Bootstrapping DEA Scores for Road Safety Strategic Analysis in Brazil

Jorge Tiago Bastos

Department of Transportation, Federal University of Paraná, Centro Politécnico s/n, Jardim das Américas
Curitiba, Paraná, 81531-990, Brazil
E-mail: jtbastos@ufpr.br; www.dtt.ufpr.br

Yongjun Shen

Transportation Research Institute, Hasselt University, Wetenschapspark, bus 6
Diepenbeek, 3590, Belgium
E-mail: yongjun.shen@uhasselt.be; www.uhasselt.be/IMOB-EN

Elke Hermans

Transportation Research Institute, Hasselt University, Wetenschapspark, bus 6
Diepenbeek, 3590, Belgium
E-mail: elke.hermans@uhasselt.be; www.uhasselt.be/IMOB-EN

Tom Brijs

Transportation Research Institute, Hasselt University, Wetenschapspark, bus 6
Diepenbeek, 3590, Belgium
E-mail: tom.brijs@uhasselt.be; www.uhasselt.be/IMOB-EN

Geert Wets

Transportation Research Institute, Hasselt University, Wetenschapspark, bus 6
Diepenbeek, 3590, Belgium
E-mail: geert.wets@uhasselt.be; www.uhasselt.be/IMOB-EN

Antonio Clóvis Pinto Ferraz

Department of Transportation Engineering, School of Engineering of São Carlos, University of São Paulo,
Trabalhador Sãocarlense Av.
São Carlos, São Paulo, 13.566-590, Brazil
E-mail: coca@sc.usp.br; www.eesc.usp.br/portaleesc/en/

Abstract

In this paper, three risk indicators on road safety are combined into a composite indicator in order to assess the overall fatality risk for the 27 Brazilian states using the so-called multiple layer data envelopment analysis model. The states are first clustered and next, a range of bootstrapped scores is generated to manifest the estimated variability in the road safety performance. Bootstrapping the original DEA scores showed to be a useful strategy to assess the robustness of the states' ranking.

Keywords: Road safety – Bootstrapping – Data Envelopment Analysis - Clustering – Brazilian states.

1. Introduction

In macro-level road safety research, country/states comparisons are used to evaluate the road safety situation. Most commonly, three indicators are used for comparison or monitored over time: traffic fatalities per inhabitants, traffic fatalities per vehicle and traffic fatalities per traveled kilometer. Although any of these indicators is useful to classify the road safety level of a place, it is possible that different conclusions are drawn depending on the adopted indicator. In other words, the resulting rank based on each of the three indicators might be considerably different and the choice of a single reference is sometimes puzzling. In order to overcome this difficulty, the Data Envelopment Analysis (DEA) methodology is a tool with convenient features to combine these three indicators into a single index or composite indicator capable to express the road safety performance in a more adequate perspective.

DEA is a term used to designate a "data oriented" approach in which mathematical programming methods are applied to handle large numbers of variables and relations (multiple input and multiple output – MIMO problems); that way, it has become an attractive tool to deal with complex problems. The first concepts on this theme emerged in 1957 for measuring the productive efficiency of industries (1), although its most widely known basic form, the CCR model, was proposed by Charnes, Cooper and Rhodes in 1978 on the evaluation of production processes through exploring the relation between the amount of inputs and outputs. The entities under study, responsible to convert inputs into outputs, are named decision-making units – DMUs (2,3).

DEA's attractiveness of no functional form requirements for input information contributed to its dissemination and application in various fields. Additionally, the DEA method involves the optimal attribution of weights for each input/output (in this study the three indicators), so that each DMU receives the weight combination that places it in the best possible position among all the other DMUs based on a general score. This score or composite indicator results from the product between the indicators and their respective weights. Therefore, the advantage of this process is that there is no excuse of unfair weight allocation and consequently unfair assumptions for ranking DMUs.

Suppose five DMUs – A, B, C, D and E (which can be five different companies producing the same product) each one responsible to transform a certain amount of inputs (e.g. raw material) into outputs (e.g. final product). The DMU capable to produce more outputs with the same amount of input is then the most efficient one of the set. According to the example in Table 1, DMU D is the most efficient one, scoring 1.00.

Table 1. Numerical example to illustrate the basic concept of DEA.

DMU	Input	Output	Efficiency score
A	5.00	2.00	0.40
B	8.00	4.00	0.50
C	9.00	5.00	0.56
D	6.00	6.00	1.00
E	4.00	3.00	0.75

The rest of the DMUs are underperforming units, and their efficiency can be expressed in relation to the most efficient DMU; this is, DMU A presents 40% of the efficiency of DMU D (being the worst performing of the five units), B presents 50%, C presents 56% and E presents 75%. Now the input (horizontal axis) and output values (vertical axis) of the five DMUs are plotted in Figure 1.

