

Benchmarking Analysis of Road Safety Levels for an Extensive and Representative Dataset of European Cities

Katerina Folla ¹, Paraskevas Nikolaou ², Loukas Dimitriou ², George Yannis ¹

¹ National Technical University of Athens, Iroon Polytechniou 9, Zografou 157 73, Athens, Greece

² Lab. for Transport Engineering, Department of Civil and Environmental Engineering, University of Cyprus, 75 Kallipoleos Str, P.O. Box 20537, 1678 Nicosia, Cyprus

Abstract. According to European urban sustainable planning guidelines, road safety corresponds to one of the most important elements of cities' performance. Several methods have been developed over the years for supporting policy-making, towards the improvement of road safety levels mainly at the national level. In this study, a methodological framework is developed, extending the macro-level (national) analysis and focusing at a higher spatial resolution, that of urban regions. The methodological approach is based on benchmarking analysis, able to suitably rank alternative cases/regions with distinctive characteristics within a multivariate comparative framework and on the investigation of the components that affect their ranking. In particular, an extensive and representative dataset from 101 European regions is collected and analyzed, incorporating their socio-economic, demographic and road infrastructure characteristics. Then, Data Envelopment Analysis (suitably adapted to road fatalities framework) has been developed, evaluating the urban regions' road safety performance over a period of 9 years. The resulted region ranking is further examined by using Tobit regression models for identifying the components that appear to affect their performance in different extents providing a valuable guiding 'tool' for experience/knowledge-transfer and policy-making. The datasets and the results are presented and discussed in detail, such as they will be useful not only for demonstration purposes but also they will be suitable as a benchmark for researchers and practitioners.

Keywords: Benchmarking Analysis, Sustainable Urban Mobility, Road Traffic Fatalities, Stochastic Frontier, Data Envelopment Analysis.

1 Introduction

Road traffic injuries have nowadays been recognized as one of the most important public health issues that require concerted efforts for effective and sustainable prevention. Worldwide, an estimated 1.35 million people are killed in road accidents, with road traffic injuries being the leading cause of death for children and young people aged between 5 and 29 years old [1]. The European Commission (EC) has committed to improve the safety of the European road network. On that purpose, EC has adopted a

Road Safety Programme which aims to halve the number of road deaths by 2020, compared to the 2010 level. This target followed an earlier target set in 2001 to cut road fatalities by 50% compared to 2001, which was almost achieved [2].

In the European Union (EU), in 2018, there were around 25,100 fatalities in road accidents in the EU-28. This is a decrease of 21% compared to 2010. With an average of 49 road fatalities per one million inhabitants, European roads are by far the safest in the world, however, EU seems to be far from reaching the target of halving the number of road deaths by 2020 [3]. Additionally, despite the significant progress in road safety improvement in the European Union, there are still significant variations between the European countries in terms of road safety performance, which may be attributed to the different socioeconomic environments, cultures and behavior characteristics, modal shares, etc.

At the regional level, the disparities may be even larger in EU. An analysis of the regional distribution of road fatalities shows that, while the overall number of fatal road accidents in EU has fallen since 1991, significant regional disparities remain, providing an insight into the conditions that favor low fatality rates in road transport [4].

Within this context, more and more countries are taking steps to improve their road safety situation, while there is also an urgent need not only for the countries but also for the regions to work together more closely, to identify the common problems and improve their road safety performance by learning lessons from existing best practices in other countries. Consequently, international benchmarking of the road safety performance of various regions could be a useful tool for national policymakers as well as international organizations working on road safety.

The objective of the current study is to evaluate the road safety performance of EU urban regions, taking into account the evolution of road safety, transport and economic characteristics over the decade 2008-2016. On that purpose, a representative sample has been created with data for 101 regions (at NUTS 2 level) from 13 European countries over the same period. Data concerning the vehicle fleet, road infrastructure and economic conditions of these regions were selected mainly from the EU CARE database and Eurostat. In this study, a methodological framework is developed, extending the macro-level (regional) analysis and focusing at a higher spatial resolution, that of urban regions. More specifically, suitably adapted to road fatalities' Data Envelopment Analysis was developed, evaluating the regions' road safety performance and identifying the best and under-performing regions. Then the road safety performance (efficiency score) of the region was concerned for measuring the effect that the different explanatory variables have on it and in this way shedding light on the factors that particular focus must be given for improving the road safety performance of the EU regions.

