2 Discrete Choice Models

To analyze public transport service quality based on the user preference data that is collected through surveys discrete choice models are used. In these models the overall evaluation of each user is considered to be the first relevant indicator that illustrates their overall satisfaction. Although, they may reveal they are satisfied with some attributes of the service or not with the other ones (de Oña and de Oña, 2014). Eboli and Mazzulla (2008), Eboli and Mazzulla (2010) and Hensher (2014) have used this method.

Discrete choice framework uses the principle of utility maximization to model individuals' choice behavior. The assumption is that each individual behaves rationally and has enough information about all the alternatives they have confronted, thus they will choose the alternative with the highest utility. This utility is related to the attributes of the alternatives and also the attributes of the decision maker (Ben-Akiva and Lerman, 1985).

The utility of any given alternative i for any given individual q is a combination of a systematic element V_{in}, and a random component ε_{in}, as shown below (Ortúzar and Willumsen, 2011):

$$U_{in} = V_{in} + \varepsilon_{in},$$

where,

U_{in} is utility of choosing alternative i for the individual q,

V_{in} is systematic component of alternative i for the individual q,

ε_{in} is error component.

And V_{in}, the parameterized function of observed attributes for individual i, can be written as

$$V_{in} = \beta_{1i}X_{1in} + \beta_{2i}X_{2in} + \cdots + \beta_{ki}X_{kin}$$ (Ortúzar and Willumsen, 2011).

$X_{1in}, X_{2in}, \ldots, X_{kin}$ are the k independent variables that include both attributes of the alternative i and socio-economic variables of the individual n; and $\beta_{1i}, \beta_{2i}, \ldots, \beta_{ki}$ are the unknown parameters which are assumed to be constant across individuals, and may vary across alternatives.

The impact of unobserved factors that influences individual's choices is captured by ε_{in} (Ortúzar and Willumsen, 2011). The error term is assumed as a random variable that follows a certain probability distribution; therefore, utilities are also random variables (Ben-Akiva and Lerman, 1985).

Considering a choice set containing two alternatives i and j, the probability that the individual n chooses the alternative i, P_{in}, is as equation below:

$$P_{in} = Probability\ \{ Uin \geq Ujn \}$$ (Ben-Akiva and Lerman, 1985).

To solve the equations a specific probability distribution is considered for error terms. Parameters of β are then estimated using maximum likelihood estimation (Ben-Akiva and Lerman, 1985).

Different types of discrete choice models are the result of different assumptions about the error component distribution (Ben-Akiva and Lerman, 1985). For example logit models have the assumption that ε follows a logistic distribution (Munizaga and Ortúzar, 1999). Logistic regression techniques like probit regression is also used for analysis of discrete choice.

Regression analysis is to model and analyze variables to explore the relationship between them and the dependent variable (Hosmer Jr et al., 2013). The outcome variable is quite often discrete and logistic regression model is the most frequently used regression model for analyzing such data. In a logit model the outcome variable is binary or dichotomous and it affects the form of the model and its assumption comparing to linear regression (Hosmer Jr et al., 2013).

To examine discrete choices the general statistical model, Mixed logit model, which has overcome the limitation of the standard logit model can be used (Train, 2009). Mixed logit model is used in plenty of studies on passenger's perception because it can help to uncover the heterogeneity of their perception. However, there is a drawback on this method as it does not take into consideration the attributes that their marginal utility may be zero and assign a non-zero parameter estimate for all the individual decision makers. Cirillo et al. (2011) used a mixed logit model with a non-parametric distribution of the coefficients to observe public transport users' perception of the quality of service in a sample consist of university students. Roman et al. (2014) applies a Mixed Logit model in combination with multinomial logit models to study the level of satisfaction of transit users on Gran Canaria in Spain.

Multiple linear regression models are another statistical method to study users of public transport. Using factor analysis results Ngoc et al. (2017) employ this model trying to develop some quality standard based on customers' point of view of the service in developing countries. LOS (level of service) concept was introduced and studied in this work. It also applied a weight analysis.

Binary logit models have also been used in this field as van Lierop and El-geneidy (2017) have studied the loyalty of transit users with this model.

A multinomial logit model is a simple extension of the binomial logit model where the dependent variable has more than two unordered outcomes (Venables and Ripley, 2002). After applying the multinomial logit models Roman et al. (2014) conclude that there is a difference in the perception of users of urban or interurban service. Tyrinopoulos and Aifadopoulou (2008) and Eboli and Mazzulla (2008) also have used this method to estimate the overall satisfaction of transportation customers.

Logit model assumes the irrelevant alternatives are independence of each other. In a multinomial logit model where each response is defined as an alternative the ordering of these alternatives does not comply with that assumption of a logit model. This issue can be overcome with ordered alternatives using nested or cross-nested models. In ordered logit models the alternatives are defined according to other alternatives that can be closer or more distant to them (Tyrinopoulos and Antoniou, 2008). Many studies (Efthymiou et al., 2017; Efthymiou et al., 2013; Efthymiou et al., 2014; Iseki and Taylor, 2010; Tyrinopoulos and Antoniou, 2008) have used Ordered Logit models to study user satisfaction of public transport services.

2.4.3 Nested Logit (studies of mode choice or PT service quality)

Fluegel et al. (2015) studies the mode choice for short trips to Oslo using a nested logit model. In their model public transport trips and their combinations have been differentiated. Excel software have been used for the analyses. They present that this model gives rational results for both overall market and submarkets.

Hensher and Ton (2006) have tried to compare neural network models with nested logit in order to investigate their transferability. Nested model has been designed for Sydney, Melbourne and a combination of these two cities. In their result they note that the choice model can better predict the overall market share and neural network model on the other hand has the capability of matching the individual market shares.

Koppelman et al. (1993) argue that nested logit models have an advantage over multinomial ones for analyzing mode choice as sensitivity of mode pairs to changes in the system can be exhibited. Three different nested logit models and a multinomial logit model have been applied in their study.

Hensher et al. (2003) have used nested choice models through experimental design to investigate the users' perception about bus service among different bus operators and tried to make a quantified service quality index.

2.4.4 Ordered Logit Model

The ordered logit model (also proportional odds model), is an ordinal regression model; that is, a regression model for ordinal dependent variables (McCullagh, 1980). When survey respondents are asked to rate something or give their opinion in an ordered way (e.g. agree, neutral or disagree) the data obtained is discreet and ordered. Using standard or nested discrete models on such data would result in loss of all the information reflected by the ordering. Although, there are some cases that using unordered probability models are preferred (Washington et al., 2010).

An appropriate model for the analysis of discrete data which are ordered is an ordered logit model, because it does not have the problem of the violation of the independence of irrelevant alternatives of a logit model resulting from overcoming the violation of errors' independency (Ben-Akiva and Lerman, 1985). An ordered logit model has the ability to estimate parameter coefficients for independent variables, as well as threshold values between the choices which are known as intercepts (Tyrinopoulos and Antoniou, 2008).

Tyrinopoulos and Antoniou (2008), Efthymiou et al. (2017, 2014) have applied ordered logit model on survey data obtained from public transport users to study what attributes of the service have more impact of overall satisfaction of the transit riders.

To model the ordinal ranking of data an unobserved variable z is defined by ordered probability models and it is typically specified as a linear function of each observation such that,

$$z = \beta X + \varepsilon\, \mu_0$$

where X is a vector of variables determining the discrete ordering for observation *n*, β is a vector of estimable parameters and ε is the error term (Washington et al., 2010).

Using this equation, observed ordinal data, y, for each observation is defined as

$$
\begin{aligned}
y &= 1 \quad \text{if } z \le \mu_0 \\
y &= 2 \quad \text{if } \mu_0 < z \le \mu_1 \\
y &= 3 \quad \text{if } \mu_1 < z \le \mu_2 \\
y &= \cdots \\
y &= I \quad \text{if } z \ge \mu_{I-1}
\end{aligned}
$$

where μ is estimable parameter which corresponds to integer ordering and defines y, and I is the highest integer ordered response (Washington et al., 2010).

Parameters μ are estimated jointly with model parameters β. Determining the probability of I specific ordered responses for each observation n is then the estimation challenge, thus an assumption on the distribution of error term is made to accomplish the task (Washington et al., 2010).

Proportional Odds Model

The most commonly used ordinal logistic model is proportional odds (PO) model. This model is based on cumulative probabilities while the continuation ratio (CR) model is based on conditional probabilities (Harrell, 2015).

The PO model is best stated as follows, for a response variable having levels 0, 1, 2, …, *k*:

$$\Pr\,[Y \geq j|X\,] = \frac{1}{1 + \exp[-(\alpha_j + X\beta)]}$$

where j = 1, 2, . . . , k. Some authors write the model in terms of Y ≤ j. When response category has two levels, this model is equivalent to the usual binary logistic model; therefore, this formulation with $Y \geq j$ makes the coefficients consistent with binary logit modle (Harrell, 2015).

In Proportional Odds model it is assumed that regression coefficients, β, are independent of j, the cutoff level for Y. The model has the assumption that the log odds that $Y \geq j$ is linearly related to each *X* and there is no interaction between Xs.

The proportional odds model is fitted using Maximum Likelihood Estimation on a pretty complex likelihood function that is dependent on differences in logistic model probabilities (Harrell, 2015).

2.4.5 Maximum Likelihood Estimation

Maximum likelihood (ML) which is used with a wide range of statistical analysis is a general technique of estimating parameters of a model given observations and it draws statistical inferences in different situations, especially nonstandard ones (Harrell, 2015). In this study in both of our models (Factor Analysis and Ordered Logit Model) this estimation method is used. The formula of the estimation is explained below from (Harrell, 2015).

Suppose in a population of individuals P is the probability of an event (Y) occurring. The probability of observing data in a sample size of n and observed responses of Y_1, Y_2, … , Y_n is given by

$$L = \prod_{i=1}^{n} P^{Y_i}(1 - P)^{1-Y_i}.$$

Let *s* be the number of "successes" which is the sum of Ys or the number of times that the event occurred (Y_i=1), thus n-s would be the number of "failures". The likelihood of the data can be simplified to

$$L = P^s(1-P)^{n-s}.$$

As working with log likelihood function is easier, for the one-sample binary response problem, the log likelihood is

$$\log L = s\log(P) + (n-s)\log(1-P).$$

What the ML estimation is trying to find is that value of P that maximizes L or $\log L$. The first derivative of $\log L$, with respective to P, is

$$\mathrm{U(P)} = \partial \log L/\partial P = s/P\ - (n-s)/(1-P).$$

$\mathrm{U(P)}$ is called score function and equating it to zero and solving it will give the value of P that maximizes the $\log L$.

In their book Dell´Olio et al. (2017) provide useful information and a general background on service quality of public transport. It explains how passenger's view is important, choice of variables, method of collecting data and finally how to measure service quality using data mining approaches.

2.4.6 Factor Analysis

Background

Factor analysis is a data-driven statistical model which is used to reduce the large number of attributes in a data set. It gives us fewer factors that describe the correlation among the large number of attributes we had in our data set and are actually unobserved (Tyrinopoulos and Aifadopoulou, 2008; Tyrinopoulos and Antoniou, 2008)**.** Normally after applying factor analysis other statistical models like logit models are used to discover more findings in the data set.

Understanding and analyzing public transport users' satisfaction and perception using factor analysis can be seen in the works of many researchers (Eboli and Mazzulla, 2013; Efthymiou et al., 2017, 2014; Fellesson and Friman, 2008; Ngoc et al., 2017; Tyrinopoulos and Antoniou, 2008)**.** Efthymiou et al. (2017) have applied this method to reduce the 23 indicators they have from their survey and reduce them to fewer factors that explain the underlying relationships between original factors. Their analysis results in three factors they named quality of service, service production/ transfer quality and ticket services. Each of these three factors explains more than 10% of the variance. Factor analysis in this example shows that the importance of service attributes in users' eyes is affected by these three factors.

There are different steps in doing factor analysis that should be considered. First of all we should make sure we are using the proper extraction method. Next things to be taken care of are, which correlation matrix should be used, how many factors to settle on, how the rotation of factors should be and finally (if needed) how to compute factor scores.

In this study factor analysis refers to Exploratory Factor Analysis (EFA) as there is also a confirmatory factor analysis. As the name suggest EFA is to explore data and its application is not to test hypotheses.

De Oña et al. (2015) uses a confirmatory factor analysis to assess the validity of the measurement model in a structural equation model. This classical factor analysis (CFA) is used to explore the relationships between different variables and the pre-chosen factors (Washington et al., 2010).

In factor analysis model the variables $X_i's$ with means μ_i's are expressed as linear functions and the formulation is as follows:

$$\begin{aligned}
X_1 - \mu_1 &= l_{11}F_1 + l_{12}F_2 + \cdots + l_{1m}F_m + \varepsilon_1 \\
X_2 - \mu_2 &= l_{21}F_1 + l_{22}F_2 + \cdots + l_{2m}F_m + \varepsilon_2 \\
\vdots \qquad & \qquad \vdots \qquad\qquad \ddots \qquad\qquad \vdots \\
X_p - \mu_p &= l_{p1}F_1 + l_{p2}F_2 + \cdots + l_{pm}F_m + \varepsilon_p .
\end{aligned}$$

Where, the factor analysis model in matrix rotation is given as

$$(\boldsymbol{X} - \boldsymbol{\mu})_{p\times 1} = \mathbf{L}_{p\times m}\mathbf{F}_{m\times 1} + \boldsymbol{\varepsilon}_{p\times 1}$$

where F's are factors, l_{ij}'s are the factor loadings, the ε_i's are error terms that are associated only with $X_i's$ and the *p* random errors and *m* factor loadings are unobservable or latent.

To solve these equations some restrictions are needed, because there are *p* + *m* unknowns and only *p* equations. The type of rotation that is chosen determines the type of model and the restrictions. Orthogonal or oblique rotations consider factor loadings close to zero or close to one and solve the equations with related conditions.

For example the following conditions are satisfied in an orthogonal model:

- **F** and $\boldsymbol{\varepsilon}$ are independent
- The mean of **F** and $\boldsymbol{\varepsilon}$ are zero $\begin{cases} E\,[\mathbf{F}] = \mathbf{0} \\ E\,[\boldsymbol{\varepsilon}] = \mathbf{0} \end{cases}$
- *COV* [**F**] = **I,** the identity matrix
- *COV* [$\boldsymbol{\varepsilon}$] = $\boldsymbol{\psi}$**,** the diagonal matrix

Rotation

EFA looks for the strongest correlations between variables and the factor. When the first best match is found it will be called Factor 1 and the search will be continued for the second set of correlations and so on. If thinking of a multidimensional variable space the axes of factors can be rotated within this space which suggest an important feature of EFA, rotation.

It is obvious that factors can also be correlated to one another and how to control this criterion depends on the type of rotation one uses in their analysis. Oblique and orthogonal rotations are two important techniqupes in factor analysis (Costello and Osborne, 2005).

Orthogonal rotations assume the factors are not correlated while on the other hand oblique rotations allow for correlation. Varimax is a type of orthogonal rotation and is the most common used among analysts. Quartimax and equamax are also orthogonal whereas direct oblimin, quartimin and promax are oblique methods (Costello and Osborne, 2005). Orthogonal results might be easier to be interpreted but this should not be a reason to choose it over oblique methods, and thus lose some valuable information where factors are correlated.

Factor Extraction Method

There are different techniques for factor extraction such as principal factors (principal axis factoring), unweighted least squares, generalized least squares, maximum likelihood, alpha factoring, and image factoring. On one hand there are different names for some of these methods in different textbooks or software packages and on the other hand advantages and disadvantages of each of them is not very easy to find in references which makes it difficult and confusing to choose the best extraction method. This might be a good reason to prefer using PCA instead of EFA (Costello and Osborne, 2005).

Fabrigar et al. (1999) argue that in general for relatively normally distributed data the best extraction method is maximum likelihood technique but if the data is significantly non-normal the method of principal axis factors is recommended. Harrell (2015) believes maximum likelihood as a more general objective function is a better solution when the variable is non-normal or polytomous. The distribution of data is the key indicator in choosing a method; for example, maximum likelihood is appropriate when there is multivariate normally distributed data, but lack of clear guidance for other methods arises the confusion in selecting a suitable method (Baglin, 2014).

In R the general command for factor analysis, factanal(), performs maximum-likelihood factor analysis but on the package of psych() the command fa() uses a minimum residual factoring method as default. In the later package the user can choose between various options of factoring techniques including weighted least squares and maximum likelihood.

Number of Factors

Scree test, Velicer's MAP criteria, Kaiser's criterion (eigenvalues greater than one) and parallel analysis are some alternative tests for factor retention (Baglin, 2014; Costello and Osborne, 2005; Velicer and Jackson, 1990).

When Kaiser's criteria retain all the factors with eigenvalue bigger than one, the scree test looks for a natural curve or break on the graph of eigenvalues (Costello and Osborne, 2005).