The slope of the line connecting each point to the origin indicates the efficiency score of the corresponding DMU; and the line connecting DMU D to the origin represents the efficiency production frontier, since no other DMU can be more efficient than DMU D. The area between the production frontier and the horizontal axis is called production possibility set, and all the DMUs are "enveloped" within this set, hence the term data envelopment analysis.

The underperforming DMUs might achieve the best attainable performance (similar to DMU D) if they are able to increase their efficiency. For example, DMU A could be as efficient as DMU D by increasing its output value maintaining its input amount unchanged (A' point) or by decreasing its input amount without changing the output value (A" point). Therefore, the production frontier constructed based on DMU D's is valuable within a benchmarking process, which is an important tool aimed at promoting the transference of good practices from high to low performing DMUs, as it can be deduced from which DMUs a particular underperforming DMU can learn.
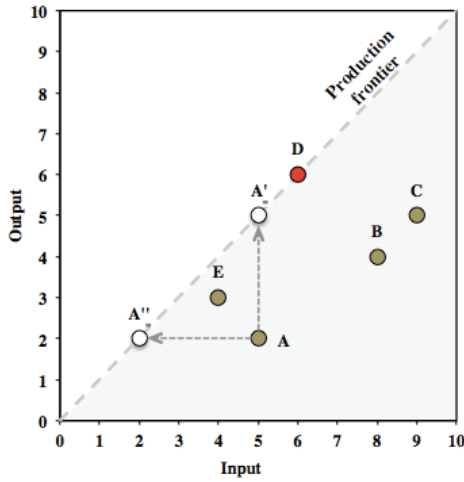
Fig. 1. Graphical representation of the efficiency production frontier.

For simple examples as the one presented previously, this graphical representation is a useful visual description of the fundamental mechanisms of DEA, yet difficult to reproduce when more parameters are considered in the investigation of the DMUs efficiency. The usual approach is to employ a linear programming formulation for DEA to estimate the efficient production frontier and compute the relative, optimum efficiency of each DMU.

However, there is still some discussion about the statistical properties of the produced scores. The obtained optimum scores might represent a very particular combination of parameters (e.g. indicators) and weights that may be unlikely to be found in realistic circumstances. Thus, testing other possible combinations is a valuable strategy to overcome this limitation. Complementarily, the seeking for more realistic results requires the imposition of weight restrictions, enveloping the mathematical problem into more plausible and realistic solutions.

The application of a deterministic-nature technique, such as DEA, to a non-deterministic problem, such as road safety, requires the index values to be submitted to a sensitivity analysis procedure (bootstrapping), after which DEA scores are expressed throughout a range of possible values, instead of a single number.

In addition to these issues, the sensitivity analysis has also the purpose of simulating different data sceneries, since the exposure and fatality data we use is subjected to some level of uncertainty. That is actually

the main motivation for testing the sensitivity in this study. Therefore, the objective of this paper is to propose this inedited approach of bootstrapping DEA scores for road safety evaluation at the state level and demonstrate the aforementioned methodological process using road safety indicator data of the 27 Brazilian states (BR-27).

The paper has the following structure. Firstly, in Section 2, the data envelopment analysis method is presented as an approach capable of offering a scientifically sound composite indicator on road safety. In the same section, the bootstrapping technique is introduced as a method to test the sensitivity of DEA scores. Afterwards, in Section 3, the 27 Brazilian states are classified into clusters. In Section 4 the DEA technique in combination with the bootstrapping process is applied for each cluster and road safety scores manifesting the overall diagnosis are obtained and discussed. Finally, Section 5 contains the conclusive comments and recommendations for future research.

2. DEA in road safety research

2.1. *Computing a composite indicator*

In the road safety framework, the technique of DEA has been applied for composite indicator investigations, e.g. (4,5), through the application of a minimization model (since traffic fatality indicators are undesirable outcomes intended to be minimized), as expressed in Eq. (1).

