Transportation Research Part A xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Transportation Research Part A

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

Improved understanding of the relative quality of bus public transit using a balanced approach to performance data normalization☆

Mark Trompet*, Richard J. Anderson, Daniel J. Graham

Railway and Transport Strategy Centre, Centre for Transport Studies, Department of Civil and Environmental Engineering, Imperial College London, Skempton Building, SW7 2AZ London, United Kingdom

ARTICLE INFO

Keywords:
Normalization
Benchmarking
Quality
Performance
Bus
Key performance indicators

ABSTRACT

In order for bus operators and/or their respective authorities to understand where service quality can improve, it is useful to systematically compare performance with organizations displaying similarities in types of services offered, operational characteristics and density of the service area. These similar characteristics enable peer organizations to benchmark performance once their operational data are normalized for differences in scale of operations. The most commonly used normalization factors for the demand side output are passenger boardings and passenger kilometres. For the supply side output these are vehicle kilometres and vehicle hours. Through twelve years of experience in the International Bus Benchmarking Group (IBBG) a better understanding of differences in service characteristics between 'similar' peers has been achieved, which highlight a challenge for the interpretation of normalized performance. It became clear that relative performance should often not be concluded from performance indicators normalized in a single dimension. Variety between peers in commercial speed, trip length, vehicle planning capacity, vehicle weight and network efficiency result in the need for a bi-dimensional or balanced approach to data normalization. This paper quantifies the variety within these operational characteristics and provides examples of the interpretation bias this may lead to. A framework is provided for use by bus organization management, policymakers and benchmarking practitioners that suggests applicable combinations of denominators for a balanced normalization process, leading to improved understanding of relative performance.

1. Introduction

Based on a definition by Lema and Price (1995) benchmarking is defined as a systematic process of continuously measuring, comparing and understanding organizations' performance and change in performance of a diversity of key business processes against comparable peers anywhere else in the world to gain information which will help the participating organizations to take action to improve their performance.

One of the prerequisites for a successful benchmarking process is the selection of peer organizations based on similar characteristics. In bus benchmarking special focus should be given to the types of bus services offered (e.g. urban, suburban, charter, schoolbus, para-transit, etc.), and the service area characteristics, in particular the density of operations and demand. These similarities enable organizations to compare performance once their data are normalized for scale.

As described by Trompet et al. (2009), variability in comparable performance is welcomed as this results in the identification of best practices, which can lead to improved performance. However, the authors also describe that, even within a fixed group of

* Corresponding author.
E-mail addresses: m.trompet@imperial.ac.uk (M. Trompet), richard.anderson@imperial.ac.uk (R.J. Anderson), d.j.graham@imperial.ac.uk (D.J. Graham).

https://doi.org/10.1016/j.tra.2017.11.019

seemingly comparable operators, there is variation in operating characteristics and environment that results in a subset of operators within that group being less directly comparable. The composition of this less directly comparable subset of operators will differ based on the key performance indicators (KPIs) produced. For example operators from low wage countries will perform well on financial efficiency KPIs when compared to peers operating in higher wage countries or cities, even when a Purchasing Power Parity index is applied. While trend analysis of all operators in the peer group will remain useful to identify best practices (e.g. cost reduction trends will pinpoint which operators are worth talking to for lessons on how to be more productive or efficient), the low wage cities primarily perform better in absolute terms than the high wage city operators due to the wage difference (which is primarily an exogenous factor) and are therefore less directly comparable in absolute financial performance achieved. Apart from differences in operating environments, a second important reason which could lead to peers in a single Key Performance Indicator to be less directly comparable is the bias due to the normalization factor chosen; this is the main focus of this research.