Baglin (2014) has some recommendations for doing factor analysis with ordinal data including choosing number of factors. They note that Kaiser's should not be used for deciding the number of factors. To get a sense of the number of dimensions scree plot can be helpful. They recommend to use parallel analysis-based methods as they perform better than Kaiser's criteria and scree plot. On the other hand Costello and Osborne (2005) indicate that despite the accuracy of parallel analysis there is a lot of difficulty in finding statistical software packages that have such methods and therefore, scree plot has been found more convenient.

Some analysts such as Washington et al. (2010) suggest doing a Principle Component Analysis to get an idea of how many factors to choose for EFA. Other propose considering one third of original variables to settle on without considering the actual underlying factors. Velicer and Jackson (1990) find the latter suggestion not valid as it is almost the same result that Kaiser's criteria would give.

As it was discussed above choosing the number of factors does not follow a very specific guideline and can depend on own judgements. There are some criteria that can help analyzers to determine how many factors to pick. Choosing fewer factors makes the work simpler while more factors gives a better fit.

Polychoric correlation matrix

The correlation matrices that are used in factor analysis (generally Pearson's method) assume that the variables are continuous but Holgado–Tello et al. (2010) argue that using polychoric correlations have more advantages and the results are more accurate to reproduce the measurement model when the model includes variables that are ordinal. Ordinal variables or ordered categorical variables are commonly used in many fields of science such as the social and health. The ordering of these type of variables is important and it is how their values can be compared (Ekstrom, 2011).

In the package of psych() the software R also suggests to use polychoric correlations rather than Pearson's for polytomous items. To do so in the command of fa() it should be added that cor="poly".

Factor score

Sometimes the result of explanatory factor analysis is used for further analysis such as some regression or predictive models. The researcher needs to compute scores to represent each participant's placement on the factors and then use them in investigating the research questions (Distefano et al., 2009).

Refined and non-refined methods are the two main classes of computing scores. Non-refined methods are easy to calculate and interpret. They are cumulative procedures to provide information about participants' placement on the distribution of factors. Refined methods on the other hand use more technical approaches; therefore, they are more complex but exact (Distefano et al., 2009).

Some methods of non-refined are 'sum scores by factor', 'sum scores above a cut-off value', 'sum scores- standardized variables' and 'weighted sum scores'. Sum scores above a cut-off value method for instance would only sums up the items with factor loadings above a threshold. This cut-off value is an arbitrary decision made by the researcher. When a loading is negative it would be subtracted in the computation (Distefano et al., 2009).

Regression scores, Bartlett scores and Anderson-Rubin Scores are three methods which are in the category of refined methods. Factor scores using these methods can be done using statistical software packages like SPSS, SAS and R (Distefano et al., 2009).

EFA or PCA

It may also be useful to take a quick look at another common method for data extraction named Principal Component Analysis which is sometimes confused with EFA.

Both are data reduction techniques; they allow you to capture the variance in variables and the output looks almost the same. These similarities have made the confusion between EFA and PCA. When the sample size is small or when there are not enough variables to associate factors the proper choice between these two becomes more critical (Velicer and Jackson, 1990).

Costello and Osborne (2005) and Widaman (1993) like many others argue that PCA as a very common extraction method is not a true method of factor analysis and should be used with a lot of caution and restrictions; however, some researchers suggest that these two methods do not vary and some other ones say PCA is preferable. Costello and Osborne's (2005) findings show EFA is the optimal method comparing to PCA.

In order to reduce data, PCA uses a linear combination of a set of variables to create index variables (components) from a larger set of variables; EFA, on the other hand, applies a fundamentally different approach for data reduction.

EFA is a model of a latent variable that cannot be measured directly using a single variable but through the relationships it causes among a set of variables (Washington et al., 2010). In a simpler way the difference can be interpreted as PCA tries to find the component that is caused by a set of variables but EFA looks for the factor that a set of variables are caused by it.

3 Data Analysis and Processing

3.1 Surveys and data sets

To study the effects of economic crisis on public transport users' satisfaction and demand in Greece the survey that was conducted in 2008 was done again in 2013 and 2017. Tyrinopoulos and Antoniou (2008) used a questionnaire for the cities of Athens and Thessaloniki which was most based on the Handbook for Measuring Customer Satisfaction and Service Quality (TRB, 1999) and 23 qualitative and operational service parameters were defined. The respondents were asked to give their opinion about these attributes both in terms of satisfaction and importance. Beside the section of quality assessment of the service the questionnaire had a part for demographic questions and another part for travel patterns such as frequency of using the transit or purpose of the trip.

Efthymiou et al. (2014) and Efthymiou and Antoniou (2017) conducted the same questionnaire with minor changes and modifications so that the direct comparison of the data sets were allowed and the results could be compared**.** Zuo (2017)**,** trying to study the same topic in Munich used the Greek questionnaire and made some changes as she also wanted to investigate the impact of weather conditions on users' choices of transit.

Having both data sets in Athens and Munich in 2017 it was needed to go deeper into the surveys of both cities and pull out the most comparable questions and results in order to have two data sets which allowed direct comparisons. Here we will take a look at the questionnaire that was done as face to face interviews of passengers in Athens and then show you the most relevant parts of the both surveys in Athens and Munich. The last part of the questions that is about the changes in users' ridership comparing to past years is omitted. The first part of the survey in Athens 2017 tries to collect some demographic information about the users which are gender, age, occupation, driving license ownership and car ownership details. In the second part some more mobility patterns are asked: frequency of using the transit, usual ticket types the user buys, main purpose to use public transport, usual time slots of using bus or metro and the main reason they use public transport.

In the third section questions about service quality is included. For the questions about quality of the service the respondents had to answer both how satisfied they are about an attribute of the service and how important each attribute is to them. They had to choose their answers from (1) not satisfactory at all/ not important at all, (2) not satisfactory/ not important, (3) moderate, (4) satisfactory/ important and (5) very satisfactory/ very important. The 25 selected attributes were classified into four categories. Definitions are from (Efthymiou et al., 2017).

- **General characteristics of the public transit system**

1. Service frequency: refers to the frequency of the service in the lines of the transit systems.

2. On-time performance: refers to the accuracy of the departure times of the vehicles at terminal stations in relation to the predefined schedule.

3. Service provision hours: refers to the operating hours of the service provision on a given day.

4. Network coverage: refers to the spatial coverage of the area under consideration with public transit services.

5. General information provision: refers to the sufficiency of the information provided to the passengers about the general characteristics of the transit services, such as the lines, terminals and stops points, departure times, tickets and passes available.

6. Types of tickets and passes: refers to the sufficiency of the various available types of tickets and passes with respect to the coverage of the needs of the public.

7. Prices of tickets and passes: refers to the price-structure of the various types of tickets and passes available.

8. Tickets selling network: refers to the sufficiency of the tickets selling network and the ease to purchase tickets from the various selling points.

9. Personnel behavior: refers to the behavior of the various types of personnel of the transport operator (e.g., drivers, station officers and ticket counter officers), when communicating and transacting with the passengers.

10. Existence of bus lanes: refers to the sufficiency and performance of the bus lanes to facilitate the efficiency of the transit service.

11. Measures for environmentally friendly public transit: refers to the contribution of public transit in the protection of the environment and the adequacy of the relevant actions and measures taken by the relevant authorities.

- **Terminals and stops**

12. Walking distance to terminals and stops: refers to the distance that passengers have to walk from the origin point to the closest terminal and stop.

13. Information provision at terminals and stops: refers to the sufficiency of the information available to the passengers about the services provided at the terminals and stops.

14. Conditions at terminals and stops: refers to the conditions of the terminals and stops concerning shelter, visibility, seating capacity, etc.

15. Safety at terminals and stops: refers to the perceived sense of safety of the passengers when waiting at the terminals and stops to use the public transit service.

16. Announcing the arrival of next vehicle from variable messages

- **Vehicles**

17. Onboard conditions: refers to the conditions inside the vehicle during the execution of a journey, mainly concerning crowded situations and the provision/condition of available facilities (e.g., seats and air-conditioning).

18. Vehicles cleanliness: refers to the level of cleanliness of the vehicles from various standpoints (seats, handles, windows, doors, floor, etc.).

19. Driving behavior: refers to the driving performance of the vehicle's driver.

20. Onboard information provision: refers to the provision of information inside the vehicle during the trip, such as next stop and estimated arrival time at the next stop.

21. Accessibility to disabled and mobility impaired people: refers to the provision of facilities by the transit operator to facilitate the accessibility of transit services by disabled and mobility impaired people.

22. Onboard safety

- **Transfer points**

23. Distance between transfer points: refers to the distance that passengers have to walk between transfer points in order to continue their trip.

24. Waiting time at transfer points: refers to the time that passengers have to wait at transfer points in order to continue their trip.

25. Information provision at transfer points: refers to the provision of information to passengers at the transfer points about the combination of the various lines and modes, and their time schedules.

Starting from the socio economic and mobility patterns parts some of the questions had to be removed. For example what was asked in Munich about driving license, car ownership or regular ticket bought was missing in questionnaire of Athens. In the question of asking for the main reason of using transit there were some different alternatives in answers of each city which made it too difficult for comparison and this question was also omitted.

With very slight changes the question about "occupation" was kept. As you can see in Table 1 two pairs of answers had to be merged in order to have comparable results about job status. Similar approach was held for the question of "purpose of trips by public transport", but as in Athens respondents could only choose one answer whereas in Munich the people had been given the chance to choose different options this question could also not be kept.

Athens	Munich
• Student	• Student
• employee	• full time job
• Self employed	• part time job
• Retired	• self employed
• Unemployed	• retired, unemployed, stay at home
• other	• other

Table 1 Occupation status alternatives in questionnaire

Before getting to the indicators in both questionnaires we will take a look at the question of overall satisfaction that is an important question and will be treated as a dependent variable in one of later analysis. The question in both cities is the same; however, the problem arises in the alternatives given to respondents to choose from where in Athens there are four choices and in Munich there are five. Table below shows a more clear representation of the question. The proportion of each answer in different cities is given in brackets.

Question: How satisfied are you with the overall public transport service?

Athens	**Munich**
(1) Not satisfactory at all (8%)	(1) Totally not satisfied (1%)
(2) Moderate (35%)	(2) Not satisfied (9%)
(3) Satisfactory (45%)	(3) Neutral (27%)
(4) Very satisfactory (12%)	(4) Satisfied (52%)
	(5) Totally satisfied (11%)

Missing the alternative “Not satisfied” in Athens might have some effects on respondents to choose “Moderate” over “Not satisfactory at all” which is a more positive approach but we do not know to what percentage the answers would split between the neighboring options. On the other side in the city of Munich the answer to option 2 is not that high and for our first scenario the Munich data from answers “Totally not satisfied” and “Not satisfied” have been merged. Therefore, for this specific question there would be 4 likert scale data instead of 5 for both cities of Athens and Munich.

As mentioned below the biggest part of the questionnaires were dedicated to transit attributes and respondents had to select between 5 likert scale answers ranging from 1 (not satisfactory at all/ not important at all) to 5 (very satisfactory/ very important). The changes and modifications in the questionnaire of Munich lead to complete eliminations of some indicators but one indicator was kept using some merging approaches. The attributes that are the same in both cities are listed below:

- On board cleanliness
- On board information
- Safety on board
- Driving manner (only for bus)
- Safety at stations while waiting
- Information at stops
- Network coverage
- Distance/ time travelling to stops (stop accessibility)
- Ticket and card price
- Ticket type
- Distance travelled while transfer
- The waiting time at the point or at the points of transfer
- Accurate timetables (punctuality)
- General information provision

Frequency of service, waiting time and temperature are three of most important factors that Zuo (2017) has found in her analysis. Temperature was one very main focus of her thesis which was not the case in Athens, thus it has not been a concern to omit it from the list. Waiting time which was in the questionnaire of Munich was not mentioned in the other questionnaire but “Punctuality” reflects the same issue of how long people have to wait or “Frequency” also shows how long the headways and waiting times between every two runs is. Hence, it would not be a problem not to have “Waiting time” on our list.

“Frequency of the service” on the other hand would be of more concern and the attempt has been made to keep this indicator in both cities. There has been only one general indicator in Athens regarding frequency but in Munich it has been asked through two attributes, frequency of service on peak hours and frequency of service on off peak hours. The idea has been to combine and make a normal average of the two answers for each observation in Munich but some statistical tests is needed to make sure that to some extend the answers are relevant and come from one population.

Among many available statistical tests t-test was chosen. Mann-Whitney-Wilcoxon Test is known to be a test when dealing with categorical or likert scale data but the results are not really different to a t-test for our example. The t-test had to be done on four pairs of answers while respondents answered questions about satisfaction and importance regarding bus or metro.

With 169 observations for bus service and 176 observations for metro service the p-values of importance of the indicators are less than the alpha ($p < 0.05$) while the p-values of satisfaction of the service are less than the alpha of 0.1. The decision has been made to compute the mean of frequency at peak hour with off-peak hour; however, the impact of this decision on the output and further analysis should be into consideration.

3.2 Data overview

In this section the sample data from both surveys in Athens and Munich is explored. Some more insight of the data is also provided in which the respondents' overall satisfaction of the service influenced by age, gender, frequency of using the service and occupation status is explored. This section is based on simple data mining and visualization. Chi square test has been used for one clustered set.

A summary of the sample characteristics is illustrated in Table 2. In Athens people who have been using bus are surveyed separately than the users of subway; therefore, there are two different columns on the table for this city.

In Athens in total 250 questionnaires have been completed by users of bus service and 313 ones by metro riders. In Munich there have been 177 respondents who gave their ideas on both bus and metro service at the same time. However, some uncompleted observations have been removed and finally 169 inquiries have been used for analysis of data in Munich.

City	Athens				Munich	
PT mode	Bus		Metro		Bus = Metro	
	n	%	n	%	n	%
Total	250	100	313	100	177	100
Gender						
Male	121	48.4	118	37.7	104	58.8
Female	129	51.6	195	62.3	73	41.2
Age						
=< 18	16	6.4	7	2.2	9	5.1
19-23	104	41.6	35	11.2	22	12.4
24-29	40	16.0	69	22.0	56	31.6
30-39	27	10.8	88	28.1	36	20.3
40-49	28	11.2	71	22.7	22	12.4
50-59	18	7.2	32	10.2	20	11.3
>= 60	17	6.8	11	3.5	12	6.8
Occupation						
Employee	71	28.4	122	39.0	91	51.4
Self employed	28	11.2	45	14.4	4	2.3
Student	106	42.4	88	28.1	58	32.8
Retired/ unemployed	42	16.8	46	14.7	15	8.5
Other	3	1.2	12	3.8	9	5.1
Frequency of using public transport						
Daily	118	47.2	188	60.1	112	63.3
Often	70	28.0	74	23.6	34	19.2
Occasionally	41	16.4	40	12.8	24	13.6
Rarely	21	8.4	11	3.5	7	4.0
Overall user satisfaction about PT service						
Very satisfied	5	2.0	65	20.8	20	11.3
Satisfied	92	36.8	165	52.7	93	52.5
Neutral	116	46.4	79	25.2	48	27.1
Not satisfied/at all	37	14.8	4	1.3	16	9.0

Table 2 Respondent profile

Gender: male respondents are less than female riders in Athens while in Munich they are more than female respondents. In Athens this difference is more highlighted among metro users where 37.7% of respondents are male and 62.3% are female. In Munich on the other hand males are about 20% more than females. Although these differences might have some reasons beneath, the sample size especially in Munich is small and it does not allow for further judgments.

Age: more than 60% of respondents at bus stops in Athens are people younger than 30 years old. The age distribution for metro users in this city follows a more normal pattern. In Munich this proportion is about 50%.

Gender and age groups: distribution of male and female respondents by age is shown in figures 2 and 3 below. In Athens male and female have almost the same age distribution. For age group 50-59 females are 12% of their gender group while males are only 5% of all the males. In Munich 36% of males are between 24 and 29 while females have 10% less proportion of this age group. Like in Athens female respondents between 50 and 59 years old are more than male ones.