$$OIS_s = \min \sum_{i=1}^{p} w_i y_{i,s} \qquad (1)$$

$$\text{subject to} \quad \sum_{i=1}^{p} w_i y_{i,s} \geq 1, \quad s = 1, \dots, n,$$

$$w_i \geq 0, \quad i = 1, \dots, p.$$

OIS_s is the optimum index score (or composite indicator – CI) of the *s-th* DMU, $y_{i,s}$ is the *i-th* indicator of the *s-th* DMU or state, w_i is the weight attributed to indicator $y_{i,s}$, *n* is the total number of DMUs and *p* the total number of indicators. The scores indicating the best performers will present an OIS value equal to one, and underperforming DMUs will present a score larger than one. Subsequential needs of expressing a hierarchy in the set of selected indicators motivated the development of the Multiple Layer DEA-based Composite Indicator model (ML DEA-CI) (6). By solving Eq. (2), the

composite indicator based on a *K*-layered hierarchy of *p* indicators can be calculated for each state *s*, where u_{fK} is the weight given to the *f-th* category in the *K-th* layer and $w_{fK(K)}$ denotes the non negative internal weights associated with the indicators of the *f-th* category in the *K-th* layer; the sum of all $w_{fK(K)}$ within a particular category is equal to one.

$$OIS_s = \max \sum_{f_{K_1}=1}^{p(K)} u_{f_K} \left(\sum_{f_{K-1} \in A_K^{(K)}} w_{K-1}^{(K-1)} \right.$$
$$\left. (1 \sum_{f_K \in A_{k+1}^{(k+1)}} w_K^{(k)} (1 \sum_{f_{k-1} \in A_{f_1}^{(3)}} w_2^{(2)} (\sum_{f_1 \in A_{f_2}^{(2)}} w_{f_1}^{(1)} y_{f_1 s})))) \right. \tag{2}$$

However, the flexibility in selecting the most favorable weights for each DMU forbids the comparison on a common basis (7,8). To allow direct comparisons, the cross-index score (CIS) could be computed by applying Eq. (3), where CIS_s is the cross-index score of the s-th DMU, $y_{i,s}$ is the *i-th* indicator of the *j-th* DMU which CIS is to be computed, w_i is the weight attributed to indicator $y_{i,s}$, *n* is the total number of DMUs and *p* the total number of indicators:

$$CIS_s = (1/n) \sum_{s=1}^{n} \sum_{i=1}^{p} (w_i y_{i,s}) \tag{3}$$

2.2. Sensitivity analysis of DEA scores

The key idea of bootstrapping, firstly addressed by Efron in 1979, is to resample from the original data to create replicate datasets, which mimic the original unknown sampling distribution of the estimates of interest (9); in other words, an empirical distribution is artificially constructed. The principle is as follows: given a specified random variable θ *(y,f)*, depending on both *y* and the unknown distribution *f*, estimate the sampling distribution of θ on the basis of the observed data *y* (10,11). In the flowchart of Fig. 2, the steps to bootstrap the index scores are summarized.

Given a set of *n* DMUs and their corresponding *p* indicators' values $(y_{1i,s},...,y_{1p,n})$, the first step is to normalize these values. Then, the DEA model is run in order to compute the original index scores $(\theta_{1s}, ..., \theta_{1n})$. These scores are resampled with replacement according to the empirical distribution, generating the set $\beta_{1s*}, ..., \beta_{1n*}$ and, afterwards, they should be smoothed, forming the set $\theta_{1s*}, ..., \theta_{1n*}$.
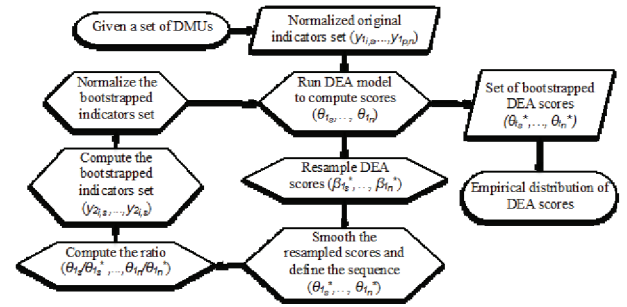

Fig. 2. Flowchart containing the steps of the bootstrapping procedure.

Smoothening is the term attributed to the process of incorporating the properties of the original set of index scores into the resampled set, through the use of reflection methods (12,13), represented by the application of Eq. (4) to (6), in which:

$$h = \left[\frac{4}{(p+q+2)} \right]^{[1/(p+q+4)]} * N^{[-1/(p+q+4)]} \tag{4}$$

$$\tilde{\theta}_i^* = \begin{cases} \beta_i^* + h.\varepsilon_i^* & if \ \beta_i^* + h.\varepsilon_i^* \geq 1 \\ 2 - \beta_i^* - h.\varepsilon_i^* & otherwise \end{cases} \tag{5}$$

$$\theta_i^* = \bar{\beta}^* + \frac{(\tilde{\theta}_i^* - \bar{\beta}^*)}{\sqrt{\left(1 + \frac{h^2}{\hat{\sigma}_\theta^2}\right)}} \tag{6}$$

$$\hat{\sigma}_\theta^2 = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \left(\hat{\theta}_i - \hat{\bar{\theta}}\right)^2 \tag{7}$$

