CONSTRUCTING THE URBAN PUBLIC TRANSPORT SYSTEM QUALITY INDICATOR

Irina Pticina, Irina Yatskiv
Transport and Telecommunication Institute, Latvia

Abstract

In this research the possibility of developing the composite indicator characterising the urban system of public transport (urban public transport quality index - UPTQI) is considered. The development of the composite indicator in terms of the initial data describing the transport system (TS) currently operating in German cities (EUROSTAT) is presented. Some variants of OECD algorithm realisation for developing the composite indicator have been applied.
As a result, this work shows an alternative of constructing the composite indicator characterising urban public transport system, serving as a basis for drawing comparison between urban public transport system quality in various cities, and for assessing the influence of various characteristics on the overall estimate.

Keywords: public transport quality, composite indicator, weighting

1 Introduction

The purpose of public transport is rendering safe, reliable, punctual, accessible, non-polluting and cost-effective transport services to people. To estimate the appropriateness of the existing services of public transport and the offers concerning introduction of new services, service quality indicators imply: punctuality and maintenance rate, time etc. So, there are many characteristics, which need to be measured for understanding the actual quality of public transport system, and sometimes it is difficult to estimate the actual situation on the basis of this multivariate set of data. The benchmarking of the urban public transport can apply a scalar or composite indicator to be constructed - instead of using a set of initial characteristics as a basis.
The problem of developing an integrated indicator of quality is a subject that quite a number of theoretical researches [1],[2],[3] are focused on. In general, the indicator is a quantitative measure received on the basis of a number of observations, capable of showing relative position of an object (for example, a country or city) in the respective field. The composite indicator is a function from sub-indicators and weights:

$$CI_i^t = w_1 x_{1,i}^t + w_2 x_{2,i}^t + \cdots + w_m x_{m,i}^t \tag{1}$$

where CI_i^t – value of composite indicator for object i (i=1…n) at time t,

$x_{j,i}^t$ – value of sub-indicator j (j=1…m) for object i at time t,

w_j – weight associated with sub-indicator j (j=1…m).

The Organization of Economic Cooperation and Development (OECD) proposed the 10 steps algorithm for constructing the composite indicator [4]. The output of a composite indicator is a set of scores indicating the relative performance of a object within a set of objects. The essential disadvantages which should be considered at construction are – the ambiguous process of a choice of scales and primary variables, the possible gap in data and also many variants of weight estimation method in scalar indicator function. The weights can be obtained by methods that can be divided up into two groups [4]: based on statistical analysis (factor analysis, regression analysis etc) and based on the opinion of experts (Conjoint Analysis, Analytic Hierarchy Process – AHP etc).
Also, there are composite indicators based on the input data (that of the subgroup) having equal weights. That is to say, each group (if the data had been already grouped before), or the initial data in total (if each indicator is considered separately) equally contributes to the composite indicator [5], [6], [7]. If all the input data (variables) initially have identical weight and will subsequently unite into groups according to certain parameters - this can lead to unbalanced structure of the composite indicator (groups including a larger quantity of variables, will make a larger contribution to the composite indicator and vice versa).

2 Theoretical framework for quality index construction

Let us consider as the object – the urban system of public transport, as composite indicator - urban public transport quality index - UPTQ and as the sub-indicators - the particular quality characteristics.
For calculating weights and aggregating primary indices into the composite indicator methods, based on Principal Components and Factor analysis (PCA/FA) model and benefit of the doubt approach (BOD) were considered.

2.1 Weights estimation based on factor analysis model

This approach used by Nicoletti G., Scarpetta S., Boylaud O. [7]. At first, the correlation structure of the data was checked. It is needed for preliminary analysis of the common factors existing. Then, according to standard approach to FA, identifying a certain number of latent factors smaller than the number of sub-indicators implies data representation and the factor structure rotation if necessary.

