Comparative Evaluation of Driving Efficiency Using Smartphone Data

Dimitrios I. Tselentis
Research Associate
National Technical University of Athens,
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Tel: +30 210 772 2210; Fax: +30 210 772 1454
E-mail: dtsel@central.ntua.gr

Eleni I. Vlahogianni, PhD
Assistant Professor
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Tel: +302107721369, Fax: +302107721454
E-mail: elenivl@central.ntua.gr

George Yannis, PhD
Professor
National Technical University of Athens
Department of Transportation Planning and Engineering
5 Iroon Polytechniou St., GR-15773, Athens, Greece
Tel: +302107721326, Fax: +302107721454
E-mail: geyannis@central.ntua.gr

ACKNOWLEDGEMENT
The authors would like to thank OSeven Telematics, for providing all necessary data to accomplish this study.

Word Count: 1728 (Tables: 2)

Paper No: 18-04182

Submission Date: 14/11/2017

INTRODUCTION

The objective of this paper is to provide a solid framework for the comparative evaluation of driving efficiency based on Data Envelopment Analysis (DEA). The analysis considers each driver as a Decision Making Unit (DMU) and aims to provide a relative efficiency measure to compare different drivers based on their driving performance. Driver's efficiency on a microscopic level has been studied in a great extent but never by making use of DEA techniques. This paper proposes a methodological framework to address the issue of measuring driver's efficiency and categorize the drivers of the sample used in three groups i.e. non-efficient, weakly efficient, most efficient.

Measuring driving efficiency has been the focus of many studies in driving behavior literature in the past *(1)* because from a safety perspective, it is extremely significant to identify driving risk parameters and quantify their influence on traffic risk. Several studies have been carried out regarding mobile phone usage distraction and methodologies for collecting and analyzing driving behavior data *(2)*. Literature revealed that the most significant parameters associated with driving risk are mobile phone usage *(3)*, speeding *(4)* and harsh events *(5)*. which are used herein to estimate efficiency on urban, rural and highway road and in an overall model. Driving analytics used, were collected using an innovative data collection scheme, which is based on the continuous recording of driving behavior analytics in real time, using smartphone device sensors. The proposed methodological framework is tested on data from fifty-six (56) drivers during an 8-months driving experiment.

DEA models are identifying the most efficient drivers that lie on the efficiency frontier and act as peers for the rest of the non-efficient drivers. Results provide a potential for classification of the driving sample based on drivers' comparative efficiency. The main characteristics of each driving group are consequently analyzed and presented to draw important conclusions on their features and provide recommendations for drivers on how to improve their driving efficiency. An additional value of the methodology proposed is that it provides the methodological framework to estimate of the optimal level of inputs and outputs that should be reached by each driver to become efficient. The impact of this methodology lies on the fact that most common inefficient driving practices are identified (aggressive, risky driving etc.) and driving behavior is comparatively evaluated and analyzed.

METHODOLOGY

DEA has become one of the most popular fields in operations research, with applications involving a wide range of context *(6)* in order to measure and compare the productivity performance of a group of DMUs *(7, 8)*. DEA has also been implemented in transport fields in assessing public transportation system performance *(9)*, as well as traffic safety studies *(10, 11)* where it was proved to be equally useful as in the fields stated above. The concept of DEA is to minimize inputs (input-oriented model) or maximize the outputs of a problem (output-oriented model). From a road safety perspective, increasing mileage increases crash risk *(2)* and, therefore, an input-oriented DEA program is being developed aiming to minimize inputs (recorded driving metrics) maintaining the same number of outputs (recorded distance). The code for data aggregation and DEA (based on *(12)*) models development was implemented in Python programming language. For the purposes of this study, drivers will be considered DMUs; this is deemed to be a correct assumption since a) all variables used are continuous quantitative variables as those used in previous DEA studies *(13)* and b) a driver should reduce his mileage and the frequency of some of his driving characteristics *(2)*.

A mobile App developed by OSeven Telematics is employed for the purposes of this study to record user's behaviour exploiting the hardware sensors of the smartphone device and a variety of APIs to read sensor data and transmit it to a central database. After data is sent and stored in the cloud server for central processing and data reduction, it is converted into meaningful behavioral and safety related indicators as a result of Big Data mining techniques and Machine Learning algorithms. Data collected from the naturalistic driving experiment implemented were anonymized so that the driving behaviour of each participant was not connected with any personal information. The large database consisted of 50,741 trips made from two hundred and thirty six (236) drivers that were randomly selected from OSeven database for the purpose of this research. For each driver, all trips that took place between August 2016 and April 2017 were selected. Participants chosen should have a) a driving sample of more than 100 driving hours, b) positive mileage on all three types of road network and c) positive input attributes (i.e. zero harsh acceleration, braking, speed limit violation, mobile phone usage). This procedure resulted to 56 driver and a total number of 34,060 trips.

FINDINGS

All driving indicators influencing accident risk along with distance travelled by drivers are recorded per road type and in total. Total driving behavior of each driver is considered equivalent to the sum of the driving characteristics that were recorded for the period examined. DEA procedure as described in *(12)* is followed separately for each of the three different road types and aggregately in an overall model. Variables used in the analysis are explained in table 1.