- *h* is the smoothing parameter defined by the number of indicators *p*, the number of outputs *q* and the number of DMUs *N* presenting the set of original optimum scores $(\hat{\theta}_{1s},...,\hat{\theta}_{1n})$,

- ε_i^* is a random drawn from a standard normal distribution,

- $\tilde{\theta}_i^*$ is the smoothed sampled score $(\tilde{\theta}_{1s}^*,...,\tilde{\theta}_{1n}^*)$ adjusted to be ≥ 1.0,

- $\hat{\sigma}_\theta^2$ is the plug in estimator of the variance of the original scores $(\hat{\theta}_{1s},...,\hat{\theta}_{1n})$,

- $\hat{\sigma}_{\theta}^2$ is the plug in estimator of the variance of the original scores $(\hat{\theta}_{1s},\dots,\hat{\theta}_{1n})$,

- β_i^* is the average of the resampled score $(\beta_{1s}^*,\dots,\beta_{1n}^*)$,

- θ_i^* is the variance corrected bootstrapped score $(\theta_{1s}^*,\dots,\theta_{1n}^*)$.

The next step is to compute the ratio between the original and the bootstrapped scores ($\theta_{1s}/\theta_{1s}^*,\dots, \theta_{1n}/\theta_{1n}^*$) and then multiply it by the original indicators set, producing the new indicators set $y_{2i,s}, \dots, y_{2p,n}$, which is normalized prior to computing the second set of index scores. By running the DEA model for the second time, the first loop (constituted by the steps described in the hexagonal forms) is completed and it is repeated t times until an adequate set of bootstrapped scores $\theta_{ts*},\dots, \theta_{tn*}$ is obtained.

3. Data set

Brazil consists of 26 member states and 1 federal district. Here, they are generically referred to as BR-27, meaning the set constituted by Paraná (PR), Rio Grande do Sul (RS) and Santa Catarina (SC) – Southern states; Espírito Santo (ES), Minas Gerais (MG), Rio de Janeiro (RJ) and São Paulo (SP) – Southeastern states; Distrito Federal (DF), Goiás (GO), Mato Grosso (MT) and Mato Grosso do Sul (MS) – Center-Western states; Alagoas (AL), Bahia (BA), Ceará (CE), Maranhão (MA), Paraíba (PB), Pernambuco (PE), Piauí (PI), Rio Grande do Norte (RN) and Sergipe (SE) – Northeastern states; Acre (AC), Amapá (AP), Amazonas (AM), Pará (PA), Rondônia (RO), Roraima (RR) and Tocantins (TO) – Northern states. The three indicators intended to be composed into a single index are presented in Table 1 for each state.

It is important to emphasize that the three possible rankings to obtain based on the mentioned indicators provide different figures. AM for example, is the state with the 2nd lowest MR, although it sustains only de 22nd place regarding the FR1 and the 17th place regarding FR2. Accordingly, taking MR as a reference, PA would be one of the best performing Brazilian states in terms of road safety; however, concerning the FR1 indicator, PA would present a quite unfavorable

performance; finally, with respect to FR2, PA would present a close to average performance.

4. Cluster analysis

In order to provide more realistic comparisons among the states, the bootstrapping was conducted separately for the defined clusters. This step is motivated by the contrasting figures existing between the Brazilian states. The gap in road safety background and standards might be substantially large between two states; at a level that this comparison is not even reasonable to draw.

Table 2. Traffic fatality related indicators for Brazilian states.

		Indicators		
		MR (fat/10^5 inh)	FR1 (fat/10^4 veh)	FR2 (fat/10^9 km)
	AM	12.90	8.55	46.68
	PA	16.46	12.88	71.96
	BA	16.65	10.32	68.94
	RJ	16.67	5.85	47.32
	SP	17.27	3.49	32.08
	RN	18.38	8.04	60.17
	AP	19.23	11.03	56.43
	AC	19.34	9.33	62.81
	RS	19.52	4.45	41.02
	MA	19.94	16.21	104.68
	PB	21.28	11.53	82.02
	PE	21.46	10.78	85.55
	MG	21.70	6.14	48.61
	CE	22.06	10.96	96.46
DMU	DF	23.94	5.04	42.50
	AL	24.73	17.70	128.19
	SE	28.43	14.03	95.89
	GO	30.38	7.79	57.88
	SC	30.47	5.66	50.43
	RS	19.52	4.45	41.02
	RR	30.61	10.98	74.45
	ES	30.97	8.65	76.79
	PR	31.44	6.53	63.60
	MS	32.33	8.22	63.61
	PI	32.48	17.48	131.41
	MT	36.35	9.74	66.62
	RO	36.52	10.13	66.08
	TO	37.12	13.14	75.66
	BR-27	21.53	6.50	63.27