The remaining of this paper is organized as follows. Section 2 presents the results of previous work related to the current study. Section 3 describes the dataset used for the current analysis. Section 4 describes the methodology followed for the purposes of the current paper. Section 5 presents the results of the applied methodology. The paper ends concluding remarks in Section 6.

2 Literature Review

This section reviews the existing studies conducted on the assessment and benchmarking of road safety performance using various methodological techniques.

In order to comparatively evaluate the road safety performance of various countries or regions, risk indicators defined as the ratio of road safety outcomes (mainly fatalities) and measures of exposure (e.g. vehicle-kilometers traveled, population, registered vehicles, etc.) are often used. However, in the vast majority of the cases, the benchmarking results of a group of countries or regions are not consistent, with the ranking position of a country or a region being different depending on the risk indicator used. Within this context, numerous road safety indicators have been developed for international or interregional comparisons and monitoring of road safety progress. For instance, *Al-Haji* [5] proposed a road safety development index (RSDI) based on three domains of road safety, the outcomes (fatality rates), the user (road user behavior) and the system (vehicles, roads, socioeconomic conditions, enforcement and organization level). In the SUNflower next study [6], three different types of performance indicators were distinguished, i.e. road safety performance indicators, implementation performance indicators and policy performance indicators which were integrated into an overall road safety index. *Yannis et al.* [7] developed an SPI for the evaluation of the safety level of the road network to be used as a benchmark for cross-region comparisons.

Additionally, Data Envelopment Analysis (DEA) has been used in various studies, as a performance measurement technique. *Shen et al.* [8] developed a model in order to evaluate the road safety performance of the 27 EU countries and assess whether the road safety outcomes registered in a country correspond to their level of exposure. *Shen et al.* [9] further analyzed the road safety performance of the EU countries in the period 2001–2010 by developing a Malmquist productivity index (MPI) and then evaluated EU countries in terms of both efficiency and technical changes. In 2018, *Nikolaou & Dimitriou* [10], analyzed the road safety performance of EU-23 countries over a decade (2005–2014) considering their socio-economic and demographic background supporting the decision-making process by using short-term and long-term targets.

At the regional level, there are few studies having implemented DEA for benchmarking road safety performance, all referring to regions within the borders of a country. More specifically, *Alper et al.* [11] estimated the relative efficiency of 197 local municipalities in traffic safety in Israel during 2004–2009, using DEA. The inputs reflected the resources allocated to the local municipalities and outputs concerning road safety outputs accidents, while safety performance indicators were also used as intermediate variables. Moreover, the mortality rate and fatality rate were aggregated into a composite indicator through a multiple-layer DEA composite indicator model, aiming to identify the optimum combination of indicators' weights for 27 Brazilian states [12]. Finally, a double frontier DEA cross efficiency method was used to evaluate the road safety performance of 31 Iranian provinces based on safety management data and the number of fatalities in 2016 [13].

For analyzing and identifying the effect of different components on the regions' road safety performance a suitable regression model was developed, namely, Tobit. The

combination of the DEA method and Tobit regression has also been developed in several researches. For instance, *Tasnim and Afzal* [14] used Tobit regression model for observing what macro factors affect the efficiency of the knowledge spillover theory of entrepreneurship. For identifying the efficiency of 59 countries DEA method was developed. Another study used DEA for measuring the efficiency of 30 university science parks and Tobit regression model for analyzing the impact of possible influential factors [15].

3 Data Description

As shown in Section 2, numerous studies are investigating the assessment of road safety performance at the international level by developing indexes/indicators in order to rank the countries based on their road safety performance. This study aims to develop a methodological framework, extending the macro-level (regional) analysis and focusing at a higher spatial resolution, that of urban regions. In order to achieve adequate and meaningful results during comparisons, similar regions have to be included in benchmarking analyses, which have a similar level of development, motorization and a similar type of transport system [5]. On that purpose, an extensive and representative dataset from 101 European regions (NUTS 2 level) was collected and analyzed, incorporating their socio-economic, demographic and road infrastructure characteristics.