Let a_{ij} - the *j* factor loading for *i* variable and $D[f_j]$ - variance explained by the *j* factor. In this case introduce the normalization of factor loading as

$$a_{i,j\,norm} = \frac{a_{i,j}^2}{D[f_j]} \tag{2}$$

and the weight for *i* variable as maximum of factor loading multiplied the proportion of total variance for corresponding factor

$$w_i = \max\left(a_{i,j\,norm}\right)\frac{D[f_j]}{\sum_{k=1}^{m} D[f_k]} \tag{3}$$

2.2 Weights estimation based on benefit of the doubt approach (BOD)

The method endogenously determines country-specific weights that explicitly take account of a country's own choices and achievements across primitive dimensions of performance. In the BOD approach, the composite indicator is defined as the actual/benchmark performance ratio and the weights are city specific. Optimal weights are obtained by solving the constrained optimisation as linear programming problem [4]:

$$CI_c^* = \max(w_{qc}) \sum_{q=1}^{Q} I_{qc} w_{qc} \qquad (4)$$

with following constraints: $\sum_{q=1}^{Q} I_{qk} w_{qk} \leq 1 \, ; w_{qk} \leq 0$;

for $\forall k = 1, \ldots, M; \quad \forall q = 1, \ldots, Q$;

where CI_c^* – value of composite indicator for city c,

I_{qc} – normalised value of sub-indicator q for city c,

w_{qk} – weight associated with sub-indicator q for city c,

M – number of cities and *Q* – number of sub-indicators.

3 Construction of urban public transport quality index

3.1 The initial data

To develop the composite indicator, let us consider the initial data describing public TS of 37 German cities (EUROSTAT). 8 sub-indicators describing the urban public transport system have been selected (see Table 1).

Table 1 List of sub-indicators

№	sub-indicators	Code
1	Proportion of journeys to work by public transport (rail, metro, bus, tram)	x1
2	Length of public transport network / land area	x2
3	Number of stops of public transport per km2	x3
4	Cost of a monthly ticket for public transport (for 5-10 km)	x4
5	Number of stops of public transport per 1000 pop.	x5
6	Number of stops per 1km of public transport network	x6
7	Proportion of public transport network on fixed infrastructure /Proportion of public transport network on flexible routes	x7
8	Proportion of the area used for transport (road, rail, air, ports)	x8

3.2 Multivariate analysis and normalisation of data

The correlation analysis of data was fulfilled and the most of Pearson correlation values lie in [0.3;0.61]. The highest correlation (-0.61) between sub-indicators x2 (Length of public transport network / land area) and x6 (Number of stops per 1km of public transport network). Also, the significant correlations between next pairs of sub-indicators: 0.60 for x8 (Number of stops per 1km of public transport network) and x1 (Proportion of journeys to work by public transport (rail, metro, bus, tram); -0.58 for x8 and x5 (Number of stops of public transport per 1000 pop.); 0.56 for x8 and x2 (Length of public transport network / land area).

The data was normalised through the Re-scaling normalising procedure and lies in (0;1). For one initial sub-indicator x4 (Cost of a monthly ticket for public transport) (for 5-10 km) the distance from maximum was calculated. This is related to our assumption that the appeal of public transport is also connected with a low fare; therefore, we have made maximum transformation with respect to x4, to make the overall quality index dependence monotonous and positive with respect to all variables – i.e., making the overall quality index value rise with the growth of sub-indicator.

3.3 Weighting and Aggregation

All the calculations for the first method were performed through using the package Statistica/Win5.5. The principal components were received and the first four components explain 85% of the total variance, while the eigenvalues of the first three components exceed 1 and that of the fourth component is almost 1. These 4 components have been left for the subsequent analysis. A structure close to a simple one was obtained at the rotation by Biquartimax normalized method; the factor loadings after rotation are presented in Table 2. The weighting factors for sub-indicators, calculated according to the formulas stated in Section 2.1, are shown in Table 3. Moreover, weight values are presented with respect to the case of equal estimation. The Table 3 highlights the weight values exceeding the average weight value and testifying that sub-indicators data have a larger significance.