The objective of cluster analysis in this study is to reach a categorical structure that fits a certain set of observations. In other words, the intention is to sort the observations into groups such that the degree of "natural association" is high among members of the same group and low for members from different groups (14).

In hierarchical clustering, the data are not partitioned into classes in a single step; instead they are first separated into a few broad classes, each of which is further split into minor classes and so on, until final classes are generated. The researcher needs to decide at which stage of the segmentation it is (more) convenient to stop (15).

A combination of six available parameters was used for the clusters construction of the BR-27: GDP *per capita*, motorization rate (rate between the vehicle fleet and the population of a place), fatalities per vehicle rate, fatalities per traveled kilometer rate, highway density, and a regional parameter. Hierarchical clusters were constructed using Ward's method, leading to four main clusters; Fig. 3 shows a map with the location of each cluster.

Efficacious actions to improve road safety in a certain state are more likely to generate desirable results in states with a similar background. This is particularly important in a country with continental dimensions and very contrasting backgrounds in the different regions. These contrasts refer to many aspects, from which the most obvious concern socioeconomic issues.
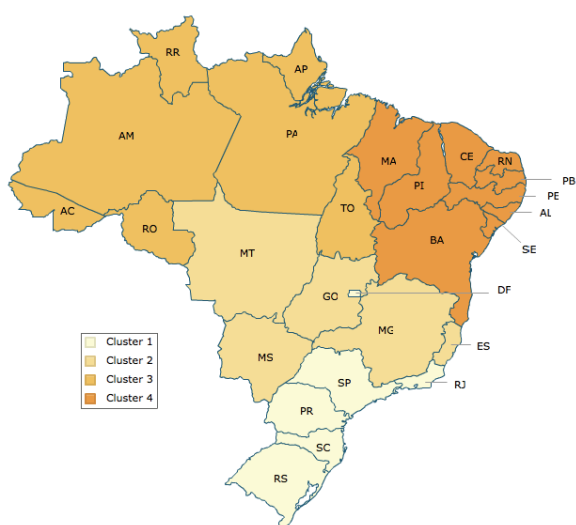
To a certain extent, the four clusters are representative of four particular regional contexts in Brazil. Cluster 1 contains the Southern, two states and the Federal District, where the national capital is located. They represent the most urbanized and developed areas of the country. Cluster 2 is formed by three Center-Western states and the other two Southeastern states, which in general present most indicators close to the national average.

Cluster 3 consists of the Amazon region, characterized by low population density, it contains isolated areas with more intense human occupation (predominantly in the states' capital cities). Cluster 4, coincident with the Northeast region, is characterized by the contrast between the more urbanized and touristic coastal areas and the interior with low population density and historic social problems due to the semi-arid climate.

5. Application and results

The ML DEA-CI was computed with the software *Lingo* (developed by *Lindo Systems*) for a set of indicators using average values between 2009 and 2011 for the 27 Brazilian states. The indicators were combined through a hierarchic structure, in which traffic fatalities per vehicle (FR1) and traffic fatalities per traveled kilometer (FR2) were combined into a single indicator (FR), which was combined with traffic fatalities per inhabitants (MR), according to Fig. 4.
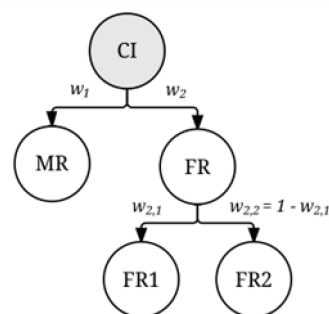

Fig. 4. Hierarchical structure used to compute the composite indicator (CI).

This hierarchy implementation was based on the more similar nature of FR1 and FR2. As stated in Section 2, they bring very similar information in different refinement levels regarding the risk exposure parameter.

Fig. 3. Clusters location.