The nomenclature of territorial units for statistics, (NUTS) is a geographical nomenclature subdividing the economic territory of the European Union (EU) into regions at three different levels (NUTS 1, 2 and 3 respectively, moving from larger to smaller territorial units). The NUTS classification is based on the following main principles: (a) the NUTS regulation defines minimum and maximum population thresholds for the size of individual NUTS regions and (b) NUTS favors administrative divisions and thus, if available, administrative structures are used for the different NUTS levels. The population size of the NUTS 2 regions ranges between 800.000 inhabitants and 3.000.000 inhabitants [4].

The importance of spatial analyses has been demonstrated in the literature, while studies have shown that smaller regions may present significant similarities in road safety outcomes across the EU, without being affected by the presence of national borders [16; 17]. *Eksler* [18] has shown that the mortality ratios of EU NUTS 2 regions show greater variations compared to those estimated at national level, which is attributed to their demographic structure and population dynamics. Population density has been found to have a significant influence on road mortality at regional level. Generally, higher level of road safety in densely populated regions can be due to lower travelling speed, modal share variations, more developed infrastructure and differences in vehicle fleet patterns [19].

Thus, data for 101 NUTS 2 regions covering 13 European countries for the period 2008-2016 were collected. Countries for which data were available are depicted in Figure 1. In detail, Figure 1 shows all NUTS 2 regions included in the database. A geographical variation of fatality rates is obvious, with higher fatality rates being presented in the Southern and Eastern European countries and the lowest fatality rates in the

Northwestern countries. At a regional level, regions in Bulgaria and Romania have the worst road safety performance in terms of fatality rates (more than 100 fatalities per million population), while regions in Sweden, Austria, and Germany present the best performance (less than 25 fatalities per population). Accident fatality rates, however, vary at a large scale among the regions within the borders of a country, indicating the need for investigating road safety problems at a regional level.

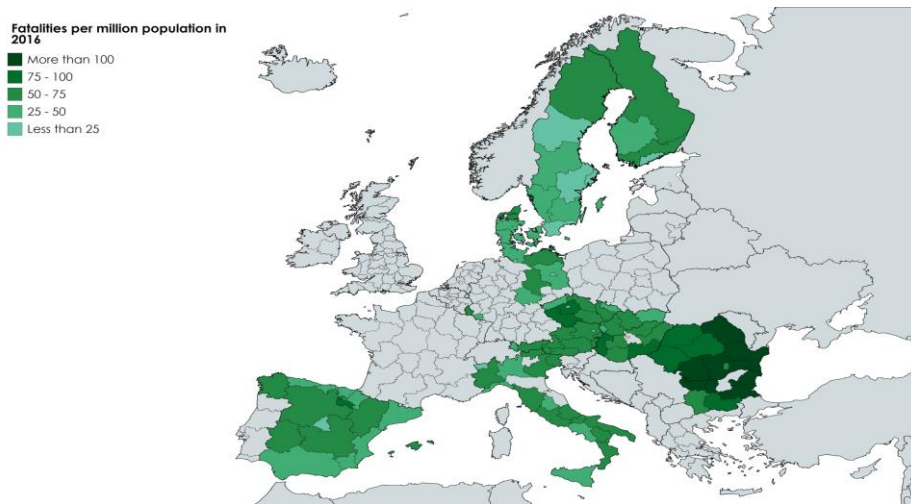


Fig. 1. Fatalities per million population by NUTS 2 region in the EU, 2017

Therefore, time-series fatality data were retrieved from the CARE database, the EU database with disaggregate road accident data. Additionally, data on population and the Gross Domestic Product (GDP) for these regions were collected from Eurostat. Finally, data related to the vehicle fleet (in total and by type of vehicle) and to the available transport infrastructure by NUTS 2 region were collected from Eurostat. Transport infrastructure related data concern the length of the motorway network as well as of the remaining road network.