Table 2 Factors loadings

	Factor 1	Factor 2	Factor 3	Factor 4
x1	**0.745454**	0.082645	0.403463	-0.207046
x2	**0.862092**	-0.097813	-0.320193	0.169933
x3	0.480532	**0.844947**	-0.106159	-0.087637
x4	-0.178090	-0.082360	0.082194	**0.885755**
x5	-0.452298	**0.673714**	-0.297925	0.321472
x6	-0.465451	**0.764284**	0.178133	-0.279356
x7	-0.005520	-0.080465	**0.908061**	0.089484
x8	**0.803666**	-0.076322	0.054367	-0.400519
Expl.Var	2.628659	1.787437	1.231352	1.213795
Prp.Totl	0.328582	0.223430	0.153919	0.151724

Table 3 Weights values

methods	X1	X2	X3	X4	X5	X6	X7	X8
PCA	0.105	**0.140**	**0.135**	**0.148**	0.086	0.110	**0.155**	0.122
Equal	0.125	0.125	0.125	0.125	0.125	0.125	0.125	0.125

To calculate weights through BOD method, the package MathCad 14 was used where the linear programming problem was solved. In our calculations we imposed the requirement for each sub-indicator to weight at least 10% and no more than 15% of the total. The specific values of weights w_{qc}, have been calculated for each city c; we have omitted them due to certain limitations imposed upon the volume of this article.

3.4 Composite indicator estimation

The final step is constructing a composite score. As aggregation methods, the Additive aggregation one was selected. Therefore, 3 indicator values - CI_{PCA}, CI_{BOD} and CI_{egual} - have been obtained with respect to each city, where the values obtained through PCA and BOD methods were used as weights; equal weights were used as well. The composite indicator values and the respective rating of a number of cities are presented in Table 4 and Figure 1.