In order to avoid a unilateral weight distribution, but at the same time still preserving adequate flexibility to the model, the shares w_1*MR and w_2*FR were limited in the model definition to vary between 10 and 90% of the index value. Likewise, to control the weight distribution on FR and avoid an exaggerated weight attribution to either FR1 or FR2, the weights $w_{2,1}$ and $w_{2,2}$ were limited to vary within a 20% maximum range. The adoption of such procedure is founded on the assumption that there is no reason to attribute much more weight on any FR indicator, since they are supposed to present high association.

Following a calibration process, additional weight restrictions were inserted on the shares w_1*MR and w_2*FR, being that w_2*FR is larger than or equal to $1.5*w_1*MR$. The intention of this intervention is to reduce the importance attributed to a somewhat biased indicator (the MR) in the composite indicator value in the case of a less-motorized state.

In more detail, the traffic fatality per inhabitant rate is substantially affected by the motorization level in a non- linear pattern. This is, until a certain stage, as higher the motorization of a place, the higher the MR tends to be, but at a certain breakdown point, this tendency is reversed, the increase in motorization is associated with a lower MR (the case of the most developed and consecrated countries in terms of road safety).

Before running the model, the data were normalized using the distance to a reference method (16). The boxplot diagrams in Fig. 4 to 7 offer an overview of the distribution of the computed CIS values.

The OIS and the CIS computed in each cluster can be consulted in Table 3, which also contains the average CIS value ($\overline{CIS}$) and the standard deviation (σ_{CIS}) considering all boostrapped scores. The results are ordered according to the $\overline{CIS}$. The boxplot diagrams in Fig. 5 to 8 offer an overview of the distribution of the computed CIS.

The table and the boxplot diagrams suggest SP, MG, AM and RN as good examples in their regional contexts, instead of the vague and predictable indication of SP as the only model state. On the other hand, PR, TO, MT and PI appear as the most underperforming states in their clusters. The availability of a range of scores instead of a merely point estimate provides a more clear classification and the identification of non-obvious performance differences between states from the same cluster; this is, best performing DMUs had their top position endorsed and underperforming ones had their ineptitude on controlling traffic fatalities emphasized.

Table 3. Numerical results.

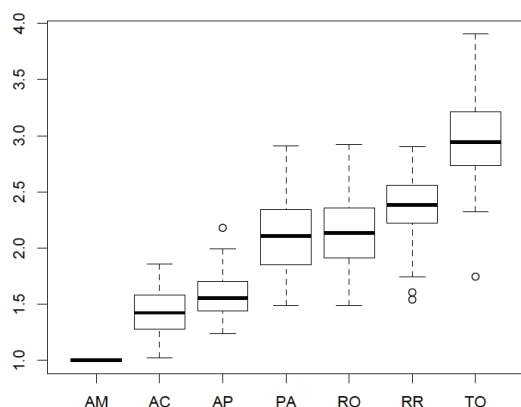| Clusters	State	OIS	CIS	$\overline{CIS}$	σ_{CIS}
Cluster 1	SP [1st]	1.00	1.00	1.00	0.02
	RS [2nd]	1.22	1.24	1.54	0.40
	RJ [3rd]	1.34	1.41	1.91	0.37
	DF [4th]	1.38	1.39	2.00	0.44
	SC [7th]	1.62	1.64	2.75	0.61
	PR [12th]	1.88	1.90	3.66	0.88
Cluster 2	MG [6th]	1.00	1.00	1.00	0.00
	GO [11th]	1.24	1.25	1.54	0.19
	MS [14th]	1.34	1.35	1.82	0.23
	ES [16th]	1.46	1.49	2.15	0.26
	MT [18th]	1.48	1.50	2.23	0.29
Cluster 3	AM [5th]	1.00	1.00	1.00	0.00
	AC [9th]	1.23	1.24	1.45	0.21
	AP [10th]	1.27	1.29	1.59	0.22
	PA [15th]	1.46	1.49	2.13	0.34
	RO [20th]	1.46	1.49	2.16	0.31
	RR [21st]	1.52	1.54	2.34	0.31
	TO [23rd]	1.72	1.75	3.00	0.41
Cluster 4	RN [8th]	1.00	1.00	1.02	0.06
	BA [13th]	1.08	1.10	1.30	0.28
	PB [19th]	1.30	1.31	1.81	0.44
	PE [17th]	1.29	1.30	1.84	0.45
	CE [22nd]	1.36	1.38	1.94	0.49
	MA [25th]	1.54	1.58	2.56	0.62
	SE [24th]	1.61	1.62	2.78	0.67
	AL [26th]	1.83	1.86	3.56	1.01
	PI [27th]	2.01	2.03	4.29	1.04