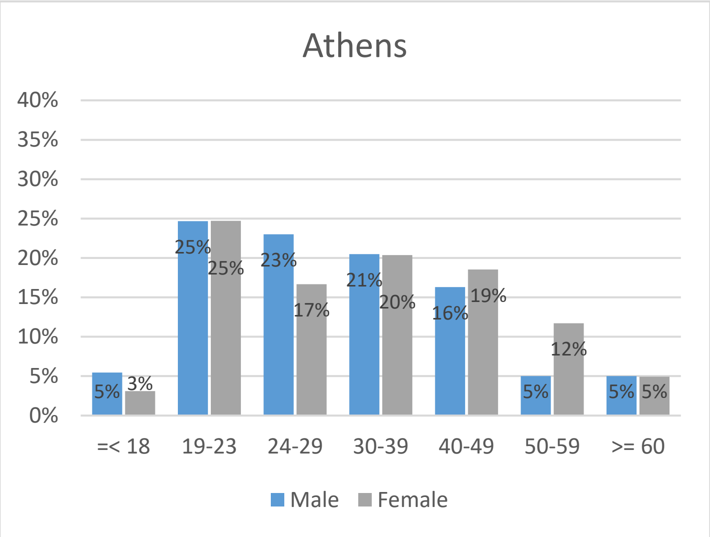


Figure 2 Distribution of each gender by age groups in Athens

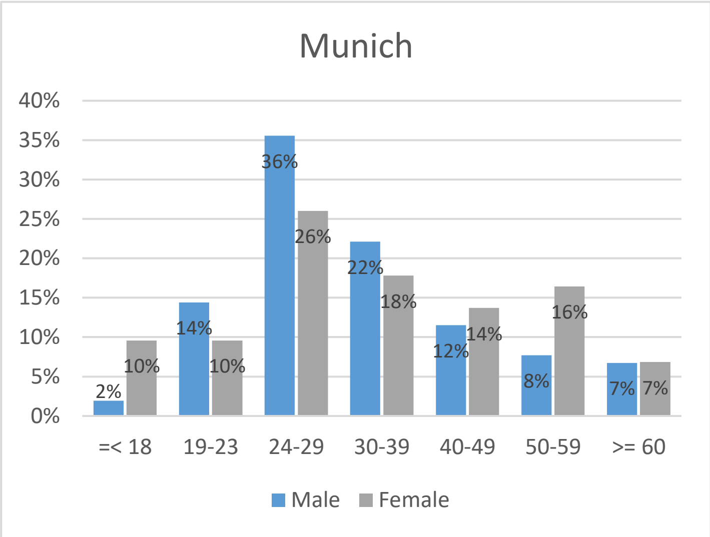


Figure 3 Distribution of each gender by age groups in Munich

Occupation: in both cities two major occupation is being an employee and being a student. In Munich about half of the sample have a full time job and about 33% are students. These numbers can be related to the age distribution of the respondents. In Athens metro users have bigger proportion of employment status while bus users have higher number of students in their sample. The other interesting proportion is the number of retired/ unemployed respondents in Athens comparing to Munich. Economic situation of each country and the cities could explain this figures.

Frequency of use: in both cities and among both transit modes most of the survey respondents are using the service every day or often. This is to be expected as the number of people that use a service more is normally higher than the ones who use it less frequently.

Overall user satisfaction about PT service: generally people would not choose extreme answers when asking about their opinions and thus the alternatives such as very bad or very good have lower frequencies comparing to good, bad or just okay. In Athens; however, 20% of the metro users are very satisfied with the overall service while only 2% of bus riders feel the same. On the other extreme about 15% of bus users are very unsatisfied with the overall transit service in Athens but less than 2% of metro users have chosen this answer. One conclusion from this raw data is that the bus service users are generally less satisfied comparing to metro users in Athens. In Munich on the other hand the figures sound more normal.

Figure 4 shows that in Athens respondents who have been using bus are less satisfied with the whole transit service. In contrary to bus users, passengers of metro indicate a higher satisfaction to the overall public transport system. A general conclusion from this comparison is that the quality of metro service is better than bus and the remarkable finding could be that how satisfaction changes the overall opinion of them about the service. When a person feels happy about the part of the service that she uses, she considers other sections to be as good as well and vice versa.

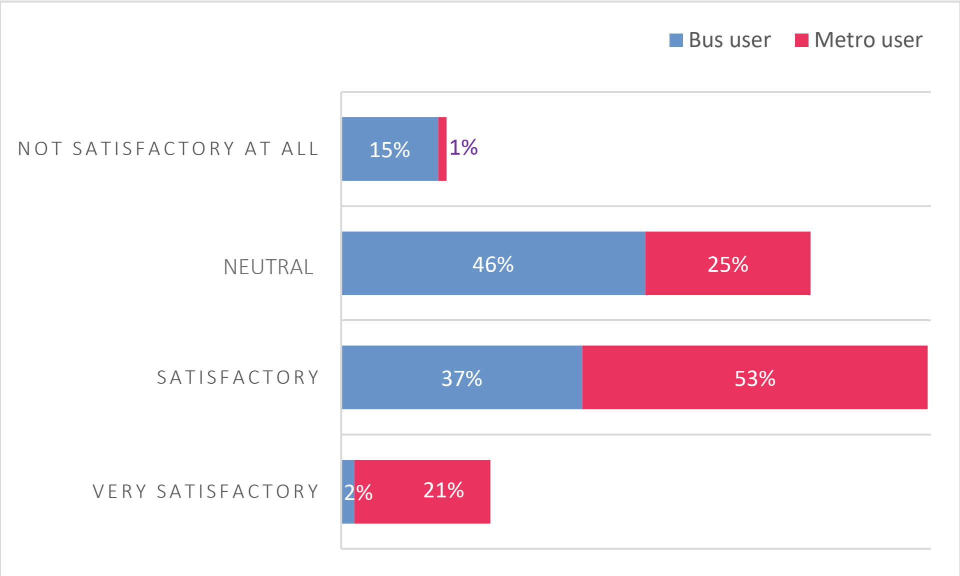


Figure 4 Overall satisfaction by different PT users in Athens

This comparison is not possible to be done in Munich as the respondents have not been distinguished regarding the mode they use and bus or metro users are not specified. But the respondents have been asked two other questions regarding their satisfaction about bus service and also metro service. The results are shown in Table 3. The difference at mid and lower levels is very slight but the highest level of satisfaction is lower about bus service in Munich. It is to be note that survey has been taken place in metro stations.

Satisfaction level	very satisfactory	satisfactory	neutral	not satisfactory/ at all
Bus service	14%	43%	28%	15%
Metro service	21%	44%	24%	12%

Table 3 Satisfaction level of bus and metro service in Munich

Overall satisfaction by different age groups

Figure 5 illustrates the overall satisfaction by different age groups comparing the two cities. People of age 24-29 and 40-49 years old in Munich seem to be more satisfied while the youngest age group in this city show the least satisfaction. It is to be noted the number of respondents younger than 19 years old are very low in number and this results cannot be easily generalized. In Athens the older age group are the least satisfied.

Overall satisfaction by different occupation groups

Figure 6 shows the overall satisfaction among different job holders between cities of Athens and Munich and on the table below you can see the distribution of occupation groups in each city.

Occupation	Employed	Self-employed	Student	Retired/ Unemployed	Other jobs
Athens-PT	0.34	0.13	0.34	0.16	0.03
Munich-PT	0.52	0.02	0.33	0.08	0.05

Table 4 Occupation proportion of Respondents, Athens and Munich

Students and employed respondents in Munich have a higher level of satisfaction from the service they are using. They together are the majority of the sample population regarding occupation status with 85% percent of proportion. 25% of self-employed people in the German city claimed they were very satisfied with the service; they are only 2% of the whole sample though.

Self-employed and retired or unemployed transit riders in Athens sample are less satisfied with their service and the employed respondents are the most satisfied group in this city. Students reveal a normal distribution of satisfaction in Athens.

Overall satisfaction regarding frequency of transit use

Figures 7 and 8 illustrate the groups with different frequency of using public transport and how satisfied they are with the service as a whole. Each figure represents one of the cities. For the city of Athens different transit riders (bus and metro) are separated.

In general for both cities and all the modes respondents are more satisfied or neutral than being dissatisfied or very satisfied which are the two extremes. In Munich the people who use the transit often seem to be less happy with the service than their other neighbors.

The rarely bus users in Athens are the least satisfied group in this comparison. It is interesting to notice that the least frequent metro users in Athens do not show dissatisfactory about the overall service which indicates how well the image of this transit mode is. On the other hand, the bus users who take the bus very rarely show a high dissatisfaction about transit service.

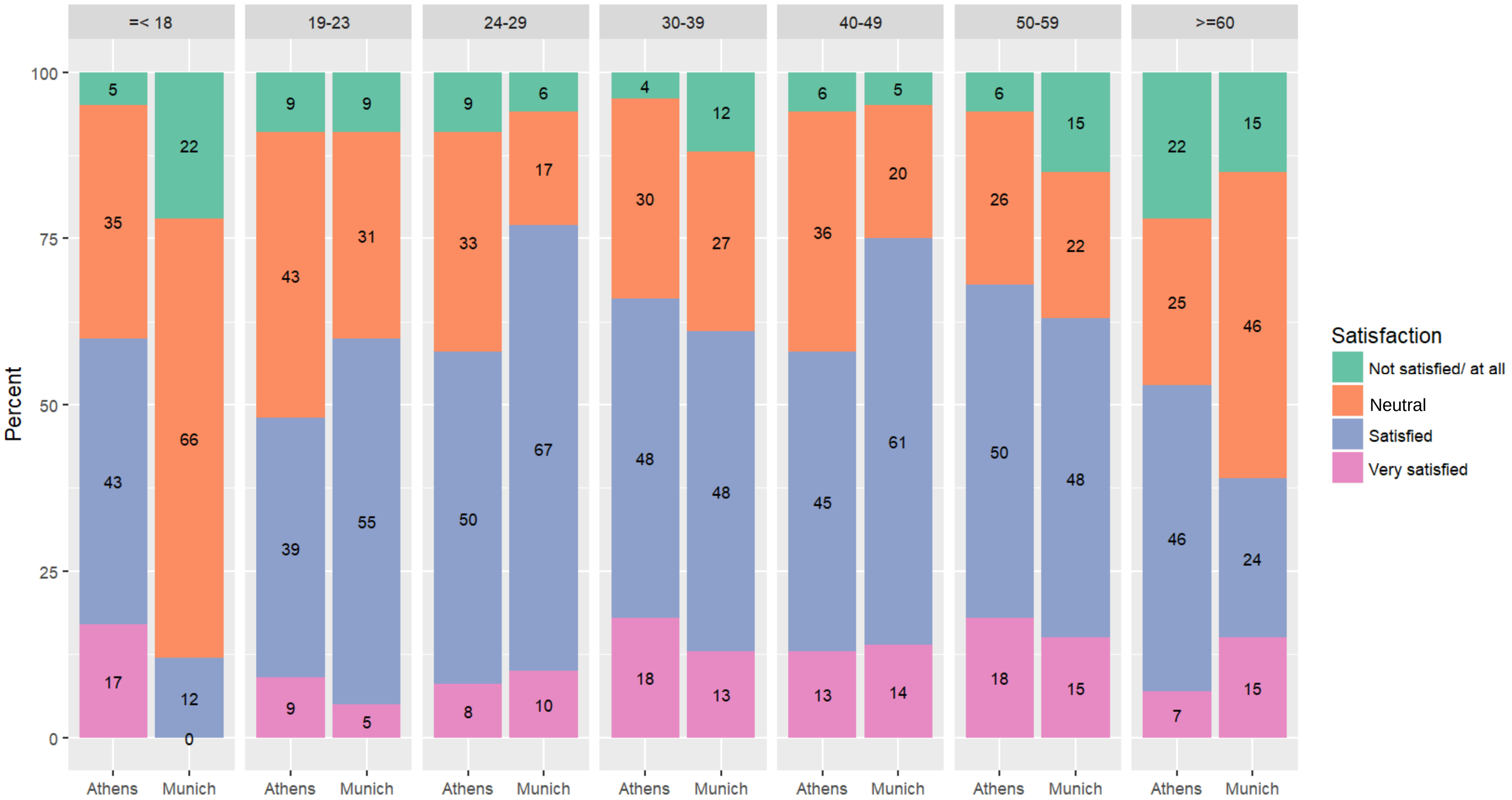


Figure 5 Comparison of overall satisfaction among different age groups between Athens and Munich.

Chi square test p value for all groups smaller than 0.05 except [Munich 50-59] which is 0.14

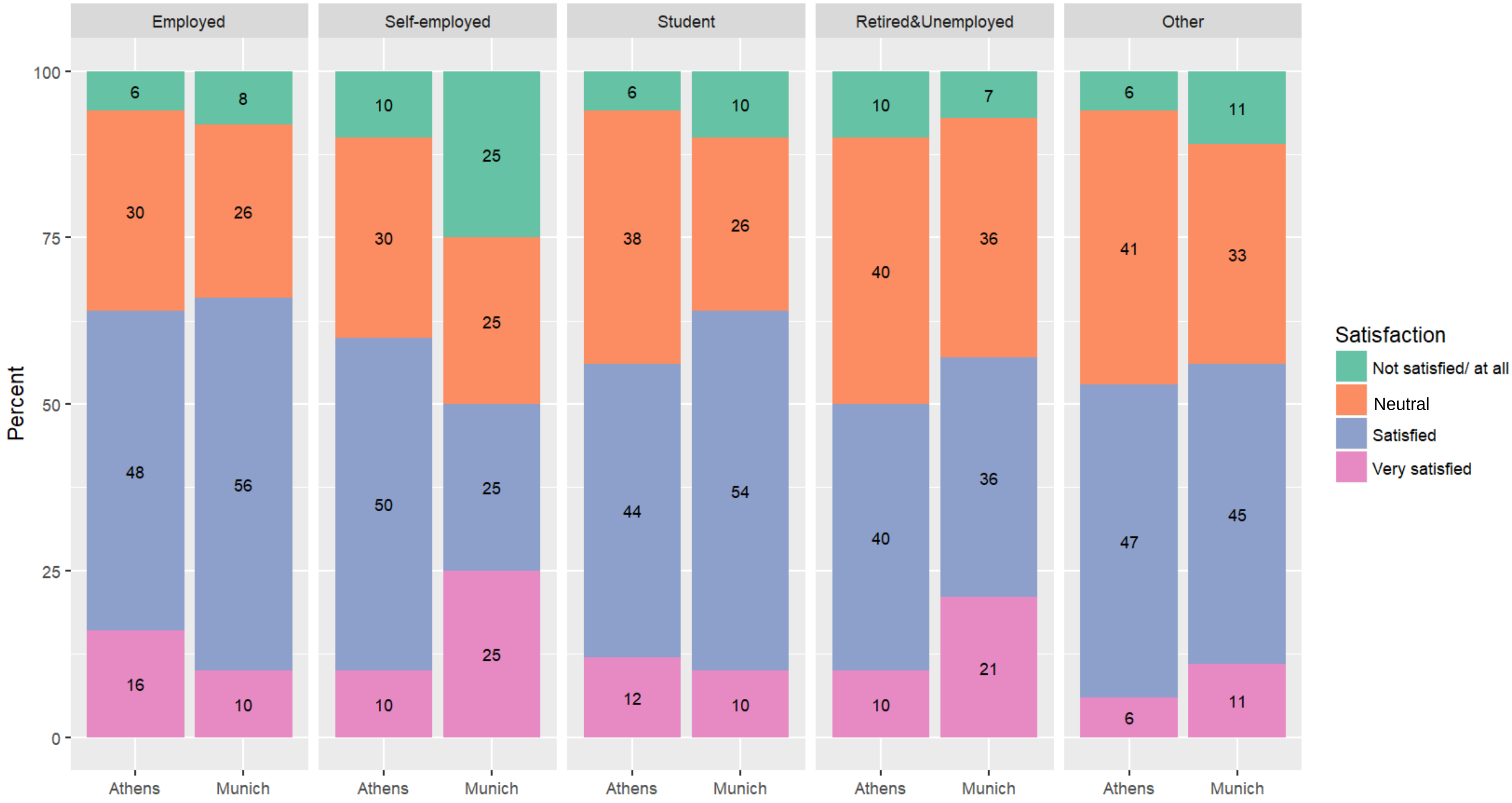


Figure 6 Comparison of overall satisfaction among different occupation groups between Athens and Munich.

