Transportation Research Part A 118 (2018) 1–13

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Bayesian networks and structural equation modelling to develop service quality models: Metro of Seville case study

Francisco Díez-Mesa[a], Rocio de Oña[a,*], Juan de Oña[a]

[a] University of Granada, Civil Engineering Department, TRYSE Research Group, Severo Ochoa s/n 18071, Granada, Spain

ARTICLE INFO

Keywords:
Structural equation modelling
Bayesian networks
Service quality
Public transport

ABSTRACT

Service Quality (SQ) in Public Transport (PT) has been a crucial aspect to improve for years because of its strong influence on user satisfaction and its capacity to attract new passengers. Different techniques have been applied for analysing SQ and Structural Equation Modelling (SEM) is one of the most widely used due to its ability to address different kinds of variables and to model a whole phenomenon occurring at one time. Nevertheless, its confirmative nature requires previous knowledge, a hurdle that can be overcome by applying Bayesian Networks (BN) as a technique that learns directly from data without pre-assumptions. The aim of this paper is to apply a novel methodological approach in the field of SQ, based on a two-step process, which combines the techniques of BN and SEM, to model SQ in the Metropolitan Light Rail Transit (LRT) Service of Seville (Spain). In other words, in this paper, the proposed methodological approach has been applied to extract and confirm, directly from data and without necessity of assumptions, the possible relationships among the LRT service characteristics and how they are related with passengers' overall SQ perception. For this purpose, firstly, a BN was automatically learnt from the data and allowed to establish relationships between SQ dimensions describing the service. SEM was then used to check the SQ model and the relationships between the dimensions extracted from the BN. The model fit parameters of SEM and its consistency with the real life expected scenario supported and validated the SQ model designed in this study. Furthermore, the different relationships among dimensions extracted from BN were analysed and support the usefulness and potential of this methodological process that could lead to the development and confirmation of new theories and models in any field of knowledge based on data and expert supervision.

1. Introduction

Over recent years the study of Service Quality (SQ) and its improvement has become a relevant and crucial factor in many fields. Companies, operators and governments have focused on providing high levels of quality in their services in order to improve the current level of customer satisfaction and to attract new clients. It is important to highlight that this relationship between SQ and Satisfaction has been tested in a variety of fields, such as Marketing (Grönroos, 1984), Tourism (Shonk and Chelladurai, 2008), and On-line (Ho and Lee, 2007), etc.

In the specific field of Public Transport (PT), governments and operators have paid great attention to the study and analysis of SQ from the perspective of passengers (Andreassen, 1995; de Oña et al., 2015; Dell'Olio et al., 2011; Nathanail, 2008; Woods and

* Corresponding author.
E-mail address: rociodona@ugr.es (R. de Oña).

https://doi.org/10.1016/j.tra.2018.08.012
Received 6 November 2017; Received in revised form 28 May 2018; Accepted 13 August 2018

Masthoff, 2017, etc.). Their main purpose, apart from addressing passenger satisfaction, has been to improve the ability of PT to compete with the private car by finding alternative modes of transport able to combat problems of mobility, traffic jams and pollution, etc. (Beirao and Cabral, 2007; de Oña and de Oña, 2015; Linda, 2003).

A wide range of techniques applied to the study of SQ in PT can be found in the literature (e.g., Celik et al., 2013; Islam et al., 2016; Kuo and Liang, 2011, etc.). Their main application has been to investigate the influence different service quality factors have on overall SQ. Such knowledge allows administrations and operators to concentrate their efforts and investment in specific areas in need of improvement or of greatest importance to the passengers.

Examples of such techniques are: Importance-Performance Analysis in Weinstein (2000); A Composite Index in de Oña et al. (2016a); Multinomial Logit models in Eboli and Mazzulla (2008); the VIKOR method in Kuo and Liang (2011), etc. Structural Equation Modelling (SEM) has, nevertheless, been one of the more widely used techniques and its application in this field has grown over recent years (e.g. Amin and Isa, 2008; Chen, 2008; Hapsari et al., 2017; Yang et al., 2012; Yilmaz and Ari, 2017, etc.). The main reason behind the increased interest in SEM is its ability to easily address large numbers of variables, both endogenous and exogenous, as well as latent variables (not observed variables) explained as a linear combination of observed variables (Golob,2003). Indeed, SEM is generally considered to be one of the best integrated methods for measuring latent variables and assessing their structural relationships (Chiou and Chen, 2012; de Oña et al., 2015).