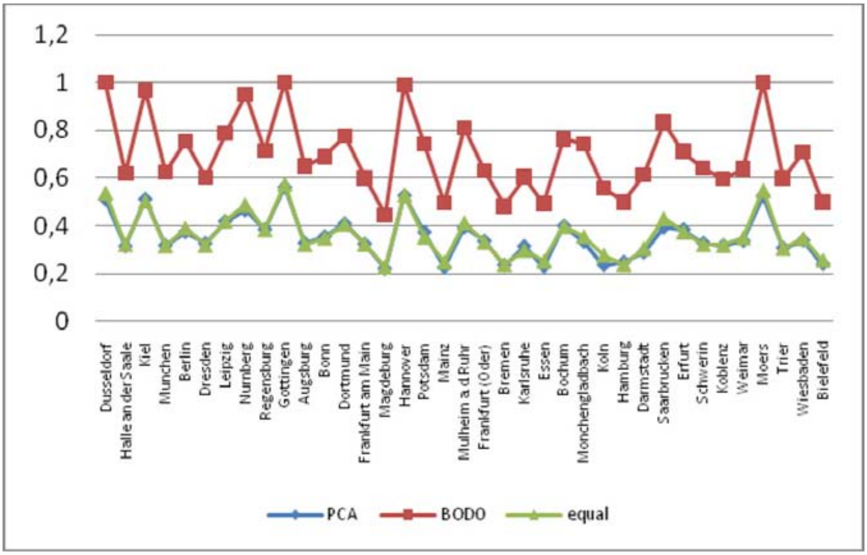

Figure 1 Values of UPTQI of 37 Germany cities

Table 4 Composite indicators values

cities	CIPCA	Rank PCA	CIBOD	Rank BOD	CIequal	Rank equal
Dusseldorf	0.56281	1	1.00013	3	0.57570	1
Halle an der Saale	0.52689	2	0.99018	4	0.53142	4
Kiel	0.52626	3	1.00126	1	0.54997	2
Munchen	0.51181	4	0.96843	5	0.50421	5
Berlin	0.51138	5	1.00035	2	0.53931	3
Dresden	0.46561	6	0.95221	6	0.48942	6
Leipzig	0.41887	7	0.78623	9	0.42066	8
Nurnberg	0.40932	8	0.77648	10	0.40730	10
Regensburg	0.39998	9	0.76306	11	0.39874	11
Gottingen	0.39414	10	0.80930	8	0.41339	9
...						
Frankfurt (Oder)	0.33274	20	0.74379	14	0.35535	15
...						
Schwerin	0.24674	31	0.49880	33	0.23769	35
Koblenz	0.23859	32	0.49901	32	0.25608	32
Weimar	0.23426	33	0.55754	31	0.27722	31
Moers	0.23402	34	0.48223	36	0.23682	36
Trier	0.22799	35	0.49178	35	0.25191	33
Wiesbaden	0.22490	36	0.49453	34	0.24746	34
Bielefeld	0.21879	37	0.44595	37	0.22644	37

Analysing the results of calculating the three indicator alternatives, we can point out that the same cities are ranking top 10 except for Augsburg that has fallen out of the top ten ratings according to the index CI_{PCA}; in terms of the other indices, however, the city remains at the 7th position. The CI_{PCA} values do not exceed 0.563, which corresponds to the city of Düsseldorf coming out to the top according to CI_{PCA}. At that, the value CI_{BOD} with respect to the city assumes the magnitude which is little different from the maximum one - 1.00013, whereas its rank is the third one. The least value CI_{PCA} has been achieved by the city of Bielefeld and equals 0,219. At that, Bielefeld is ranking last also in terms of CI_{BOD} and CI_{egual}. At that, the largest difference in ranks pertains to the city of Frankfurt (Oder).

4 Conclusions

Composite indicators are very common in some fields like economic and business statistics. They are used in a variety of policy domains such as industrial competitiveness, sustainable development. In this research, the urban public transport quality index (UPTQI) was constructed on the basis of different approaches to weight estimation: PCA, BOD and - the most criticised – equal weighting. There is a high correlation between all the pairs of indicator values – namely, 0.99 (see Fig. 2). The significant correlation show robustness of results. Also for correct comparising we can note that CI_{BOD} values are located within the interval (0.4;1.0), while CI_{PCA} does not exceed 0.6. For the sake of convenience, the indicator values can be normalized (to the interval 0;1) (see Figure 3).