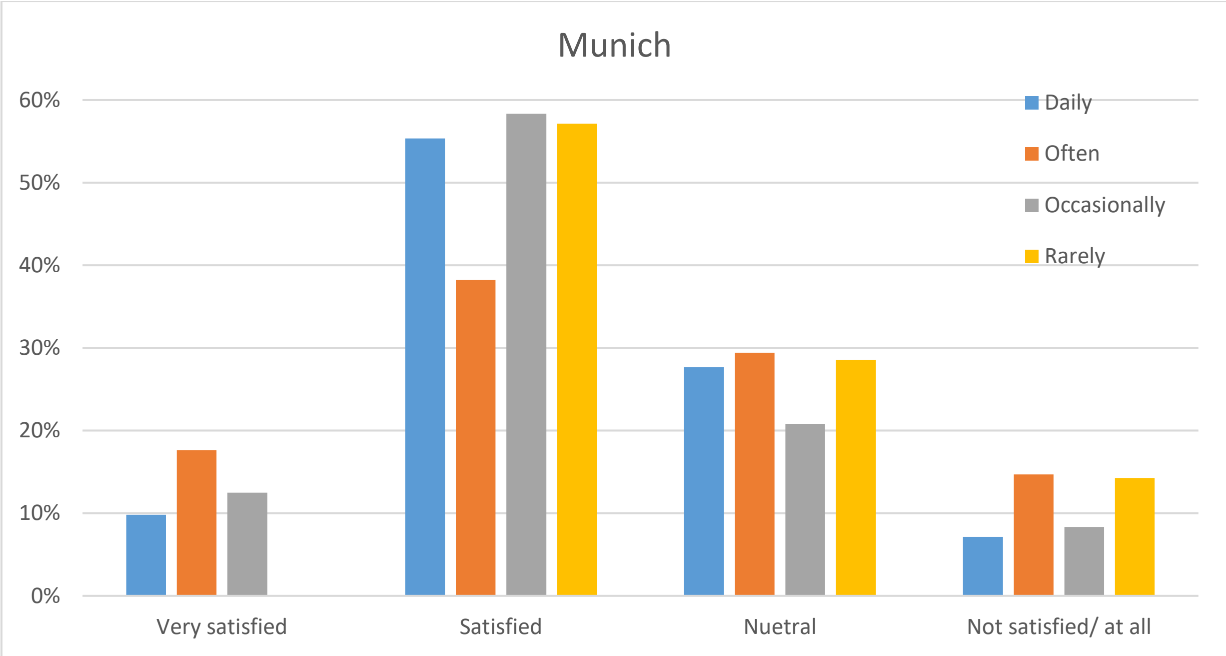


Figure 7 Overall Satisfaction Regarding Frequency of Use, Munich

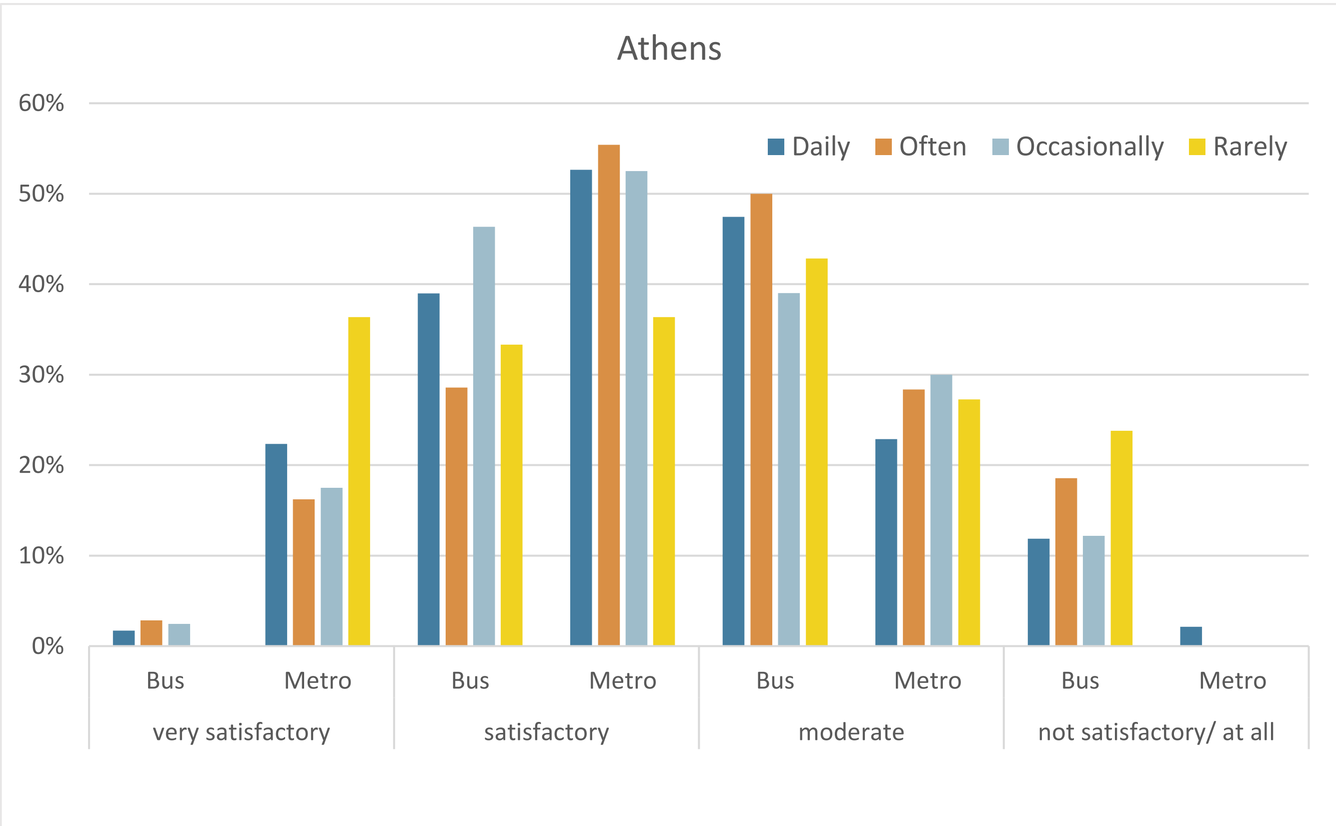


Figure 8 Overall Satisfaction Regarding Frequency of Use, Athens

Quality attributes data

As mentioned in first section of this chapter 15 quality indicators for bus and 14 indicators for metro have been chosen for further analysis. Table below shows the list of abbreviation and their definition used in this study for each of these attributes.

Shortened form	**Explanation**
Frequency	Frequency of service
Punctuality	On time performance of the service
Coverage	Network coverage of transit lines
Info	Availability of general information of service for passengers
Tickettype	Variability of ticket type
Price	Current ticket price
Distancetostop	Distance between stops and home/destination
Stopinfo	Timetable and line information at stops
Stopsafety	Safety of passengers at stops
Cleanliness	Cleanliness and hygiene on board
Driver	Driving style of bus drivers
Onboardinfo	Availability of information on vehicle
Onboardsafety	Safety of passengers on board
Transferwalk	Distance needed to walk while transfer
Transferwait	Waiting time at stops while transfer
Note: letters “I” and “S” in front of any of the shortened phrases refer to “Importance of the indicator” and “Satisfaction level toward the indicator” respectively.	

Table 5 Abbreviations used for indicators

Dealing with categorical and ordered data needs different considerations than continuous data. Finding out that which statistical illustration or test gives a better understanding of the data set is a bit complicated, since there is not an exact or certain literature and recommendations on the topic. The other fact that makes finding the best test and illustration more difficult is the nature of our data which is respondents’ opinion and perception about a service. The frequency diagrams do not always look normally distributed.

Figure 9 is two examples of frequency diagrams for two indicators of importance of network coverage and satisfaction about on-board safety on buses in Athens. The same diagrams had been made for all of the indicators, but because a useful conclusion could not be driven they are not used in this study. However, some general statistical description of the indicators such as the mean, standard deviation and skewness is given in table 6. The rest of the indicators for the city of Munich can be found in Appendix A.

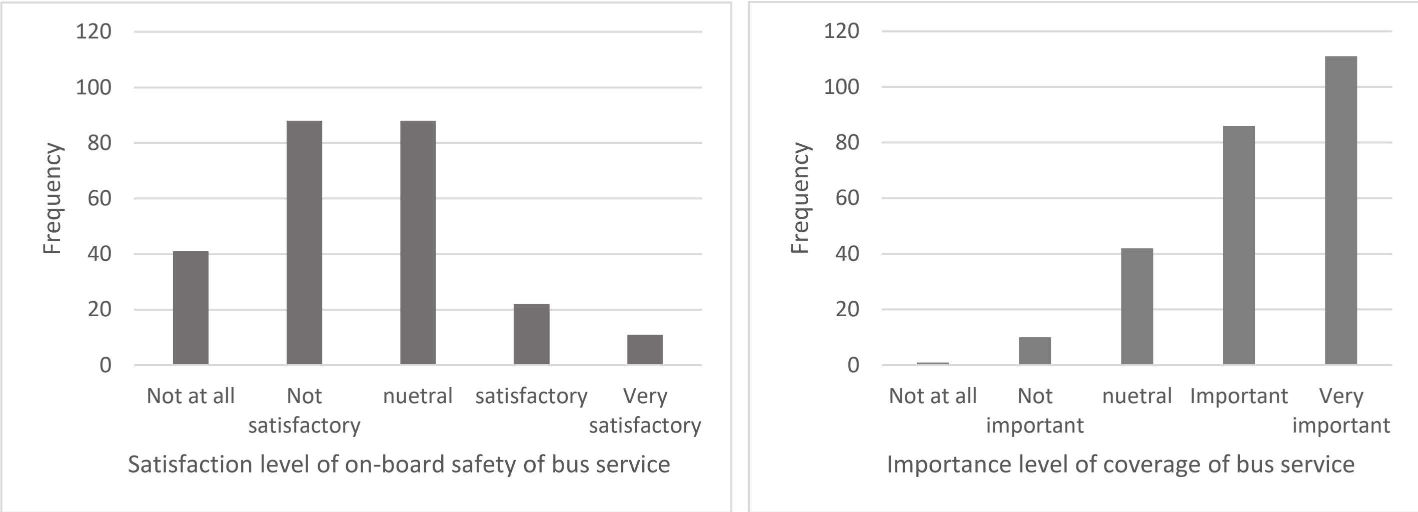


Figure 9 Frequency diagrams for two indicators of quality

Taking a look at medians it shows that using a box plot for such data is not appropriate and it is not correct to remove any data considering them as outliers.

Skewness is a measure of the asymmetry and kurtosis is a measure of tailedness (or peakedness) of the probability distribution of a variable (Kim, 2013). Doing tests on them can clarify the normality of distribution to some extent but this is not the case for our data. (Seltman, 2015) looking at skewness and kurtosis for quantitative variables and possibly for ordinal variables is useful but for categorical variables they do not make sense (Seltman, 2015).

What mentioned above about the frequency distribution of our data set shows the same problem about skewness or kurtosis of data. We can use and explain them but we cannot reject some data for the reason that its kurtosis value is bigger or less than the recommended thresholds. Some recommend acceptable values would be between ± 1 but some others suggest bigger spans such as ± 2. Despite IPuncuality, IFrequency in metro of Athens and IPunctuality in Munich, the rest of the data have acceptable values. These three should be taken into consideration when dealing with or interpreting the output in further analysis.

By looking at the data it is seen that in general satisfaction indicators have lower mean and median values comparing to importance of the quality indicators. Many median values for importance of the services in Athens are 5 (very important) but they are mostly 3 (neutral) for satisfaction data.

Correlation among indicators and also between indicators and other characteristics of sample such as age or occupation is calculated and plotted. All the illustrations are added to this report as Appendix B. Correlation of the explanatory variables before doing most of the statistical analysis such as regression models is recommended but it is not that clear which method to use for categorical, likert or ordered data. The correlation among our data set is in an acceptable range but for the reason mentioned I do not relate to them for further analysis.

	Athens Bus (250 observations)					Athens Metro (313 observations)				
Importance	mean	sd	median	skew	kurtosis	mean	sd	median	skew	kurtosis
Frequency	4.43	0.89	5	-1.47	1.13	4.53	0.82	5	-1.95	3.85
Punctuality	4.37	0.78	5	-1.04	0.35	4.56	0.77	5	-1.88	3.52
Coverage	4.18	0.88	4	-0.85	0.06	4.19	1.15	5	-1.46	1.26
Info	4.01	0.91	4	-0.57	-0.13	4.09	1.12	4	-1.19	0.66
Tickettype	3.42	1.08	3	-0.08	-0.76	3.77	1.17	4	-0.69	-0.31
Price	4.15	0.89	4	-0.90	0.40	4.15	1.20	5	-1.29	0.59
Distancetostop	3.93	0.85	4	-0.49	-0.02	4.17	0.96	4	-1.02	0.48
Stopinfo	4.38	0.81	5	-1.54	2.88	4.19	0.95	4	-1.16	0.97
Stopsafety	4.08	0.95	4	-1.11	1.16	4.29	1.07	5	-1.50	1.44
Cleanliness	4.44	0.76	5	-1.80	4.53	4.30	1.06	5	-1.53	1.58
Driver	4.28	0.78	4	-0.98	1.21	-	-	-	-	-
Onboardinfo	3.70	0.92	4	-0.51	0.12	4.11	0.93	4	-0.88	0.24
Onboardsafety	4.39	0.83	5	-1.41	1.67	4.28	1.03	5	-1.41	1.24
Transferwalk	4.00	0.82	4	-0.31	-0.76	3.88	1.08	4	-0.84	0.28
Transferwait	4.08	0.81	4	-0.51	-0.41	4.08	0.93	4	-0.86	0.31
Satisfaction										
Frequency	2.63	0.97	3	0.14	-0.25	3.60	0.95	4	-0.57	0.10
Punctuality	2.80	1.03	3	0.05	-0.52	3.85	0.99	4	-0.80	0.37
Coverage	3.18	1.14	3	-0.24	-0.67	3.20	1.33	3	-0.31	-1.10
Info	3.25	0.93	3	-0.03	-0.33	3.45	1.01	4	-0.56	-0.10
Tickettype	3.31	0.95	3	-0.23	-0.16	3.19	1.06	3	-0.41	-0.49
Price	3.05	1.11	3	-0.06	-0.64	2.79	1.22	3	0.02	-1.00
Distancetostop	3.15	0.99	3	-0.23	-0.28	3.49	1.02	4	-0.61	-0.04
Stopinfo	3.11	1.00	3	-0.27	-0.50	3.72	0.92	4	-0.63	0.31
Stopsafety	2.61	1.05	3	0.29	-0.40	3.43	1.07	4	-0.53	-0.29
Cleanliness	2.36	1.01	2	0.51	-0.21	3.21	1.13	3	-0.35	-0.65
Driver	3.07	0.98	3	-0.35	-0.25	-	-	-	-	-
Onboardinfo	3.04	0.97	3	-0.22	-0.31	3.66	0.95	4	-0.63	0.24
Onboardsafety	2.50	1.01	2	0.43	-0.08	3.27	1.09	3	-0.46	-0.39
Transferwalk	3.05	0.89	3	0.01	0.00	3.64	0.91	4	-0.58	0.41
Transferwait	2.64	0.91	3	-0.07	-0.42	3.37	0.95	3	-0.47	0.15

Table 6 some statistical information for quality indicators in data set of Athens for both bus and metro service and importance or satisfaction indicators

4 Modeling Framework

4.1 Factor Analysis

The software package R offers different packages and commands to do different types of factor analysis. This study benefits from the package psych() and the command fa() for its analysis in this section. Maximum likelihood extraction method has been used for factor analysis in our study. In this section the details of factor analysis procedure is explained.

The polychoric correlation has been chosen to be used in this work. However, for some sets of data both correlation methods were used to see the differences. Although very small difference in magnitude of some loadings of variables could be observed but these differences did not change the final decisions on choosing factors. In other words when applying polychoric correlations we had pretty much the same loadings and exactly the same factors were determined.

Below you can see one example from importance of bus service indicators in users' opinions in Athens in Table 7. Loadings below 0.3 has been removed and more than 0.6 have been highlighted. Table on the right benefits from polychoric method and it has been noticed the loadings are a bit stronger. As mentioned above in our work the final result and decision is not affected by using different correlation matrices though.

Indicators	Factor1	Factor2	Factor3
IFrequency		**0.70**	
IPunctuality		**0.64**	
ICoverage		0.54	0.33
IInfo	0.43		
ITickettype		0.42	
IPrice	0.30		
IDistancetostop			0.46
IStopinfo	0.54	0.33	
IStopsafety	**0.63**		
ICleanliness	**0.62**		
IDriver	**0.62**		
IOnboardinfo	0.45		
IOnboardsafety	**0.71**		0.33
ITransferwalk			**0.63**
ITransferwait			0.49

Indicators	Factor1	Factor2	Factor3
IFrequency		**0.83**	
IPunctuality		**0.73**	
ICoverage		**0.60**	
IInfo	0.45		
ITickettype		0.41	
IPrice			
IDistancetostop			0.46
IStopinfo	0.46	0.3	
IStopsafety	**0.60**		
ICleanliness	**0.61**		
IDriver	**0.79**		
IOnboardinfo	0.48		
IOnboardsafety	**0.82**		
ITransferwalk			**0.67**
ITransferwait			0.49

Table 7 Factor loadings of importance indicators of bus service in Athens. Left: using default correlation method with factanal() command, Right: Factor loadings using polychronic correlation with fa() command

In this study factor analysis for different number of factors has been done and comparison between the results and loadings and some of the tests such as parallel analysis have brought into consideration to select the best number of factors with each analysis. Example below that is for the data set of Importance of service indicators about bus in Munich will explain how four factors seemed to be the best choice.

A quick PCA in R gives figure 10 in which two components have proportion of variance above 10% but the next two ones are very close to 10%; therefore, two to four factors might give the best result according to this approach.