4 Methodology

This study develops two procedure concepts of benchmarking analysis (DEA) together with Tobit regression. The concept of DEA was conducted for measuring the efficiency of the regions' road safety performance considering their socio-economic, demographic and road infrastructure context.

As adapted from *Charnes et al.* [20] output-oriented DEA measures the efficiency scores for different Decision Making Units (DMUs) based on the concept of maximizing the product (i.e., output/s) keeping constant the input data. However, the concept developed in this study is that road fatalities (output) must be minimized and not maximizing as the original formation of the DEA method. Therefore, the suitable adapted to road safety DEA method is presented below:

$$\begin{aligned}
& \min \sum_{i=1}^p w_i y_{i,s} \\
& \text{Subject to } \sum_{i=1}^p w_i y_{i,s} \geq 1, s = \\
& \quad 1, \dots, n \\
& \quad w_i \geq 0, i = 1, \dots, p
\end{aligned} \tag{1}$$

Where, $y_{i,s}$ is the i^{th} indicator of the s^{th} DMU (region), w_i is the weight attributed to indicator $y_{i,s}$, n is the total number of DMUs (i.e., 101) and p is the total number of indicators. The best-performing regions (in terms of road safety) will be presented with an efficiency equal to one and the under-performing regions (in the same terms) will be presented with efficiency scores below to one.

The concept of Tobit was employed for analyzing what determinants affect the efficiency of the region's road safety performance. The reason for developing Tobit and not a classic regression model was due to the capability of Tobit of analyzing censored data and thus Tobit is also known as the censored regression model. The mathematical formation of Tobit model is based on *Tobin* [21] work and can be seen below:

$$\begin{aligned}
y_t &= X_t \beta + u_t & \text{if } X_t \beta + u_t > 0 \\
&= 0 & \text{if } X_t \beta + u_t \leq 0 \\
& & t = 1, 2, \dots, N.
\end{aligned} \tag{2}$$

Where, N is the number of observations (i.e., 101), y_t is the dependent variable (efficiency scores), X_t is a vector of independent variables, β is a vector of unknown coefficients, and u_t is an independently distributed error term assumed to be normal with zero mean and constant variance σ^2 .

Therefore this study followed the above methodological framework for identifying the road safety performance of the 101 EU regions in each time instant and analyzed the measured efficiencies for estimating the effect that each component has on them. The next section presents all the results from the DEA-Tobit analysis.

5 Results

This section presents the results from the developed methodological framework. The first implementation was the development of the DEA method which was suitably adapted to the road safety framework. In particular, road fatalities were concerned as output in DEA and the socio-economic, demographic and road infrastructure characteristics of the regions were concerned as inputs. The goal of this DEA, as mentioned above, was the minimization of road fatalities taking as constant the input characteristics. Figure 2 presents the outcome of this implementation. As can be seen 11 out of the 101 regions appeared to best-perform out of the years and thus policymakers should take a closer look at what they are doing in terms of road safety in order to adopt strat-

egies and policies that are helping them to create and maintain their road safety performance.

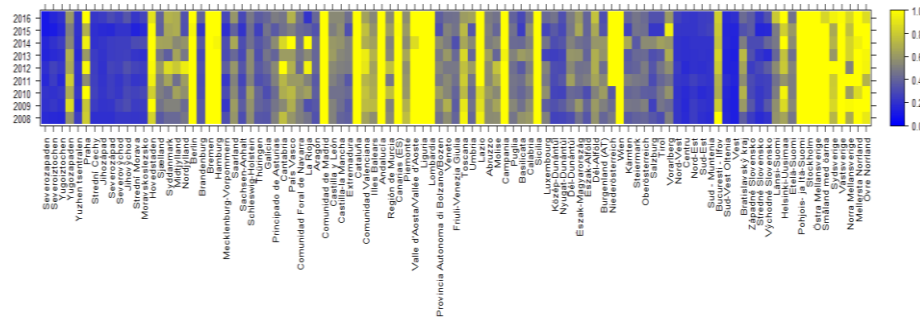


Fig. 2. Road Safety Performance of the 101 EU Regions Throughout the Years (2008-2016).