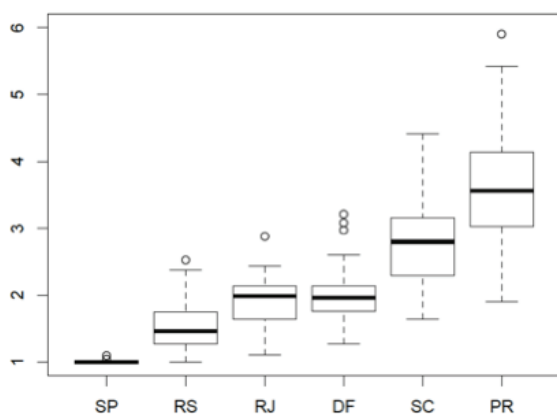Fig. 5. Cluster 1 bootstrapped CIS.

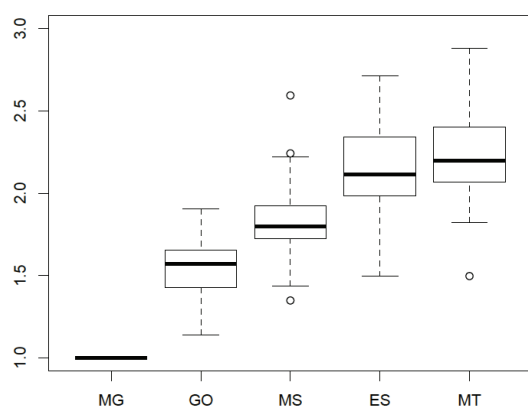Fig. 6. Cluster 2 bootstrapped CIS.

Fig. 7. Cluster 3 bootstrapped CIS.

Similar procedures have been applied for analyzing road safety throughout the Brazilian states, however with different focuses (17, 18).
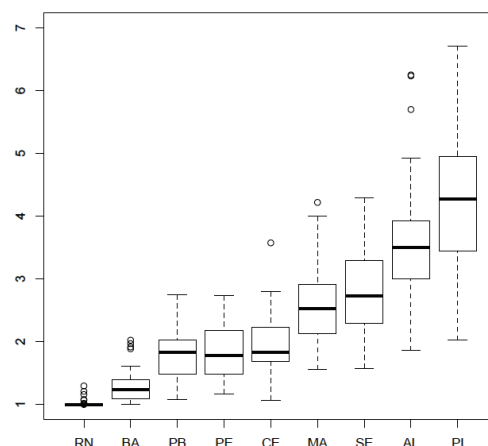

Fig. 8. Cluster 4 bootstrapped CIS.

6. Conclusions

This paper described a set of procedures aiming to complement and improve the application of DEA research on ranking DMUs with respect to road safety. This analysis was motivated by the absence of such model application for evaluating developing areas such as Brazil. The possibility to present a methodology capable to deliver a composed traffic fatality related diagnosis and to deal with data uncertainty issues, a common dilemma in road safety, constitutes an important contribution in the field.

The DEA model computes the optimum solution for every state considering its respective cluster; thus, there is no excuse that the performance of a particular state was underestimated due to an unfair weight allocation process. This results in the OIS value, which is the CI used to identify the best-performing (safer) state; and the performance of the underperforming states is measured as a share of the best-performing state score.

DEA possesses some features that make its use very attractive for the topic of this paper. Firstly, obtaining a rank is not the only purpose of DEA; it can be seen as a benchmarking tool, since the performance of the DMUs is evaluated in relation to the best-performing ones. In terms of road safety, it is particularly valuable, because states or countries that do not show good road safety results can learn from successful initiatives taken in best-performing states/countries.

The application of DEA techniques in association with cluster analysis and bootstrapping procedures

brought singular and valuable insights for diagnosing the traffic fatality situation in a "per state" perspective. DEA research allowed the use of combined road safety indicators, which were traditionally considered individually, telling only one part of the story.

Clustering is important to provide a more realistic comparison; in other words, one state can better learn from another's performance if they have similar contexts. This topic is particularly important in a huge country like Brazil, where significant contrasts in terms of road safety background can be found in different states, requiring more realistic comparisons and demonstrate possible achievable examples of relatively good performances to be followed by underperforming states.

Last but not least, bootstrapping the DEA scores, a procedure for the first time applied to road safety DEA results in a developing country, was the chosen technique to test the sensitivity of the results to possible variations in road safety Brazilian data, which might be subjected to the influence of unfavorable aspects with respect to data quality and availability. It showed how susceptible the road safety related data is to uncertainties and to which extent it affects the research results.