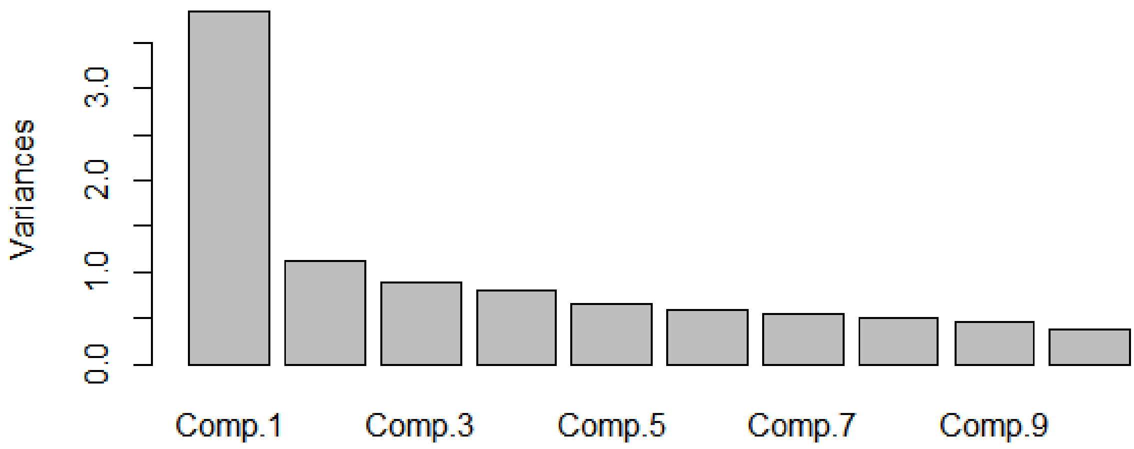


Figure 10 Principle component analysis output for Importance indicators of bus service in Munich

A parallel analysis with Scree plot using eigenvalues in R suggest four factors to be optimal. Package nFactors can be used for this analysis. Number of components/factors to retain according to optimal coordinates, parallel analysis and the Kaiser rule give four factors for this data set. According to the acceleration threshold only one component or factor is optimal.

Now let us take a look at the results for two to five number of factors which are shown in Table 8.

	2 Factors		3 Factors			4 Factors				5 Factors				
Indicator	Factor 1	Factor 2	Factor 1	Factor 2	Factor 3	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
IFrequency	**0.66**		0.67			**0.77**								**0.62**
IPunctuality	**0.72**		**0.73**			**0.69**							0.33	0.50
ICoverage	**0.66**		**0.72**			0.32	**0.61**				**0.69**			
IInfo	0.48		0.36				**0.74**				0.59			
ITickettype		**1.01**			**0.99**				**0.97**			**0.99**		
IPrice		**0.74**			**0.75**				**0.78**			**0.76**		
IDistancetostop	0.48		0.56				0.45				**0.68**			
IStopinfo	**0.78**		0.58				0.51						**0.96**	
IStopsafety	0.59			**0.87**				**0.84**		**0.76**				
ICleanliness	0.48		0.33	0.51				0.52		0.44				
IDriver	0.45		0.40				0.32						0.32	
IOnboardinfo	0.38		0.31				**0.68**				0.42		0.34	
IOnboardsafety	**0.61**			**0.93**				**0.95**		**0.99**				
ITransferwalk	**0.63**		0.64			0.44					0.50			0.30
ITransferwait	**0.77**		**0.74**			**0.71**								0.52
SS loadings	4.92	1.94	3.84	2.23	1.9	2.49	2.44	2.23	1.82	2.09	2.15	1.85	1.74	1.42
Proportion Var	0.33	0.13	0.26	0.15	0.13	0.17	0.16	0.15	0.12	0.14	0.14	0.12	0.12	0.09
Cumulative Var	0.33	0.46	0.26	0.4	0.53	0.17	0.33	0.48	0.6	0.14	0.28	0.41	0.52	0.62

Table 8 Factor Analysis results for different numbers of factors. Indicators of Importance of bus service in Munich

It is recommended to observe the changes when adding or decreasing the number of factors in a data set; changes like which indicators would load on the new factor and how much of variance they are explaining. As a rule of thumb proportion variance less than 10 percent is too small but you may keep a factor with 9.5% of explanation of variance, because the indicators loaded on this factor make a lot of sense.

The loadings below 0.3 have been removed from the results and the ones higher than 0.6 are highlighted to make the decision making process easier. As mentioned before, these thresholds do not follow a specific rule and it is up to the analyzer. One researcher might keep smaller loadings when most of the loadings are in lower range. When an indicator is not loaded in any of the factors (for example if all its loadings are below 0.3) it can be removed from the data and the analysis can be redone.

In our example from three to four factors the information becomes more important but from four to five we only get a new factor with 9% of proportion of variance. I have chosen to have four factors but selecting five factors can be logical as long as you have a good reason to keep the added factor.

4.2 Ordered Logit Model

Ordered logit model in this study has been done in software package R and the command polr() which is short for proportional odds logistic regression. The command can be found in the package MASS.

At first the model has been applied only on satisfaction indicators of bus and metro in Athens and Munich. Later predictors such age , gender, occupation and frequency of use has been added and instead of using 14 or 15 attributes for each model the factor scores have been used. Factor scores has been computed from factor analysis results using the method of sum scores above a cut-off value (which is 0.3).

The technique that has been used for ordered logit model is putting all the predictors at once and removing the variables which are not statistically significant one after next. In all of our models the dependent variable is overall satisfaction of users with public transport service. The p-value used for significance threshold is 0.1.

5 Results and Discussion

5.1 Factor Analysis Results

From importance of service indicators to the level of satisfaction of each indicator, the results of factor analysis between the cities of Munich and Athens are not very different. One finding that should be considered is that people may not pay much attention to some attributes of the service as long as they are in an acceptable quality range but the same indicators might get into attention when the quality of the rest of the service improves.

Tables 9, 10 and 11 illustrate the results of factor analysis with the number of factors to be considered the most optimal for each data set. Below each a table of correlation of factors is shown. This table is useful to check if the rotation method that had been used for analysis was appropriate or not.

In Munich service production and transfer quality, information, quality of service (cleanliness, safety, drivers and etc.) and ticket service are four factors that are important to users of both bus and metro service.

In Athens transfer quality is a separated factor of importance for both bus and metro users comparing to Munich that this dimension of quality was loaded together with service production. Bus users in Athens have not given a big importance to ticket service while it is a factor of importance for metro users.

The findings of Zuo (2017) also shows quality of service and production are two important factors for transit users in Munich. In the three factors she applied for her analysis another factor that got importance was weather which was eliminated from our data set.

Our findings are in line with the results of Efthymiou et al. (2017) in Athens for the importance of bus service indicators. It should be note that in both cities there have been plenty of indicators that was omitted from our data set to keep the opportunity of comparison between cities. Efthymiou et al. (2017) stated that quality of service, service production and transfer quality and also ticket services are the three factors that seem to be important for bus users in Athens.

Efthymiou and Antoniou (2017) applied a factor analysis on survey data of years 2008 and 2013 and their results indicate service production, transfer quality and transfer quality are three factors of importance for bus users; however the order of these factors are different in different years.

In the ternary diagram below three factors of importance has been chosen and the loadings for both cities and both modes can be seen and compared. The dotted red line shows how to read loadings for each data set.

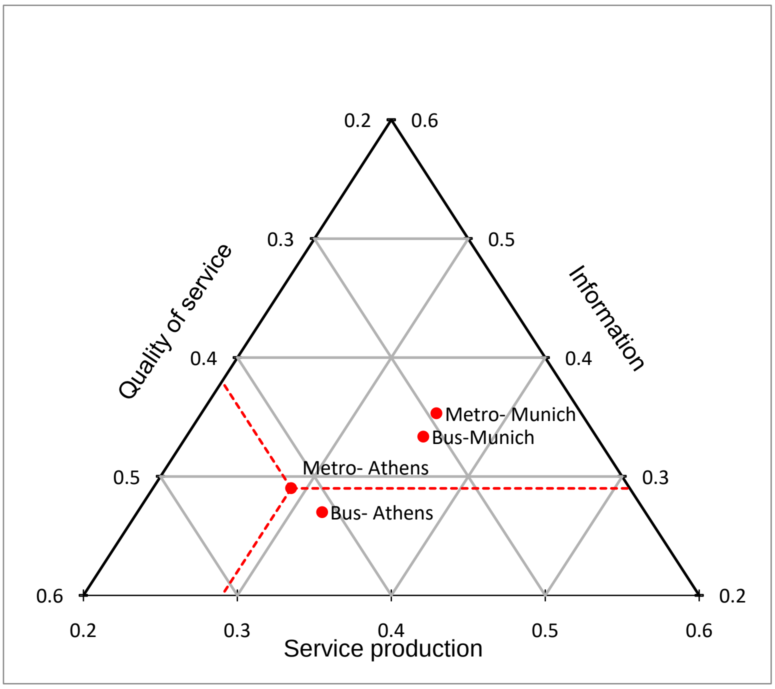


Figure 11 Ternary diagram for three factors of importance of service in Athens and Munich

In this study factor analysis have been done on satisfaction indicators, too. In Munich five and six factors are chosen for final results. For bus service satisfaction the factor of transfer quality has its own separate factor. For satisfaction on metro service accessibility of the stations and network coverage are loaded on a new factor.

Satisfaction with bus driver's behavior in the work of Zuo (2017) is noted as a dimension of quality of service but in our work it is not that significant and it is included in quality of service as a whole.

The results for the satisfaction indicators in Athens can be found in Appendix C. Quality of service, service production, ticket service and information and transfer quality are the four factors of satisfaction of quality indicators in this city.

Service production which has the factor of reliability in itself is important in all of the results in both cities and for both modes of transit. This finding is not new referring to the literature.

	Importance of bus service, Munich				Importance of metro service, Munich			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
IFrequency	**0.77**					**0.88**		
IPunctuality	**0.69**					**0.66**		
ICoverage	0.32	**0.61**			**0.62**	0.31		
IInfo		**0.74**			**0.81**			
ITickettype				**0.97**				**0.98**
IPrice				**0.78**				**0.78**
IDistancetostop		0.45			0.30	0.34		
IStopinfo		0.51			**0.62**			
IStopsafety			**0.84**				**0.81**	
ICleanliness			0.52				0.58	
IDriver		0.32						
IOnboardinfo		**0.68**			**0.71**			
IOnboardsafety			**0.95**				**0.94**	
ITransferwalk	0.44				0.30	0.43		
ITransferwait	**0.71**					**0.66**		
SS loadings	2.49	2.44	2.23	1.82	2.5	2.47	2.16	1.79
Proportion Var	0.17	0.16	0.15	0.12	0.18	0.18	0.15	0.13
Cumulative Var	0.17	0.33	0.48	0.6	0.18	0.36	0.51	0.64
Factor interpretation	*Service production/ transfer quality*	*Information*	*Quality of service*	*Ticket service*	*Information*	*Service production/ transfer quality*	*Quality of service*	*Ticket service*

Factors correlation	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1	0.41	0.41	0.26
Factor 2	0.41	1	0.46	0.37
Factor 3	0.41	0.46	1	0.33
Factor 4	0.26	0.37	0.33	1

Factors correlation	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1	0.46	0.49	0.36
Factor 2	0.46	1	0.41	0.31
Factor 3	0.49	0.41	1	0.39
Factor 4	0.36	0.31	0.39	1

Table 9 Factor analysis results on importance indicators in Munich for bus and metro. Factor loadings and the factor correlation matrices

	Satisfaction about bus service, Munich				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
SFrequency					**0.99**
SPunctuality					0.51
SCoverage	0.47				
SInfo	**0.98**				
STickettype			**0.72**		
SPrice			**0.98**		
SDistancetostop					
SStopinfo	**0.63**				
SStopsafety		**0.70**			
Scleanliness		0.56			
Sdriver	0.31	0.42			
SOnboardinfo	**0.65**				
SOnboardsafety		**0.87**			
STransferwalk				0.56	
STransferwait				**0.97**	
SS loadings	2.72	2.06	1.76	1.61	1.51
Proportion Var	0.18	0.14	0.12	0.11	0.1
Cumulative Var	0.18	0.32	0.44	0.54	0.64
Factor interpretation	*Information*	*Quality of service*	*Ticket service*	*Transfer quality*	*Service production*

	Satisfaction about metro service, Munich					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
SFrequency				**0.97**		
SPunctuality	0.36			0.43		
SCoverage					0.53	
SInfo	**0.74**					
STickettype			**0.77**			
SPrice			**0.95**			
SDistancetostop					**0.82**	
SStopinfo	**0.96**					
SStopsafety		**0.73**				
Scleanliness		0.55				
Sdriver						
SOnboardinfo	**0.69**					
SOnboardsafety		**0.90**				
STransferwalk						0.44
STransferwait						**0.90**
SS loadings	2.49	1.99	1.73	1.51	1.37	1.29
Proportion Var	0.18	0.14	0.12	0.11	0.1	0.09
Cumulative Var	0.18	0.32	0.44	0.55	0.65	0.74
Factor interpretation	*Information*	*Quality of service*	*Ticket service*	*Service production*	*Access*	*Transfer quality*

Factors Cor.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1	0.45	0.34	0.47	0.41
Factor 2	0.45	1	0.27	0.36	0.35
Factor 3	0.34	0.27	1	0.28	0.25
Factor 4	0.47	0.36	0.28	1	0.52
Factor 5	0.41	0.35	0.25	0.52	1

Factor Cor.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Factor 1	1	0.5	0.3	0.5	0.48	0.5
Factor 2	0.5	1	0.29	0.39	0.26	0.32
Factor 3	0.3	0.29	1	0.34	0.24	0.28
Factor 4	0.5	0.39	0.34	1	0.24	0.52
Factor 5	0.48	0.26	0.24	0.24	1	0.34
Factor 6	0.5	0.32	0.28	0.52	0.34	1

Table 10 Factor analysis results on satisfaction indicators in Munich for bus and metro. Factor loadings and the factor correlation matrices

	Importance of bus service, Athens				Importance of metro service, Athens				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
IFrequency		**0.93**						**0.65**	
IPunctuality		**0.61**						**0.64**	
ICoverage		0.47		0.30					
IInfo			**0.68**				0.35		
ITickettype			0.54				**0.98**		
Iprice							0.4	0.5	
IDistancetostop				0.51		0.32			
IStopinfo	0.36								**0.98**
IStopsafety	**0.73**				**0.91**				
ICleanliness	0.55				**0.64**				
IDriver	0.57		0.34						
IOnboardinfo			0.44		0.4				0.3
IOnboardsafety	**0.83**				0.49				
ITransferwalk				**0.64**		**0.8**			
ITransferwait				0.49		**0.64**			
SS loadings	2.46	1.97	1.63	1.37	2.1	1.68	1.44	1.44	1.42
Proportion Var	0.18	0.14	0.12	0.1	0.16	0.13	0.11	0.11	0.11
Cumulative Var	0.18	0.32	0.43	0.53	0.16	0.29	0.4	0.51	0.62
Factor interpretation	*Service quality*	*Service production*	*Information*	*Transfer quality*	*Service quality*	*Transfer quality*	*Ticket service*	*Service production*	*Information*

Factors correlation	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1	0.36	0.5	0.44
Factor 2	0.36	1	0.41	0.3
Factor 3	0.5	0.41	1	0.28
Factor 4	0.44	0.3	0.28	1

Factors Cor.	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Factor 1	1	0.61	0.34	0.56	0.46
Factor 2	0.61	1	0.36	0.41	0.45
Factor 3	0.34	0.36	1	0.17	0.42
Factor 4	0.56	0.41	0.17	1	0.23
Factor 5	0.46	0.45	0.42	0.23	1

Table 11 Factor analysis results on importance indicators in Athens for bus and metro. Factor loadings and the factor correlation matrices

5.2 Ordered Logit Model Results

In tables 12 to 15 the results of applying the model on indicators of satisfaction for different modes in both cities is shown. It has been expected that the values be positive which means improvements in the satisfaction about an attribute of a service while other variables are fixed results into a higher satisfaction about the dependent variable which is the overall satisfaction.