This paper provides evidence to bus organization management, policymakers and benchmarking practitioners that conclusions from relative quality and other performance differences observed in a single Key Performance Indicator (e.g. only normalized by a single denominator) can be biased. Due to variety between peers of operating characteristics such as commercial speed, trip length, vehicle capacity, vehicle weight and network efficiency (e.g. variety in amount of deadheading) it is necessary that performance is at least reviewed from two different dimensions, in order to obtain an improved and more balanced understanding of relative performance. This is especially important for bus organizations that have 'extreme' values in any of the five identified variable operating characteristics.

To achieve this objective the remainder of this paper is structured as follows. Section 2 reviews normalization factors used or suggested in a number of relevant previous and on-going bus public transport benchmarking initiatives. Section 3 describes the International Bus Benchmarking Group (IBBG) dataset which is used for this study. Section 4 describes the variability of five operational characteristics that can lead to a skewed perspective of relative performance when only one denominator is used for normalization. In Section 5 a framework is then presented that schematically suggests which combinations of two denominators could be used to obtain a more realistic, balanced view of relative performance. The section also elaborates on how this framework fits within the steps of the whole benchmarking process and how the resulting improved understanding of relative performance will help bus organization and/or authorities identify and subsequently prioritize areas for improvement and will facilitate the identification of peers which are worth reaching out to for information on best practices. Section 6 discusses the use of revenue service planning capacity kilometres as a normalization factor in more detail, as within the IBBG, this has especially resulted in improved performance comparability and understanding. Conclusions are drawn in Section 7.

2. Normalization in transport benchmarking and literature

An overview of public transport benchmarking initiatives has been provided in a variety of reports and papers (e.g. TCRP, 2010; EQUIP, 2000; Geerlings et al., 2006; Gudmundsson et al., 2005). With regards to urban bus performance measurement specifically, Mulley (2004) describes the process and lessons learned from the UK Bus benchmarking Group which was based on the benchmarking handbook developed by EQUIP (2000). Alongside the references mentioned above, the papers and project reports listed under references (Hensher and Daniels, 1995; Fielding et al., 1985a, 1985b, 1978; Stappenhorst, 2009; Phillips, 2004; TCRP, 2003; ISOTOPE, 1998; MARETOPE, 2003; MVA Limited, 2003; Urban Transport Benchmarking Initiative, 2004; Badami and Haider, 2007; Mackie and Nash, 1982) have also been reviewed to understand how practitioners have normalized for scale. The focus in the review was to find recommendations towards improved normalization factors and on discussions in relation to the possible bias of using a single denominator for normalization.

The reports and papers generally underline the importance of peer selection and normalization for successful benchmarking. Complete sets (or examples) of main key performance indicators used are often provided. These confirmed that the most commonly used normalization factors for the demand side output are passenger boardings and passenger kilometres. For the supply side output these are vehicle kilometres and vehicle hours. Useful discussions have been provided on the use of vehicle hours as a preferred denominator over vehicle kilometres (Hensher and Daniels, 1995; Fielding et al., 1985a). The majority of reviewed papers also mention the use of seat kilometres for supply side normalization. This led to the addition of Section 6 to this paper in which the authors discuss the role and possible bias of seat capacity kilometres as a normalization factor in a benchmarking exercise.

The benchmarking projects and literature reviewed did not explain in some level of detail why a certain normalization denominator was chosen (over an alternative one) and what bias could be involved in using that single denominator. This paper aims to contribute to this area.

Fielding et al. (1985b) and Stappenhorst (2009) use cluster analysis to create comparable sub groups to improve comparability within performance indicators. This is feasible in a situation where the total number of participating organizations in a benchmarking group is sufficiently large; the size of each cluster of 'more comparable' peers should also be sufficient for performance comparison. The downside to clustering peers into more comparable subsets before benchmarking analysis is that possibly valid lessons and best practices from 'lesser' comparable peers are discarded. Clustering is done before the performance comparison is executed. This paper discusses a post performance comparison alternative, by understanding relative performance through a bi-dimensional normalization process.

3. The data

The data used for this study have been collected through the International Bus Benchmarking Group (IBBG), which is facilitated