Overall from the DEA analysis, we were able to identify under and best-performing regions in terms of road safety. This identification is very important to local authorities to observe the overall “picture” of their road safety performance throughout the years and adopt the strategies that best-performing regions (benchmarks) are following. But the question raised here is: Which counties are considered as benchmarks?

In order to answer this question, we must clarify that for every different explanatory variable to road fatalities create a different effect on their road safety performance and thus for the different socio-economic, demographic and road infrastructure characteristics we have different benchmark regions. For instance, Figure 3 presents the DEA frontier concerning the explanatory variable “Motorcycles”. Taking examples from this figure, an under-performing region should focus on the benchmarking regions (e.g. Catalunya, Lazio, etc.), which are best-performing, and adopt the strategies that might they follow especially on the context of motorcycles (laws, prohibitions, etc.).

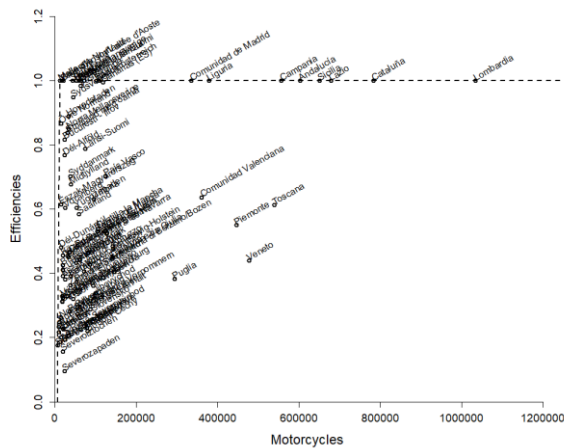


Fig. 3. DEA Frontier of the Region Based on Motorcycles.

As was discussed in the previous section for analyzing the efficiency scores of the regions' road safety performance is important incorporating this information as the dependent variable and measuring the effect that the explanatory variables have on their performance (efficiency). In this case, the Tobit regression model was considered more than appropriate. Table 1 presents the outcome of the Tobit model for each time instant.

Table 1. Results from Tobit regression models

Variables/Year	2008	2009	2010	2011	2012	2013	2014	2015	2016
Intercept	2.537e-01*** (5.05e-02)	2.75e-01*** (4.94e-02)	3.52e-01*** (4.93e-02)	2.38e-01*** (4.83e-02)	3.49e-01*** (5.13e-02)	2.39e-01*** (5.03e-02)	3.25e-01*** (4.97e-02)	1.36e-01** (5.23e-02)	2.11e-01*** (4.66e-02)
Population	-1.490e-07*** (3.65e-08)	-1.53e-07*** (3.62e-08)	-1.53e-07*** (3.67e-08)	-1.36e-07*** (3.67e-08)	-1.88e-07*** (4.49e-08)	-2.06e-07*** (3.95e-08)	-1.92e-07*** (4.59e-08)	-1.65e-07*** (4.86e-08)	-1.56e-07*** (4.57e-08)
Vehicles	-1.866e-06* (9.44e-07)	-2.078e-06* (9.23e-07)	-2.32e-06* (9.23e-07)	5.84e-07* (2.42e-07)	-	-	-	-	-1.72e-06 (8.51e-07)
Lorries	2.451e-06* (9.94e-07)	2.70e-06** (9.75e-07)	3.17e-06** (9.73e-07)	-	1.19e-06*** (2.89e-07)	1.10e-06*** (2.89e-07)	1.14e-06*** (2.87e-07)	1.01e-06*** (2.97e-07)	2.59e-06 (8.93e-07)
Motorcycles	9.095e-07* (3.87e-07)	8.340e-07* (3.69e-07)	8.41e-07* (3.59e-07)	1.19e-06*** (3.43e-07)	8.12e-07* (3.62e-07)	3.98e-07 (2.30e-07)	8.33e-07* (3.44e-07)	9.08e-07** (3.51e-07)	6.42e-07 (3.25e-07)
Passenger Cars	1.790e-06 (10.00e-07)	1.991e-06* (9.79e-07)	2.21e-06* (9.82e-07)	-8.43e-07** (2.88e-07)	-1.99e-07 (1.17e-07)	-	-1.86e-07 (1.13e-07)	-2.50e-07* (1.15e-07)	1.57e-06 (9.00e-07)
Buses	4.52e-05** (1.41e-05)	4.54e-05** (1.38e-05)	3.29e-05* (1.43e-05)	5.369e-05*** (1.418e-05)	3.96e-05** (1.50e-05)	3.977e-05*** (1.403e-05)	3.90e-05** (1.44e-05)	5.89e-05*** (1.47e-05)	5.05e-05*** (1.39e-05)
Motorway Density	4.75e-03*** (8.26e-04)	4.53e-03*** (8.09e-04)	4.22e-03*** (7.94e-04)	4.888e-03*** (7.987e-04)	4.49e-03*** (8.56e-04)	4.718e-03*** (8.365e-04)	4.375e-03*** (8.14e-04)	4.79e-03*** (8.47e-04)	4.48e-03*** (7.28e-04)
Other Roads	1.02e-05** (2.18e-06)	1.22e-05** (2.13e-06)	7.36e-06 (2.10e-06)	1.234e-05 (2.029e-06)	9.34e-06 (2.11e-06)	1.085e-05 (2.031e-06)	9.40e-06 (2.01e-06)	1.42e-05 (2.10e-06)	1.23e-05 (1.90e-06)
GDP	3.11e-06** (1.03e-06)	3.59e-06** (1.05e-06)	3.76e-06** (9.54e-07)	3.445e-06** (8.982e-07)	4.14e-06** (9.48e-07)	3.118e-06** (8.086e-07)	4.31e-06** (8.76e-07)	4.31e-06** (8.77e-07)	4.31e-06** (7.66e-07)
Log-Lik.	28.21	30.61	31.11	32.08	24.26	26.92	28.15	24.88	32.49
AIC	-34.42	-39.23	-40.22	-44.16	-28.52	-35.83	-36.29	-29.76	-42.91