Athens Bus	Value	t value	p value
SPrice	0.29	2.47	1.3E-02
SFrequency	0.56	3.71	2.1E-04
STransferwait	0.38	2.45	1.4E-02
Intercepts:			
1|2	1.36	2.70	6.9E-03
2|3	3.90	7.00	2.5E-12
3|4	7.66	9.91	3.9E-23
Residual deviance	494.48		
AIC	506.48		

Table 12 Ordered logit estimation for satisfaction variables of bus service in Athens

Athens Metro	Value	t value	p value
SFrequency	0.70	5.26	1.4E-07
SCoverage	0.20	2.25	2.4E-02
STickettype	0.29	2.42	1.6E-02
SPrice	0.44	4.22	2.4E-05
STransferwait	0.41	3.05	2.3E-03
Intercepts:			
1|2	1.56	1.97	4.9E-02
2|3	5.32	7.66	1.9E-14
3|4	8.44	10.43	1.8E-25
Residual deviance	557.08		
AIC	573.08		

Table 13 Ordered logit estimation for satisfaction variables of metro service in Athens

Munich Bus	Value	t value	p value
SFrequency	0.52	2.94	3.2E-03
SPrice	0.22	1.64	1.0E-01
STransferwalk	0.33	1.69	9.0E-02
SCoverage	0.29	1.63	1.0E-01
SDriver	0.52	3.18	1.5E-03
Intercepts:			
1|2	3.57	3.89	1.0E-04
2|3	5.66	5.83	5.6E-09
3|4	9.06	7.89	3.0E-15
Residual deviance	331.04		
AIC	349.04		

Table 14 Ordered logit estimation for satisfaction variables of bus service in Munich

Munich Metro	Value	t value	p value
SFrequency	0.50	2.59	9.6E-03
SPunctuality	0.31	1.92	5.5E-02
SDistancetostop	0.30	2.07	3.9E-02
Intercepts:			
1|2	1.15	1.64	1.0E-01
2|3	3.16	4.40	1.1E-05
3|4	6.18	7.34	2.1E-13
Residual deviance	374.57		
AIC	386.57		

Table 15 Ordered logit estimation for satisfaction variables of metro service in Munich

Frequency of service as a part of reliability of a service is statistically significant in all the model results above and the values related to this attribute is the biggest. Price and waiting time while transferring are seen in both modes of transit in Athens.

In Munich satisfaction with driving has a value of 0.6 and shows its influence on overall satisfaction. Transfer-walk, network coverage and distance to stop all have something in common which is accessibility and is resulted in both modes in Munich. It can be concluded that people normally will to walk less or spend less time off-board.

The analysis on indicators of importance is added to the report in Appendix D.

In second part of ordered logit analysis factor scores calculated from the factor analysis results have been used. At the very first attempt all possible interactions of these factors with mode, gender, age groups, occupation status and frequency of use has been applied as predictors of the model but as not all of them have been statistically significant different interactions have been removed, added again or only independent variables without being multiplied with other variables have been used as the input of the models.

For example one finding which is shown in table 16 is how age and retired respondents would illustrate sort of relationship about overall satisfaction toward the transit service. When a person is retired/ unemployed with one unit increase of age the overall satisfaction log odds changes negatively (0.02 - 0.03= - 0.01) but very slightly. When a person is not retired (Retired/ Unemployed = 0) the log odds changes is positive; however, very small (0.02).

	Value	t value	p value
Age	0.02	2.64	8.4E-03
Retired/ Unemployed	0.71	1.23	2.2E-01
Retired/ Unemployed: Age	-0.03	-2.24	2.5E-02
Intercepts:			
1|2	-2.00	-6.97	3.2E-12
2|3	0.25	0.96	3.4E-01
3|4	2.56	9.01	2.0E-19
Residual Deviance	1312.31		
AIC	1324.31		

Table16 Ordered logit estimation for Age and Retired predictors of overall satisfaction in Athens

Another model is set using only the factors extracted from factor analysis. This model has been done for both cities and results are shown in Tables 17 and 18. In Athens all the factors are statistically significant but in Munich factor of information has not been significant. Service production in both cities has a very positive relationship with satisfaction. The more people are satisfied with the frequency, punctuality and other aspect of reliability of the service the more they are satisfied about the whole transit service.

Transfer quality, ticket system and quality of the service in Athens are also positively related to overall satisfaction. In Munich the same factors with lower values are important in overall satisfaction as well.

Variable	Value	t value	p value
Service Quality	0.41	2.16	3.1E-02
Service production	0.99	5.74	9.3E-09
Ticket system	0.97	3.95	8.0E-05
Transfer quality	1.04	3.06	2.2E-03
Intercepts:			
1|2	2.83	6.28	3.4E-10
2|3	5.67	11.64	2.7E-31
3|4	8.79	15.26	1.3E-52
Residual Deviance	1076.99		
AIC	1090.99		

Table 17 Ordered logit model estimation using factors as predictors for overall satisfaction in Athens

Variable	Value	t value	p value
Service quality	0.35	1.78	7.4E-02
Service production	0.90	4.57	4.8E-06
Ticket system	0.19	1.53	1.3E-01
Transfer quality	0.36	1.98	4.8E-02
Intercepts:			
1|2	1.78	3.17	1.5E-03
2|3	3.76	6.44	1.2E-10
3|4	6.85	10.20	2.0E-24
Residual Deviance	723.50		
AIC	737.50		

Table 18 Ordered logit model estimation using factors as predictors for overall satisfaction in Munich

Adding demographic variables or frequency of transit to the factors of service quality and the interaction with mode of public transport has been a bit challenging. The results in Athens is in shown in Table 19.

In Munich despite age groups there has been no statistically significant output from the regression analysis. I questioned the repetition of demographic variables for bus and metro and tried to do the logit model separately for each mode; however, the results did not show any major changes.

For the results in Athens the base for occupation status is “Employed” and except being a student the other job categories have a negative coefficient. Although for the not significant p-values the discussion about these values is not useful. Other jobs may have a small enough p-value but the fact is that only a few observations are in this category and the sample seems to be too small for further interpretations.

The base for age groups is the youngest group (<=18 years old). The changes from on age group to the next one does not really show a constant increase or decrease and taking a look at the estimation where only Age and Retirement have been used the value is very close to zero.

Most of the estimation of interaction of a binary variable which is metro when Metro=1 and bus when Metro=0 and the service quality factors which are continues variables have small enough p-values. Improvements in satisfaction of service quality, production of service and ticket system for bus service would result in higher satisfaction in total but transfer quality has a negative value which is strange but as it is not statistically significant we would not interpret it. Metro service, on the other hand, has all positive values for the factors of quality of service. The values can be computed with adding the value of “Factor” variable with the value of “Factor: Metro” variable.

Variable	Value	t value	p value
Service quality	0.57	1.66	9.7E-02
Mode: Metro	-0.96	-0.98	3.3E-01
Service production	1.92	4.76	2.0E-06
Ticket system	0.85	2.07	3.9E-02
Transfer quality	-0.84	-1.25	2.1E-01
Occupation:			
Other jobs	-1.04	-1.91	5.7E-02
Retired/ Unemployed	-0.28	-0.93	3.5E-01
Self-employed	-0.34	-1.17	2.4E-01
Student	0.01	0.02	9.9E-01
Age groups			
>=60	-0.75	-1.12	2.6E-01
19-23	-0.74	-1.61	1.1E-01
24-29	-1.00	-2.02	4.3E-02
30-39	-0.62	-1.11	2.7E-01
40-49	-0.85	-1.51	1.3E-01
50-59	-0.56	-0.95	3.4E-01
Metro: Service Quality	-0.23	-0.55	5.8E-01
Metro: Service production	-1.53	-3.30	9.7E-04
Metro: Ticket system	0.97	1.75	8.0E-02
Metro: Transfer quality	2.46	3.02	2.6E-03
Intercepts:			
1|2	1.36	1.72	8.6E-02
2|3	4.28	5.18	2.2E-07
3|4	7.58	8.83	1.0E-18
Residual Deviance	1042.26		
AIC	1088.25		

Table 19 Ordered logit model estimation for overall satisfaction in Athens. Highlighted variables are statistically significant

6 Conclusions and Recommendations

Summary

This thesis attempted to assess the most influential aspects of transit service quality from public transport users' point of view in two European cities, Athens and Munich. For this, a factor extraction method and a choice model was built on observations that were based on individuals' respondents on different attributes of a transit service. The models were based on the data available from conduction of surveys in both cities in 2017. There were some differences in questionnaire designs that needed to be generalized so that the comparison of results between Athens and Munich could be possible.

After excluding or approximating some variables from data sets four demographic and travel behavior related variables and fifteen service quality indicators were left for further analysis.

Exploratory factor analysis as the chosen extraction method was used on service quality attributes using software package of R and the package psych(). The analysis was done separately for each city and each mode as well as for indicators of satisfaction separated from importance. Service quality, production of the service, information and transfer quality are the four factors that have been seen in all data sets of indicators of importance. Ticketing system have been observed as important in three data sets but not in Athens metro. Accessibility to the stops has been a factor that got importance for metro service in Munich.

The factors that have been estimated from factor analysis are also seen in literature. Factors addressing environmental concerns are getting more important but the data used in this study lack such information. Reliability of the service, quality of the service it provides such as cleanliness and safety, walking and access time/ distance, information provision and ticketing system and price are all factors that seem to have impact on satisfaction of transit users.

Could we generalize the findings to other cities? Yes and no. The factors are pretty much the same; for example, almost everyone cares for reliability of a service or does not want to waste time waiting or walking long distances to get to a service but the level of satisfaction differs from city to city. Even in our findings the loading on each factor and the order of their extraction is different. How well a transit is serving at the time has a big influence on users' perception. In Athens people seem to be very satisfied with metro service and have a very positive opinion about it which builds royalty. Bus service, on the other hand, needs improvements of its image. Another example can be the information factor in Munich where it is more important for metro than for bus. The bus users are normally locals and use the service in shorter distances between one or two neighborhoods but at metro more random users can be seen and they travel longer distances to more unfamiliar areas which makes information provision more important. On the other hand, the penalty of missing a stop on the bus is much less than the metro.

The choice model used in this study was ordered logit model, also known as proportional odds logistic regression and was built in software R using the package MASS(). The overall satisfaction was set as dependent variable and satisfaction attributes with demographic variables as predictors of the model.

Model estimation on service attributes showed roughly the same factors that was seen from factor analysis. Information; however, was not statistically significant. Satisfaction values got positive values while importance intercepts which can be seen in Appendix D were more negative and this was expected.

Model estimation in Athens using factors from factor analysis and other individual variables such as age, gender and occupation for a combination of modes revealed some results that showed how metro and bus service were slightly different regarding overall satisfaction. Age groups also showed a very small relationship. Unfortunately the model estimation with the same approach for Munich was not reliable enough.

Limitations and recommendations

Some of the limitations of the work carried out in this thesis and some recommendations concluded from these limitations are listed below. In future research, an attempt should be made to address these limitations in order to improve the reliability and validity of the results predicted.

- Many attributes of service quality and individual demographic or travel behavior characteristics could not be included in the data sets. The major reason was that the attempt to have similar attributes and predictors in both cities of Athens and Munich led to elimination of a lot of data. A few observations in Munich due to lack of some information had to be omitted.

- The sample size in Munich has been quite smaller than of Athens and another difference in data sets of these two cities has been that in Athens observation related to bus service had been taken from different groups of respondents while in Munich one group of transit riders had been surveyed for both modes of public transport.
- The model only involves subjective variables of service quality and due to biases in different stages of the study from designing the questionnaires to collecting data and model assumptions it is more recommended to have objective measures and attributes as well. It is to be noted that the nature of individual's judgment is heterogeneous itself.
- People who use a service and people who do not use it have different opinions about the same service (Laura Eboli and Mazzulla, 2011); therefore, it is more realistic to involve both transit riders and non-riders in data collection phase.
- In the questionnaire of Athens there have been four alternatives for the question of overall satisfaction which is not recommended. Having an odd number of alternatives for Likert type answers are better. The difference regarding this question between two cities led to some changes in observations in Munich which increases the chance of biases.
- Waiting time could not be included in data set of Munich due to the goal of having similar attributes in both cities. This indicator has been important in the results of previous study in Munich.
- The relationship between people's opinions on the importance of an indicator and their level of satisfaction about it could be studied further. The effort was made to do a model using the product of these two groups of observations but due to time limitations and difficulty in finding literature it could not be finished.
- The validity of the predictors could be explored more carefully using more complex statistical methods. A more precise calibration or validation of variables prior to modelling would be done when a bigger sample is available.
- To model each city separately and to compare the results is the approach of this thesis; however, a more sophisticated method is to do simultaneous analysis. An integrated choice and latent variable model that is done simultaneously is more efficient than doing it sequentially. These are types of structural equation models. A Hybrid model can also be used for further analysis. The results of factor analysis can give a proper idea for choosing factors for a confirmatory factor analysis which is a part of a hybrid model or a structural equation model.

Bibliography

Baglin, J. (2014). Improving Your Exploratory Factor Analysis for Ordinal Data : A Demonstration Using FACTOR. *Practical Assessment, Research & Evaluation*, *19*(5).

Beirao, G., & Sarsfield Cabral, J. . (2007). Understanding attitudes towards public transport and private car: a qualitative study. *Transport Policy*, *14*(6), 478–489.

Ben-Akiva, M. E., & Lerman, S. R. (1985). *Discrete choice analysis: theory and application to travel demand. The MIT Press*. Cambridge, Massachusetts.

Berry, L. ., & Parasuraman, A. (1997). Listening to the customer- the concept of a service quality information system. *MIT Sloan Management Review*, *38 (3)*, 65.

Bickel, P., Friedrich, R., Link, H., Stewart, L. and Nash, C. (2006). Introducing environmental externalities into transport pricing: Measurement and implications. *Transport Reviews*, *26 (4)*, 389–415.

Biesok, G., & Wyród-wróbel, J. (2012). Indicative methods of customer satisfaction surveys in public transport services. In *15th QMOD Conference* (pp. 194–209).

Buehler, R., & Pucher, J. (2012). Demand for Public Transport in Germany and the USA: An Analysis of Rider Characteristics. *Transport Reviews*, *32*(5), 541–567. https://doi.org/10.1080/01441647.2012.707695

Chou, J., & Kim, C. (2009). A structural equation analysis of the QSL relationship with passenger riding experience on high speed rail : An empirical study of Taiwan and Korea. *Expert Systems With Applications*, *36*(3), 6945–6955. https://doi.org/10.1016/j.eswa.2008.08.056

Chowdhury, S., Hadas, Y., Gonzalez, V. A., & Schot, B. (2018). Public transport users' and policy makers' perceptions of integrated public transport systems. *Transport Policy*, *61*(October 2017), 75–83. https://doi.org/10.1016/j.tranpol.2017.10.001

Cirillo, C., Eboli, L., & Mazzulla, G. (2011). On the asymmetric user perception of transit service quality. *International Journal of Sustainable Transportation*, *5*(4), 216–232. https://doi.org/10.1080/15568318.2010.494231

Costello, A. B., & Osborne, J. W. (2005). Best Practices in Exploratory Factor Analysis : Four Recommendations for Getting the Most From Your Analysis. *Practical Assessment, Research & Evaluation*, *10*(7).

de Oña, J., & de Oña, R. (2014). Quality of Service in Public Transport Based on Customer Satisfaction Surveys: A review and assessment of methodological approaches. *Transportation Science*, *49(3)*, 605–622. https://doi.org/10.1287/trsc.2014.0544

de Oña, J., de Oña, R., & Calvo, F. (2012). A classification tree approach to identify key factors of transit service quality. *Expert Systems With Applications*, *39*(2011), 11164–11171.

de Oña, R., Machado, J. L., & de Oña, J. (2015). Perceived Service Quality, Customer Satisfaction, and Behavioral Intentions: A Structural Equation Model for the Metro of

Seville, Spain. *Transportation Research Record: Journal of the Transportation Research Board*, *2538*, 76–85. https://doi.org/10.3141/2538-09

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., de Oña, J., & de Oña., R. (2017). *Public Transportation Quality of Service: Factors, Models, and Applications*. Elsevier.

Distefano, C., Zhu, M., & Mîndrilă, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research & Evaluation*, *14*(20), 1–11. https://doi.org/10.1.1.460.8553

Downs, A. (2005). *Still Stuck in Traffic*. Brookings Institution Press.

Eboli, L., & Mazzulla, G. (2008). A Stated Preference Experiment for Measuring Service Quality in Public Transport A Stated Preference Experiment for Measuring Service Quality in Public Transport. *Transportation Planning and Technology*, *31*(5), 509–523. https://doi.org/10.1080/03081060802364471

Eboli, L., & Mazzulla, G. (2010). How to capture the passengers' point of view on a transit service through rating and choice options. *Trasnport 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. https://doi.org/10.1016/j.tranpol.2010.07.007

Eboli, L., & Mazzulla, G. (2012). Structural equation modelling for analyzing passengers' perceptions about railway services. *Procedia- Social and Behavioural Science*, *54*, 96–106.

Eboli, L., & Mazzulla, G. (2013). Performance indicators for an objective measure of public transport service quality. *European Transport*, (51), 1–21.

Efthymiou, D., & Antoniou, C. (2017). Understanding the effects of economic crisis on public transport users' satisfaction and demand. *Transport Policy*, *53*(August 2016), 89–97. https://doi.org/10.1016/j.tranpol.2016.09.007

Efthymiou, D., Antoniou, C., & Tyrinopoulos, Y. (2017). Factors Affecting Bus Users ' Satisfaction in Times of Economic Crisis. *Transportation Research Part A: Policy and Practice*.