These characteristics are all crucial in the study of SQ as it has been defined as such a complex, fuzzy and abstract concept (Carman, 1990; Parasuraman et al., 1985). SQ also depends on a series of underlying observed and unobserved variables. These unobserved variables are commonly denominated as dimensions and are used to provide a better understanding of how customers perceive various service attributes (de Oña et al., 2013).

The suitability of using SEM to study SQ is, therefore, more than justified. Nonetheless, due to the confirmative nature of this technique (de Oña and de Oña, 2015; Golob, 2003), its use requires previous knowledge about how the different dimensions of the SQ models are related. This means that the bibliography or expert knowledge is usually required to develop the models (e.g.,Bagozzi, 1994; Eboli and Mazzulla, 2012; Fillone and Montalbo, 2005; Kamaruddin et al., 2012, etc.) and, depending on the context, this could cause users to miss certain important relationships in the explanation of SQ.

The above has not usually been considered in the specific case of studying SQ in PT (de Oña et al., 2017), where the scientific community agrees that SQ is directly related to and influenced by all its dimensions. Nevertheless, there is evidence in the bibliography to suggest that SQ dimensions could be interrelated and influence each other, and that their relationship with overall SQ or user satisfaction is not always direct but sometimes indirect. Evidence of this can be found in different ways (i.e., directly or indirectly) of relating the dimensions in the PT SQ models found in the literature (e.g., Chou and Kim, 2009; de Oña et al., 2017; Eboli and Mazzulla, 2012; Rahman et al., 2016, etc.). For example, de Oña et al. (2017) grouped service attributes into two latent dimensions: primary attributes (transport service factors), and secondary attributes (comfort and convenience factors). They showed that the secondary factor dimension exerted an effect on the primary attribute dimension, and its relationship with user satisfaction was not direct, but rather indirect through the primary factors dimension.

To overcome this handicap, this paper proposes a two-step procedure based on the combined use of Bayesian Networks (BN) and SEM in a single methodology, as yet unrecorded in the field of SQ. The combined use of BN and SEM has already been applied in the health sector with outstanding results (e.g., Duarte et al., 2011; Scheines et al., 1999; Trentini et al., 2015). Moreover, Kenett and Salini (2011) have presented Bayesian networks and latent variable models as suitable methodologies for analysing customer surveys.

The first step in this approach is to use BN to find certain hypotheses about the relationships between the dimensions which define SQ, followed by a second step which applies SEM to validate these relationships and the model as a whole.

Both techniques can be considered as being complementary due to the different characteristics inherent to each one. On the one hand, BN is an exploratory technique which learns directly from data without the need for pre-assumptions (Heckerman, 1998). Thus, using BN solves the problem of requiring previous knowledge about dimensional relationships. On the other hand, SEM is a confirmatory technique which allows the modelling of a phenomenon in which a set of unidirectional effects or relationships between observed and unobserved variables are established by researchers (de Oña et al., 2015; Golob, 2003). Moreover, this technique examines more than one relationship and tests a set of hypotheses considering a large amount of information at the same time (de Oña et al., 2013; Hair et al., 2010). Therefore, SEM allows researchers to check and validate the relationships of the SQ framework extracted from the BN.

Thus, the proposed two-step methodological approach is appropriate for studying SQ. This is particularly relevant in the field of PT, due to the probable existence of several relationships between the SQ dimensions which might not be discovered and tested by other means.

The proposed research presented here is oriented towards the study of SQ in the Metropolitan Light Rail Transit Service (LRT) of Seville (Spain) by applying this two-step methodology. The main goal is to achieve a model which explains the SQ of this PT service from the point of view of passengers by using data from a Customer Satisfaction Survey (CSS) carried out on 2014.

This paper is structured as follows: Section 2 presents the PT case study, the data collection procedure and the sample characteristics. Section 3 describes the BN and SEM approaches and the two-step methodological process. The results obtained are explained and discussed in Section 4. Finally, the conclusions are reported in Section 5.

2. Methodology

This section describes all the processes that have been applied in this study. First, BN and SEM techniques are defined, followed by a detailed analysis of the two-step methodological process.