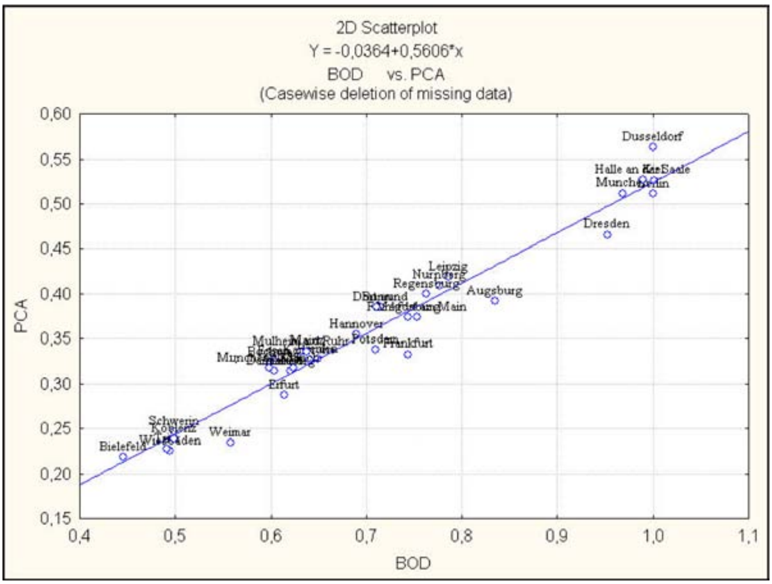

Figure 2 Scatterplot diagram

Comparison of the endogenous weight method (BOD) with the method of using fixed weights (PCA) indicates an insignificant bias, in terms of indicating a lower relative performance, being more diverse in their achievements among primitive performance dimensions. When the endogenous weight method is used, the performance ranking of countries is altered so that countries with greater diversity improve their relative performance while the relative performance of countries having less diversity may either rise or fall. In general, the indicators show the relative position of development of each object (Urban Public Transport System - UPTS); in particular, relating to the ones with the best or least performance – both the overall one and the one pertaining to different fields. This is measured with scores and ranks. The relative performance of each object is also indicated by its ranking among all the other ones. The index had to be updated regularly when new information is available and this is especially important when policy makers use the results of this benchmarking. The obtained results stated in the article provide a basis for the further investigation of composite indicator construction methods to be used for estimating the Public Transport system.

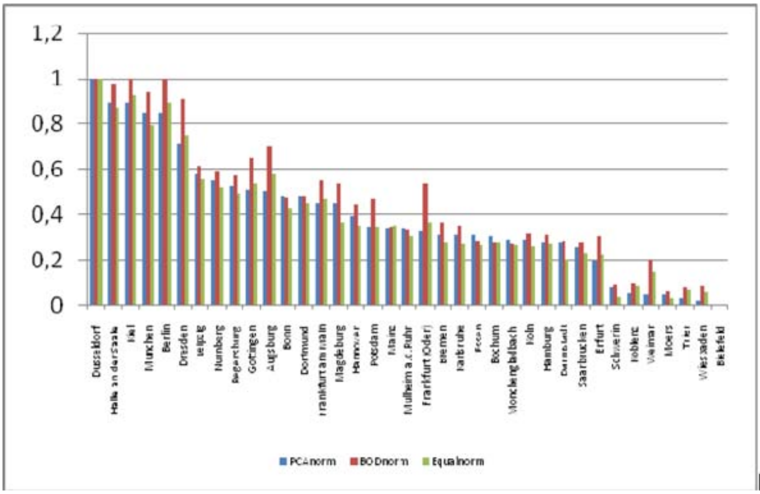

Figure 3 Normalized values of UPTQI

Acknowledgments

The article is written with the financial assistance of European Social Fund. Project Nr. 2009/0159/1DP/1.1.2.1.2/09/IPIA/VIAA/006 (The Support in Realisation of the Doctoral Programme "Telematics and Logistics" of the Transport and Telecommunication Institute).

References

[1] Borodkin F. M., Aivazyan S. A. Social indicators. Moscow: UNITY, 2006. 603 p.

[2] I.Gertsbakh, I.Yatskiv. Dynamic Classification: Economic Welfare Growth in EU During 1995-2004. \\ Proceeding of International Conference "Data mining-2006", 11-13 July 2006, Prague. WITPress. 2006, p.53-62.

[3] Ilya Gertsbakh, I.Yatskiv, O.Platonova. Constructing Social and Economic Indicators for EU Countries Using Dynamic Classification: Case Studies. Proceeding of International Conference "Data mining-2008", 26-28 May 2008, Cadiz, p.153-162

[4] Nardo M. a.o. Handbook on Constructing Composite Indicators: Methodology and User Guide. OECD Statistics Working Paper, 2005/3. 108 p.

[5] Osberg L., Sharpe A. International Comparisons of Trends in Economic Well-Being. http://www.lisproject.org/publications/liswps/242.pdf

[6] Smith J. Guide to the Construction and Methodology of the Index of Economic Well – Being. http://www.csls.ca/iwb/iewb-guide.pdf

[7] Nicoletti G., Scarpetta S. and Boylaud O. (2000) – Summary indicators of product market regulation with an extension to employment protection legislation, OECD Economics Department Working papers No. 226.