Note: Parenthesis denotes the standard error of the variables

-. denotes the non-statistically variables that were omitted from the model

As it appeared, regions with a high population record more fatal accidents and therefore their efficiency in road safety performance is dropping. Additionally, the other explanatory variables that seem to have a negative relation to the regions' efficiency is the number of vehicles, except from the year 2011 where the particular variable appeared to have a positive relationship with regions' efficiency and the passengers cars which is negatively correlated with efficiency in 2011, 2012, 2014 and 2015. The other variables appeared with a positive relationship with the regions' road safety performance. Special focus should also be given to the explanatory variable "GDP" which shows that regions with high GDP have also a good efficiency fact which means that they probably spent money on improving their road safety performance. In overall, the results obtained from Tobit regression models can be used from policymakers for supporting their work on focusing to particular factors (explanatory variables) that seem to affect the regions' road safety performance in either way (positively or negatively) not only for improving the performance of under-performing regions but also to maintain the good performance of best-performing regions.

Comparing the models based on the fit index Akaike Information Criterion (AIC) it seems that the Tobit model that has the best fit is the Tobit model of 2011 which we should pay close attention to for interpretation.

6 Conclusions

The road safety performance of regions and countries play a major role, especially in the European Health and Safety agenda. For decreasing or even eliminating the number of roads fatalities are important to identify the factors that either decrease or increase this number and find a “cure” to this “illness”. However, before this step is essential to identify which regions/countries best and worst performed throughout the years.

This study has developed a straight forward methodological framework for analyzing and identifying the road safety performance of 101 EU regions (NUTS2) over a 9 years’ period taking into consideration their socio-economic, demographic and road infrastructure context.

In detail, a benchmarking analysis, namely DEA, was first implemented for ranking and identifying best-performing and under-performing regions. DEA analysis was suitable adapted to road safety’s framework. From this implementation, we are able to assist the underperforming regions by suggesting policies, laws, and enforcements that best-performing regions are following. For this reason benchmarking regions have been determined (Figure 3).