Hence, the bootstrapping permitted the road safety assessment with the consideration of uncertainties probably present in the indicator data for the country, manifested by the delivered values' ranges. Generally, the median or average values of the bootstrapping results provided quite similar rankings in comparison to the original ranking based on the DEA index scores (both OIS and CIS). The advantage now is that we by means of Fig. 4 to 7 graphs revealed some knowledge on how likely each rank configuration could be obtained.

For future research, the combination of DEA and bootstrapping for road safety research is still capable of delivering evaluation throughout other relevant perspectives, e.g. traffic fatality target setting and safety performance indicators research. Thus, the proposal of investigating the sensitivity of road safety related information is applicable to many issues inside this field due to the frequent issue of incomplete or not entirely ideal information to base the studies on.

References

1. Farrell MJ. The measurement of productive efficiency. Journal of the Royal Statistical Society. 1957;120:253–90.
2. Cooper WW, Lawrence M. Seiford, Joe Zhu. Data envelopment analysis: history, models and interpretations. Handbook on Data Envelopment Analysis [Internet]. Springer; 2011 [cited 2013 Oct 29]. p. 1–39. See: http://www.springer.com/business+%26+management/operations+research/book/978-1-4419-6150-1.
3. Cooper WW, Seiford LM, Tone K. The basic CCR model. Data Envelopment Analysis: A Comprehensive Text With Models, Applications, References and Dea-Solver Software. Springer; 2000. p. 21–40.
4. Bax C, Wesemann P, Gitelman P, Shen Y, Goldenbeld C, Hermans E, et al. Developing a Road Safety Index [Internet]. 2012 [cited 2013 Feb 8] p. 185. See: http://www.dacotaproject.eu/Deliverables/DaCoTA_D4.9_developing%20a%20RSI%20deliverable.pdf.
5. Shen Y, Hermans E, Ruan D, Wets G, Brijs T, Vanhoof K. Evaluationg trauma managment performance in Europe: a multiple-layer data envelopment analysis model. Transportation Research Record. 2010;2148:739–56.
6. Shen Y, Hermans E, Brijs T, Wets G. Data Envelopment Analysis for Composite Indicators: A Multiple Layer Model.
7. Doyle J, Green R. Efficiency and Cross-Efficiency in DEA: Derivations, Meanings and Uses. The Journal of the Operational Research Society. 1994 May;45(5):567.
8. Sexton TR, Silkman RH, Hogan AJ. Data envelopment analysis: Critique and extensions. New Directions for Program Evaluation. 1986;1986(32):73–105.
9. Efron B. Bootstrap Methods: Another Look at the Jackknife. The Annals of Statistics. 1992 Jan 1;7:569–93.
10. Chernick MR. What is bootstrapping? Bootstrap methods: a practitioner's guide. John Wiley & Sons, Inc.; 1999. p. 1–22.
11. Davison AC, Hinkley DV. The basics bootstraps. Bootstrap methods and their application. 2nd Revised edition. Press Syndicate of the University of Cambridge; 2003. p. 11–69.
12. Simar L, Wilson PW. Sensitivity Analysis of Efficiency Scores: How to Bootstrap in Nonparametric Frontier Models [Internet]. 1998 [cited 2013 Oct 29]. See: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.44.1.49?journalCode=mnsc.

13. Walden JB. Estimating vessel efficiency using a bootstrapped data envelopment analysis model. Marine Resource Economics. 2006;21 (The MRE Foundation): 181–92.

14. Anderberg, M. R. The broad view of cluster analysis. Cluster analysis for applications. London: Academic Press Inc. Ltd.; 1973. p. 1–9.

15. Everitt B. Chapters 1 and 3. Cluster analysis. United Kingdom: British Library C.I.P. Data; 1980.

16. OECD. Handbook on constructing composite indicators: methodology and user guide [Internet]. Organisation for Economic Co-operation and Development. 2008 [cited 2013 Oct 29]. See: http://www.oecd.org/els/soc/handbook onconstructingcompositeindicatorsmethodologyanduser guide.htm

17. Bastos, J. T. (2014). *Road safety strategic analysis in Brazil: indicator and index research* (Doctorate thesis). Hasselt University/ University of São Paulo, Diepenbeek/ São Carlos. Retrieved from http://www.teses.usp.br/teses/disponiveis/18/18144/tde-08042015-103747/pt-br.php

18. Bastos, J. T., Shen, Y., Hermans, E., Brijs, T., Wets, G., & Ferraz, A. C. P. (2015). Traffic fatality indicators in Brazil: State diagnosis based on data envelopment analysis research. *Accident Analysis & Prevention*, *81*, 61–73.