Efthymiou, D., Antoniou, C., & Waddell, P. (2013). Factors affecting the adoption of vehicle sharing systems by young drivers. *Transport Policy*, *29*, 64–73.

Efthymiou, D., Kaziales, M., & Antoniou, C. (2014). Measuring the Effects of Economic Crisis on Users ' Perceptions of Public Transport Quality. *Transportation Research Record, Journal of the Transportation Research Board, Washington D.C.*, *2415*, 1–12. https://doi.org/10.3141/2415-01

Ekstrom, J. (2011). On the relation between the polychoric correlation coefficient and Spearman's rank correlation coefficient. Retrieved from https://escholarship.org/uc/item/7j01t5sf

Eurostat. (n.d.). Statistical office of the European Union. Retrieved from http://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=urb_lpop1&lang=en

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, *4*(3), 272–299.

Fellesson, M., & Friman, M. (2008). Perceived satisfaction with public transport service in Nine European cities. *The Transportation Research Forum*, *47*(3), 93–103. https://doi.org/10.5399/osu/jtrf.47.3.2126

Fluegel, S., Caspersen, E., Angell, T., Fearnley, N., & Kwong, C. K. (2015). A nested logit model for instantaneous travel mode choice analyses in the Oslo area. *Manuscript. Institute of Transport Economics, Oslo*.

Friman, M., & Fellesson, M. (2009). Service Supply and Customer Satisfaction in Public Transportation : The Quality Paradox. *Journal of Public Transportation*, *12*(4), 57–69.

Garrido, C., Oña, R. De, & Oña, J. De. (2014). Neural networks for analyzing service quality in public transportation. *Expert Systems With Applications*, *41*, 6830–6838. https://doi.org/10.1016/j.eswa.2014.04.045

Golob, T. F. (2003). Structural equation modeling for travel behavior research. *Transportation Research Part B*, *37*, 1–25.

Hall, C. M., & Gerike, R. (2014). Analysis of Visitor Satisfaction with Public Transport in Munich. *Journal of Public Transportation*, *17*(3), 68–85.

Harrell, F. E. (2015). *Regression Modeling Strategies*. Springer. Retrieved from https://books.google.de

Hensher, D. A. (2014). The relationship between bus contract costs, user perceived service quality and performance assessment. *International Journal of Sustainable Transportation*, *8*(1), 5–27.

Hensher, D. A., Stopher, P., & Bullock, P. (2003). Service quality - developing a service quality index in the provision of commercial bus contracts. *TRANSPORTATION RESEARCH PART A*, *37*(2003), 499–517. https://doi.org/10.1016/S0965-8564(02)00075-7

Hensher, D. A., & Ton, T. T. (2006). A comparison of the predictive potential of arti ® cial neural networks and nested logit models for commuter mode choice. *Transportation Research Part E*, *36*(2000), 155–172.

Holgado–Tello, F. ., Chacón–Moscoso, S., Barbero–García, I., & Vila–Abad, E. (2010). Polychoric versus Pearson correlations in exploratory and confirmatory factor analysis of ordinal variables. *Quality & Quantity*, *44*(1), 153.

Hosmer Jr, D. ., Lemeshow, S., & Sturdivant, R. . (2013). *Applied logistic regression, Third Edition*. John Wiley & Sons. Retrieved from https://books.google.de

Iseki, H., & Taylor, B. D. (2010). Style versus Service? An Analysis of User Perceptions of Transit Stops and Stations. *Journal of Public Transportation*, *13*(3), 39–63. https://doi.org/http://dx.doi.org/10.5038/2375-0901.13.3.2

Jakobsson, C., Gamble, A., Hagman, O., Polk, M., Gärling, T., & Olsson, L. E. (2011). Affective – symbolic and instrumental – independence psychological motives mediating effects of socio-demographic variables on daily car use. *Journal of Transport Geography*, *19*(1), 33–38. https://doi.org/10.1016/j.jtrangeo.2009.11.006

Kim, H. (2013). Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics*, *38*(1), 52–54. Retrieved from http://dx.doi.org/10.5395/rde.2013.38.1.52

Koppelman, F., Cities-duluth, T., States, U., Conference, I., & Modeling, B. T. (1993). Application and Interpretation of Nested Logit Models of Intercity Mode Choice. *Transportation Research Record 1413*, 98–106.

McCullagh, P. (1980). Regression Models for Ordinal Data. *Journal of the Royal Statisticl Society. Series B (Methodological)*, *42*(2), 109–142.

Munich, C. of. (2013). *Munich: Future Perspective*. Retrieved from file:///C:/Users/Candell/Downloads/PM_Magazin_en_web.pdf

Munizaga, M. A., & Ortúzar, J. d. D. (1999). Nested Logit Modelling: Some Hard Facts. *27th European Transport Conference*, *P324*, 25–36.

Ngoc, A. M., Hung, K. V., & Tuan, V. A. (2017). Towards the Development of Quality Standards for Public Transport Service in Developing Countries: Analysis of Public Transport Users' Behavior. *Transportation Research Procedia*, *25*, 4564–4583. https://doi.org/10.1016/j.trpro.2017.05.354

Ortúzar, J. d. D., & Willumsen, L. G. (2011). *Modelling Transport (Fourth Edition)*. Chichester, West Sussex, United Kingdom: John Wiley & Sons.

Redman, L., Friman, M., Gärling, T., Hartig, T. (2013). Quality attributes of public transport that attract car users: a research review. *Transport Policy*, *25*, 119–127.

Roman, C., Martin, J. C., & Espino, R. (2014). Using stated preferences (SP) to analyze the service quality of public transport. *International Journal of Sustainable Transportation*, *8 (1)*, 28–46.

Seltman, H. J. (2015). Exploratory data analysis (Chapter 4). In *Experimental Design and Analysis* (pp. 61–100). https://doi.org/0201076160

Shen, J., & Li, W. (2014). Discrete Hopfield Neural Networks for Evaluating Service Quality of Public Transit Jinxing. *International Journal of Multimedia and Ubiquitous Engineering*, *9*(2), 331–340.

Stuart, K. R., Mednick, M., & Bockman, J. (2000). Structural Equation Model of Customer Satisfaction for the New York City Subway System. *Transportation Research Record 1735*, 133–137.

Tang, L., & Thakuriah, P. . (2012). Ridership effects of real-time bus information system: A case study in the City of Chicago. *Transportation Research Part C: Emerging Technologies*, *22*, 146–161.

Taylor, B. D., Miller, D., Iseki, H., & Fink, C. (2009). Nature and/ or nurture? Analyzing the determinants of transit ridership across US urbanized areas. *Transportation Research Part A Policy and Practice*, *43*(1), 60–77. https://doi.org/10.1016/j.tra.2008.06.007

The National Archives. (2011). Commission for Integrated Transport. Retrieved from http://webarchive.nationalarchives.gov.uk/20110310110545/http://cfit.independent.gov.uk/pubs/2001/ebp/ebp/stage1/07.htm

Train, K. . (2009). *Discrete choice methods with simulation*. Cambridge university press. Retrieved from https://books.google.de

TRB. (1999). *A Handbook for Measuring Customer Satisfaction and Service Quality. TRCP Report 47*. Washington, D.C.

TRB. (2003). *A Guidebook for Developing a Transit Performance-Measurement System. TCRB Report 88*. Washington, D.C.

TRB. (2004). *Transit Capacity and Quality of Service Manual (2nd Edition). TCRB Report 100*. Washington, D.C.

TRL Limited. (2004). *The demand for public transport: a practical guide. TRL Report TRL593*. *TRL Limited*. Crowthorne, UK. Retrieved from http://westminsterresearch.wmin.ac.uk/479/

Tyrinopoulos, Y., & Aifadopoulou, G. (2008). A complete methodology for the quality control of passenger services in the public transport business. *European Transport*, *38*, 1–16.

Tyrinopoulos, Y., & Antoniou, C. (2008). Public transit user satisfaction: Variability and policy implications. *Transport Policy*, *15*, 260–272. https://doi.org/10.1016/j.tranpol.2008.06.002

van Lierop, D., & El-geneidy, A. (2017). PUBLIC IMAGE : Determining the effects of users ’ image of public transit on loyalty. *96th Transportation Research Board Annual Meeting*.

Velicer, W. F., & Jackson, D. N. (1990). Component analysis versus common factor analysis: some issues in selecting appropriate procedure. *Multivariate Behavioral Research*, *25*(1), 1–28.

Venables, W. N., & Ripley, B. D. (2002). *Modern Applied Statistics with S. Fourth Edition*. Springer. Retrieved from www.springer.com

Washington, S. P., Karlaftis, M. G., & Mannering, F. (2010). *Statistical and econometric methods for transportation data analysis*. Chapman and Hall/CRC.

Widaman, K. F. (1993). Common Factor-Analysis Versus Principal Component Analysis - Differential Bias in Representing Model Parameters. *Multivariate Behavioral Research*, *28*(3), 263–311.

Ye, R., & Titheridge, H. (2017). Satisfaction with the commute: The role of travel mode choice, built environment and attitudes. *Transportation Research Part D Transport and Environment*, *52*, 535–547. https://doi.org/10.1016/j.trd.2016.06.011

Zeithaml, V. A., Parasuraman, A., & Berry, L. L. (1990). *Delivering quality service: Balancing customer perceptions and expectations*. Simon and Schuster.

Zuo, Y. (2017). *Impact of Weather on Public Transport Users’ Satisfaction*. Technical University of Munich. Retrieved from https://mediatum.ub.tum.de/doc/1432707/843181382906.pdf

List of Abbreviations

EFA	Exploratory Factor Analysis
CFA	Confirmatory Factor Analysis
PCA	Principle Component Analysis
sd	Standard Deviation
Cor.	Correlation
ML	Maximum Likelihood
AIC	Akaike Information Citerion
PT	Public Transportation
Comp.	Component

.

.

List of Tables

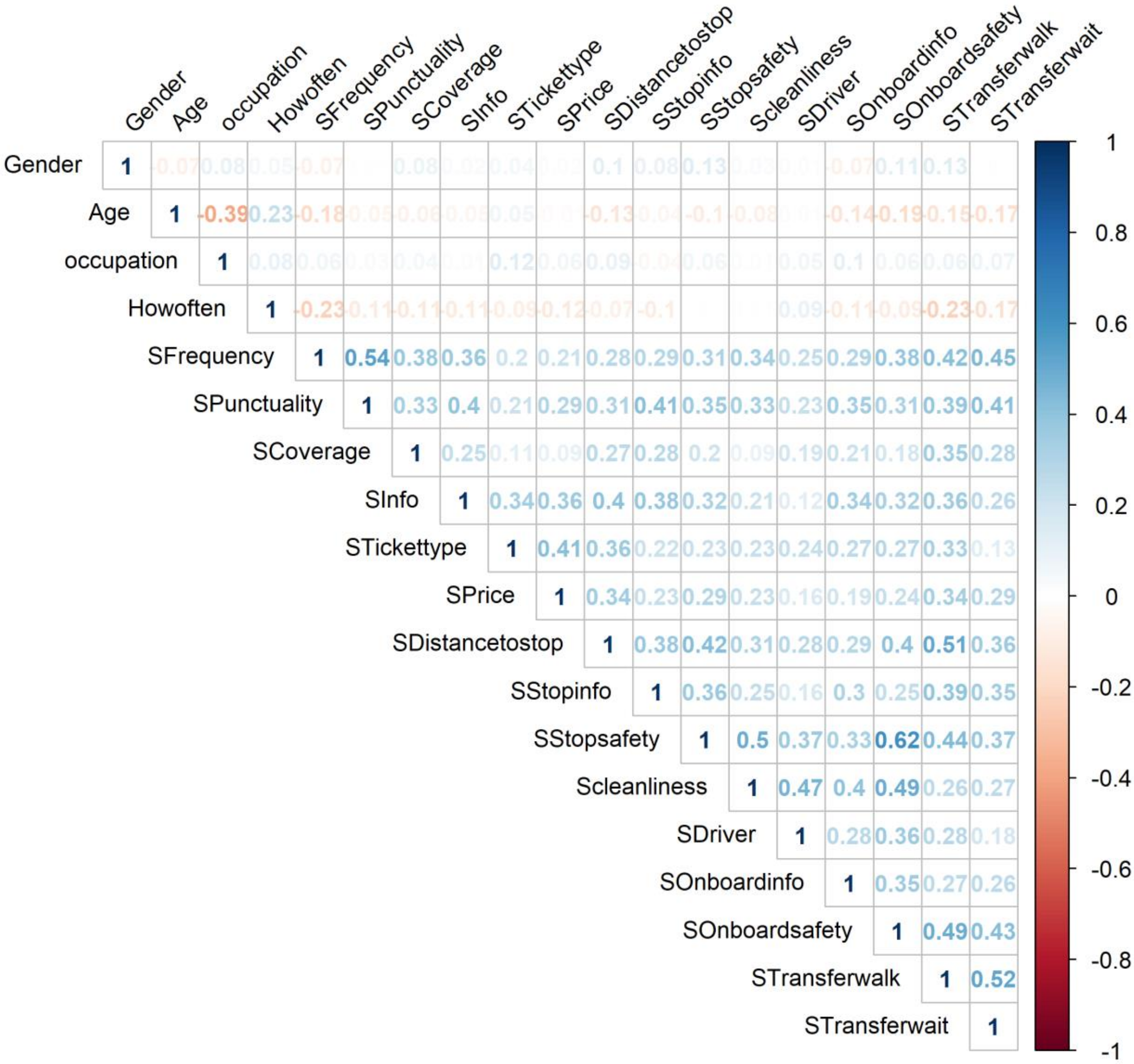
List of Figures

Appendix A: Statistics summary of quality indicators in Munich for Bus and Metro services