The following procedure followed in this paper was the measurement of the effect that these explanatory variables have on the regions’ road safety performance (efficiency) by developing Tobit regression models. Therefore, the outcome from Tobit models revealed the effect of the explanatory variables on the regions’ efficiency. In this way, we are able to support the address of policymakers for improving the region’s road safety performance. Special focus should be given from local authorities for improving their Public transportations (buses) because it can be seen that these factors show the highest positive relationship with the regions’ efficiency.

References

1. WHO (2018). Global status report on road safety 2018. World Health Organization, Geneva.
2. ETSC (2016), Ranking EU progress on road safety - 10th Road Safety Performance Index Report, European Transport Safety Council, Brussels.
3. EC (2019). 2018 road safety statistics: what is behind the figures. Press Release. European Commission, Brussels.
4. Eurostat (2018). Eurostat regional yearbook, 2018 edition. European Union, Luxembourg.
5. Al-Haji, G., 2007. Road Safety Development Index (RSDI): Theory, Philosophy and Practice (PhD dissertation). Department of Science and Technology, Campus Norrköping, Linköping University, Norrköping.
6. Wegman F., Commandeur J., Doveh E., Eksler V., Gitelman V., Hakkert S., Lynam D., Oppe S., 2008. SUNflowerNext: Towards a Composite Road Safety Performance Index, Deliverable D6.16 of the EU FP6 project SafetyNet.

7. Yannis G., Weijermars W., Gitelman V., Vis M., Chaziris A., Papadimitriou E., Azevedo C.L., 2013. Road safety performance indicators for the interurban road network. *Accident Analysis Prevention* 60, 384–395.
8. Shen Y., Hermans E. Bao Q., Brijs T., Wets G., Wang W. (2015). International benchmarking of road safety: State of the art. *Transportation Research Part C* 50, 37–50.
9. Shen Y, Hermans E, Bao Q, Brijs T, Wets G. Road safety development in Europe: a decade of changes (2001–2010). *Accident Analysis Prevention* 60, 85–94.
10. Nikolaou P. & Dimitriou L. (2018). Evaluation of road safety policies performance across Europe: Results from benchmark analysis for a decade. *Transportation Research Part A* 116, 232–246.
11. Alper D., Sinuany-Stern Z., Shinar D. (2015). Evaluating the efficiency of local municipalities in providing traffic safety using the Data Envelopment Analysis. *Accident Analysis and Prevention* 78, 39–50.
12. Bastos J.T., Shen Y., Hermans E., Brijs T., Wets G., Clovis Pinto Ferraz A. (2015). Traffic fatality indicators in Brazil: State diagnosis based on data envelopment analysis research. *Accident Analysis and Prevention* 81, 61–73.
13. Ganji S.R.S., Rassafi A., Xu D. (2019). A double frontier DEA cross efficiency method aggregated by evidential reasoning approach for measuring road safety performance. *Measurement* 136, 668–688.
14. Tasnim, N. and Afzal, M., 2018. An empirical investigation of country level efficiency and national systems of entrepreneurship using Data Envelopment Analysis (DEA) and the TOBIT model. *Journal of Global Entrepreneurship Research*, 8(1).
15. Wu, Wenqing & Xie, Fei & Zhao, Li-ming. (2010). A study on the operation of Chinese university science park based on DEA-Tobit model. 10.1109/ICIEEM.2010.5646090.
16. Eksler V. (2008). Exploring Spatial Structure behind the Road Mortality of Regions in Europe. *Appl. Spatial Analysis* 1, 133–150.
17. Papadimitriou E., Eksler V., Yannis G., Lassarre S. (2012). Modelling the spatial variation of road safety in Greece, *Proceedings of ICE – Transport*, 166:1, 49-58.
18. Eksler V. (2006). Road mortality in Europe: How sensitive is it to demographic structure and population dynamics? *IATSS Research* 31:1, 80-88.
19. Eksler V., Lassarre S., Thomas I. (2008). Regional analysis of road mortality in Europe. *Public Health* 122:9, 826-837.
20. Charnes, A., Cooper, W. and Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), pp.429-444.
21. Tobin, J., 1958. Estimation of Relationships for Limited Dependent Variables. *Econometrica*, 26(1), p.24.