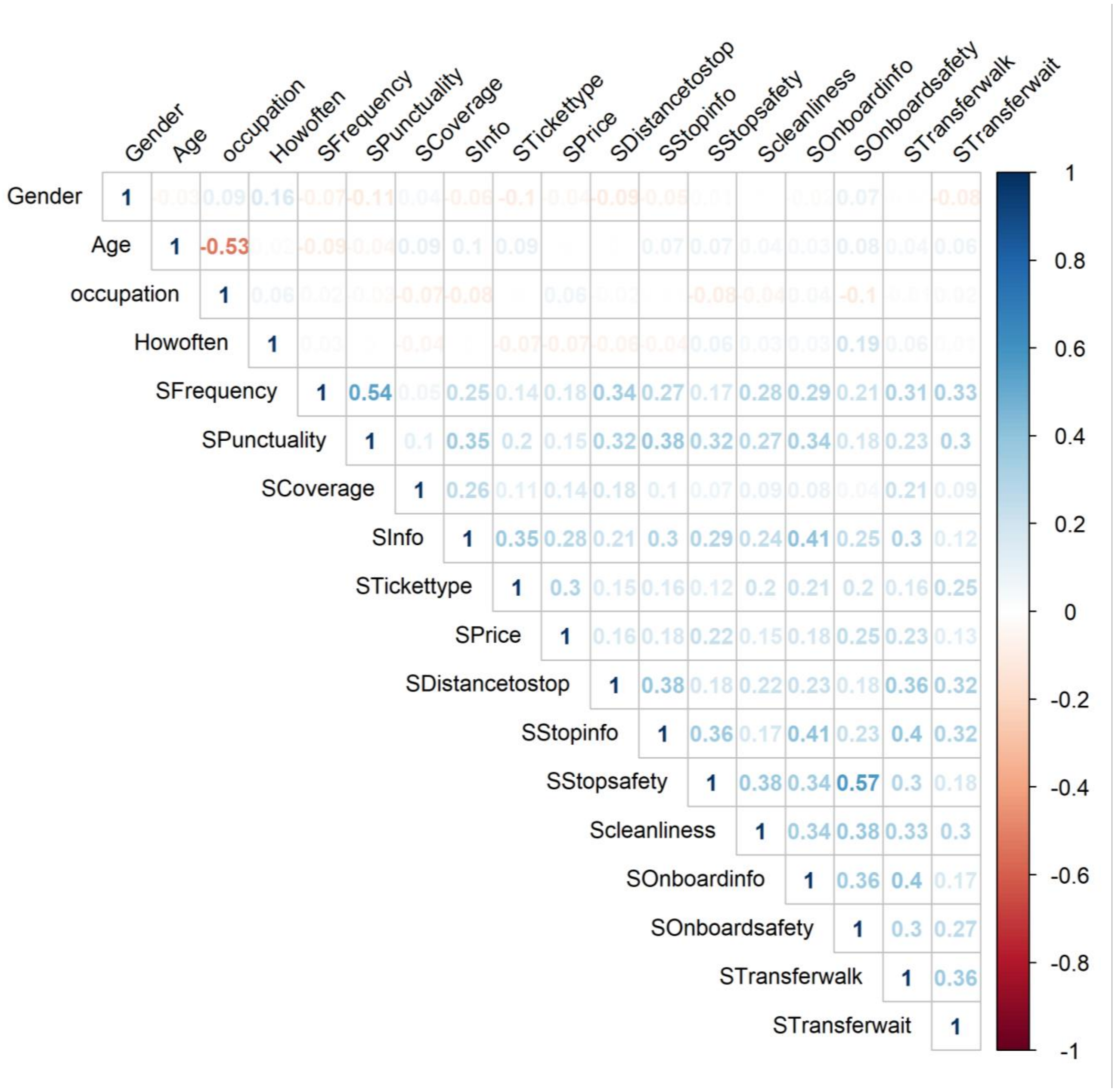
	Munich Bus (169 observations)					**Munich Metro (177 observations)**				
Importance	mean	sd	median	skew	kurtosis	mean	sd	median	skew	kurtosis
IFrequency	4.04	0.88	4	-0.70	-0.01	3.99	0.96	4	-0.76	-0.05
IPunctuality	4.43	0.92	5	-1.79	2.84	4.40	1.02	5	-1.82	2.65
ICoverage	3.94	1.09	4	-0.88	0.14	3.99	1.14	4	-1.03	0.32
IInfo	4.08	1.19	5	-1.09	0.13	4.06	1.19	5	-1.09	0.17
ITickettype	3.91	1.19	4	-0.78	-0.40	3.87	1.22	4	-0.78	-0.45
IPrice	4.04	1.14	4	-0.92	-0.14	4.01	1.14	4	-0.89	-0.19
IDistancetostop	3.90	1.04	4	-0.69	-0.08	3.84	1.10	4	-0.70	-0.20
IStopinfo	4.34	1.09	5	-1.69	1.94	4.41	1.03	5	-1.88	2.82
IStopsafety	3.93	1.18	4	-0.78	-0.43	3.89	1.17	4	-0.72	-0.49
ICleanliness	4.04	1.03	4	-1.06	0.74	4.05	1.02	4	-0.94	0.28
IDriver	3.82	1.05	4	-0.72	-0.02	-	-	-	-	-
IOnboardinfo	3.88	1.23	4	-0.82	-0.30	3.86	1.23	4	-0.81	-0.30
IOnboardsafety	3.96	1.16	4	-0.85	-0.25	3.98	1.17	4	-0.91	-0.16
ITransferwalk	3.92	0.97	4	-0.57	-0.35	3.92	0.98	4	-0.52	-0.61
ITransferwait	4.01	0.96	4	-0.62	-0.50	3.97	1.00	4	-0.77	0.11
Satisfaction										
SFrequency	3.20	1.00	3	-0.22	-0.25	3.33	1.02	3	-0.17	-0.36
SPunctuality	3.22	1.14	3	-0.04	-0.87	3.34	1.16	3	-0.27	-0.80
SCoverage	3.57	1.00	4	-0.53	0.20	3.34	1.07	3	-0.35	-0.27
SInfo	3.83	1.08	4	-0.62	-0.29	3.88	1.11	4	-0.76	-0.16
STickettype	3.21	1.16	3	-0.14	-0.66	3.23	1.18	3	-0.18	-0.71
SPrice	2.85	1.28	3	0.11	-1.01	2.84	1.23	3	0.19	-0.85
SDistancetostop	3.69	0.97	4	-0.37	-0.19	3.59	1.08	4	-0.37	-0.48
SStopinfo	3.56	1.25	4	-0.37	-1.01	3.76	1.27	4	-0.74	-0.57
SStopsafety	3.63	1.02	4	-0.24	-0.52	3.67	1.00	4	-0.20	-0.62
Scleanliness	3.67	1.08	4	-0.38	-0.74	3.55	1.09	4	-0.29	-0.69
SDriver	3.63	0.97	3.5	-0.55	0.43	-	-	-	-	-
SOnboardinfo	3.78	1.13	4	-0.65	-0.32	3.69	1.17	4	-0.55	-0.57
SOnboardsafety	3.75	1.08	4	-0.44	-0.45	3.71	1.09	4	-0.43	-0.46
STransferwalk	3.57	0.92	4	-0.33	-0.19	3.59	1.01	4	-0.33	-0.50
STransferwait	3.30	1.06	3	-0.17	-0.57	3.42	1.10	4	-0.37	-0.50

Appendix A.1 Data statistics for quality indicators in Munich

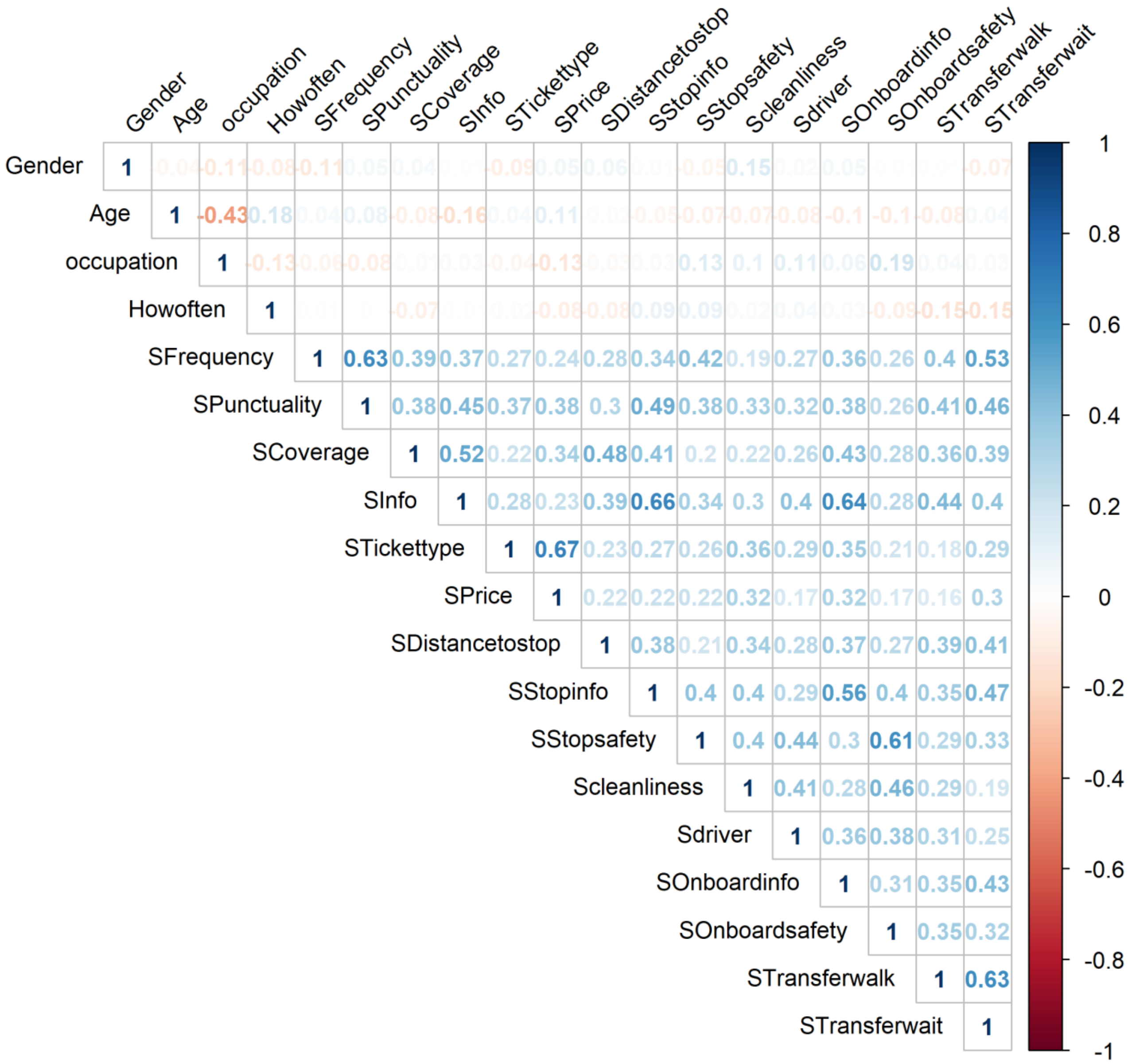
Appendix B: Correlation matrices for satisfaction indicators in Athens and Munich



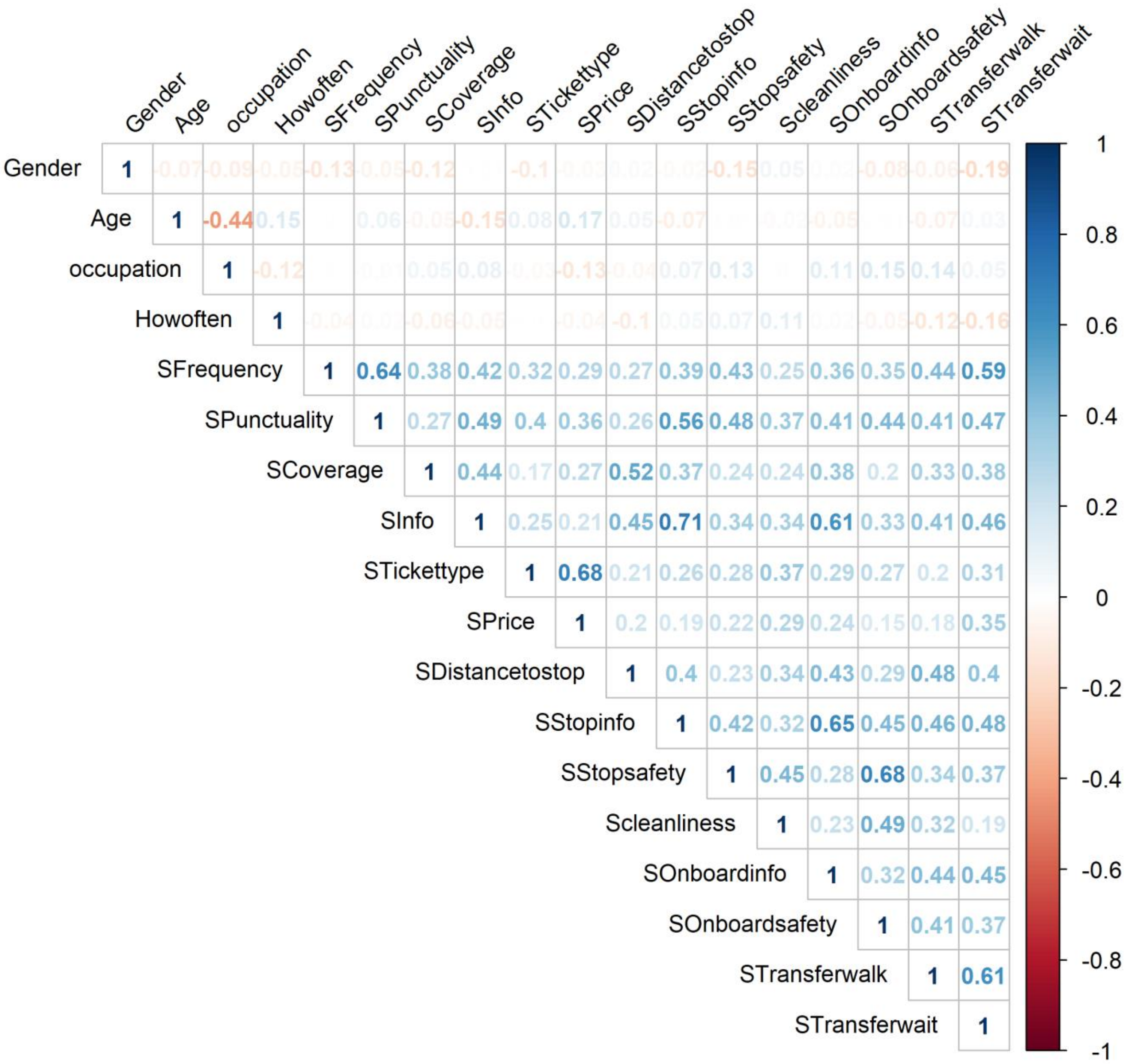
Appendix B 1 Correlation matrix for Satisfaction indicators of bus service in Athens



Appendix B 2 Correlation matrix for Satisfaction indicators of metro service in Athens



Appendix B 3 Correlation matrix for Satisfaction indicators of bus service in Munich



Appendix B 4 Correlation matrix for Satisfaction indicators of metro service in Munich

Appendix C: Factor loadings of factor analysis on satisfaction indicators in Athens for both modes

	Satisfaction about bus service, Athens				Satisfaction about metro service, Athens			
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 1	Factor 2	Factor 3	Factor 4
SFrequency		**0.75**				0.48	0.32	
SPunctuality		**0.67**				**0.97**		
SCoverage		0.47						
SInfo			0.44					**0.82**
STickettype			**0.71**					0.47
SPrice			**0.6**					0.33
SDistancetostop			0.37	0.32			0.44	
SStopsafety	**0.63**				**0.86**			
Scleanliness	**0.79**				0.36			
Sdriver	0.53							
SOnboardinfo								0.38
SOnboardsafety	0.59			0.32	**0.68**			
STransferwalk				**0.6**			0.56	
STransferwait		0.35		0.41			**0.65**	
SS loadings	1.95	1.76	1.5	1.15	1.62	1.43	1.56	1.43
Proportion Var	0.15	0.14	0.12	0.09	0.13	0.12	0.13	0.12
Cumulative Var	0.15	0.29	0.4	0.49	0.13	0.25	0.38	0.5
Factor interpretation	*Quality of service*	*Service production*	*Ticket service (/info)*	*Transfer quality*	*Quality of service*	*Service production*	*Transfer quality*	*Information (/Ticket)*

Factor correlation	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1	0.48	0.44	0.34
Factor 2	0.48	1	0.43	0.42
Factor 3	0.44	0.43	1	0.34
Factor 4	0.34	0.42	0.34	1

Factor Cor.	Factor 1	Factor 2	Factor 3	Factor 4
Factor 1	1	0.33	0.38	0.45
Factor 2	0.33	1	0.37	0.42
Factor 3	0.38	0.37	1	0.38
Factor 4	0.45	0.42	0.38	1

Appendix D: Ordered logit model output for indicators of importance

Athens Bus	Value	Std. Error	t value	p value
IFrequency	-0.28	0.17	-1.62	1.1E-01
IPunctuality	0.43	0.19	2.24	2.5E-02
IDistancetostop	-0.21	0.15	-1.38	1.7E-01
ICleanliness	-0.46	0.18	-2.60	9.4E-03
ITransferwait	-0.34	0.16	-2.08	3.7E-02
1|2	-5.47	1.09	-5.01	5.4E-07
2|3	-3.13	1.05	-2.97	3.0E-03
3|4	0.43	1.11	0.39	7.0E-01
Residual deviance	520.26			
AIC	536.26			

Appendix D.1 Ordered Logit Model output on indicators of Importance of bus service in Athens

Athens Metro	Value	Std. Error	t value	p value
IStopsafety	0.23	0.10	2.35	1.9E-02
1|2	-3.39	0.64	-5.27	1.4E-07
2|3	-0.04	0.43	-0.10	9.2E-01
3|4	2.35	0.46	5.16	2.5E-07
Residual deviance	662.51			
AIC	670.51			

Appendix D.2 Ordered Logit Model output on indicators of Importance of metro service in Athens

Munich Bus	Value	Std. Error	t value	p value
IPrice	-0.32	0.14	-2.32	0.021
ITransferwalk	0.42	0.18	2.27	0.023
ITransferwait	-0.48	0.20	-2.48	0.013
IOnboardsafety	0.36	0.15	2.45	0.014
1|2	-2.61	0.90	-2.91	0.004
2|3	-0.80	0.88	-0.92	0.360
3|4	2.11	0.88	2.39	0.017
Residual deviance	369.98			
AIC	383.98			

Appendix D.3 Ordered Logit Model output on indicators of Importance of bus service in Munich

Munich Metro	Value	Std. Error	t value	p value
IPrice	-0.26	0.14	-1.91	0.056
ITransferwait	-0.29	0.18	-1.64	0.101
ICoverage	0.27	0.15	1.74	0.082
IOnboardsafety	0.30	0.14	2.08	0.038
1|2	-2.50	0.88	-2.83	0.005
2|3	-0.57	0.86	-0.67	0.505
3|4	2.23	0.87	2.56	0.011
Residual deviance	392.59			
AIC	406.59			

Appendix D.4 Ordered Logit Model output on indicators of Importance of metro service in Munich

Declaration Concerning the Master's Thesis

I hereby declare that this master's thesis is my own work and I have documented all sources and material used. This report has not previously been submitted elsewhere for purposes of assessment.

29 August 2018 — Fatemeh Fathisardehaee

Place and date — Signature