TABLE 1: Variables recorded during the experiment

Variable name	Variable short description
ha_X	number of harsh acceleration events in X road type
ha_{urban}	number of harsh acceleration events in urban road
ha_{rural}	number of harsh acceleration events in rural road
$ha_{highway}$	number of harsh acceleration events in highway
hb_X	number of harsh braking events in X road type
hb_{urban}	number of harsh braking events in urban road
hb_{rural}	number of harsh braking events in rural road
$hb_{highway}$	number of harsh braking events in highway
$speeding_X$	total seconds of speed limit violation in X road type
$speeding_{urban}$	total seconds of speed limit violation in urban road
$speeding_{rural}$	total seconds of speed limit violation in rural road
$speeding_{highway}$	total seconds of speed limit violation in highway
$mobile_X$	total seconds of mobile phone usage in X road type
$mobile_{urban}$	total seconds of mobile phone usages in urban road
$mobile_{rural}$	total seconds of mobile phone usage in rural road
$mobile_{highway}$	total seconds of mobile phone usage in highway
$distance_X$	total distance driven in X road type
$distance_{urban}$	total distance driven in urban road
$distance_{rural}$	total distance driven in rural road
$distance_{highway}$	total distance driven in highway

As shown in table 2, model 1, 2 and 3 represents speed limits violation, mobile phone distraction and driving aggressiveness respectively whereas model 4 is the overall model that includes all traffic safety parameters and accounts for the overall safety profile of the driver.

In order to classify drivers, the percentile thresholds of 25% and 75% of θ_B were used, which splits the drivers into non-efficient, weakly efficient and most efficient. The average attributes of each class, weighted on distance (harsh acceleration/braking) or driving time (speeding/mobile usage) travelled by each driver, are shown in table 2 where the models developed per type, road type and overall are presented based on the inputs that were used in each model. For better understanding, results are presented as a percentage of driving time for speeding/mobile usage and as events per 100 kilometers driven for harsh acceleration/braking.

The average metrics for models 1, 2 and 3 in every road type is reducing while drivers become more efficient. Additionally, harsh events occurring in urban road are significantly more than in rural and highway and those occurring in rural road are more than in highway. The same is noticed for mobile usage but not for speeding where drivers of all classes tend to drive over the speed limits approximately the same in urban and rural and more in highway.

In general, it can be concluded from model 1 that speed limit violation does not fluctuate and is limited to less than 6.5% of driving time for most efficient drivers in all road types whereas for non-efficient drivers it ranges from 20% to over 32%. As for the set of weakly efficient drivers speed limit exceedance is around 12% - 14%. In terms of mobile usage distraction, it appears that non-efficient drivers use their mobile phone significantly more than the other two classes averaging at 16% while most efficient drivers use it less than 1.5% in average. Weakly efficient group of drivers make mobile usage of less than 7%. It is also noticeable from model 3 that drivers of all ranges of aggressiveness have a 2-3 times larger number of harsh acceleration than braking events per 100km of driving.

As the number of inputs and outputs increases while the number of DMUs remains low, the number of efficient DMUs that are found to be efficient is radically increased. This is the case of the overall model 4, where 38 drivers with unit efficiency were found and this is why it is not considered significant enough to be presented.

After DEA LPs are solved and the efficiency index θ_B and coefficients λ_i are estimated for each DMU the efficient level of inputs for each DMU is calculated as the product sum of the lamdas and the input values of each of the identified peers. In order to find the efficient level of outputs for the same DMU, each output value should be divided by theta value. For the purpose of brevity, lamdas and thetas calculated are not presented herein.

TABLE 2: Driving characteristics of efficiency groups per road type and overall

Model		Efficiency classes		
		1: 0 - 25 % percentile	**2: 25 - 75 % percentile**	**3: 75 - 100 % percentile**
Urban	**1**	20.08 % $speeding_{urban}$	11.95 % $speeding_{urban}$	6.51 % $speeding_{urban}$
	2	19.48 % $mobile_{urban}$	6.80 % $mobile_{urban}$	2.31 % $mobile_{urban}$
	3	45.97 $ha_{urban}/100km$ 17.38 $hb_{urban}/100km$	27.40 $ha_{urban}/100km$ 8.99 $hb_{urban}/100km$	10.71 $ha_{urban}/100km$ 5.08 $hb_{urban}/100km$
	4	41.06 $ha_{urban}/100km$ 16.75 $hb_{urban}/100km$ 17.77 % $mobile_{urban}$ 15.79 % $speeding_{urban}$	22.85 $ha_{urban}/100km$ 8.43 $hb_{urban}/100km$ 6.78 % $mobile_{urban}$ 13.02 % $speeding_{urban}$	24.72 $ha_{urban}/100km$ 6.81 $hb_{urban}/100km$ 4.05 % $mobile_{urban}$ 8.66 % $speeding_{urban}$
Rural	**1**	23.79 % $speeding_{rural}$	14.21 % $speeding_{rural}$	6.33 % $speeding_{rural}$
	2	15.10 % $mobile_{rural}$	5.69 % $mobile_{rural}$	1.64 % $mobile_{rural}$
	3	23.65 $ha_{rural}/100km$ 11.43 $hb_{rural}/100km$	14.28 $ha_{rural}/100km$ 6.96 $hb_{rural}/100km$	6.36 $ha_{rural}/100km$ 3.00 $hb_{rural}/100km$
	4	20.31 $ha_{rural}/100km$ 8.71 $hb_{rural}/100km$ 10.28 % $mobile_{rural}$ 20.58 % $speeding_{rural}$	12.32 $ha_{rural}/100km$ 6.26 $hb_{rural}/100km$ 6.51 % $mobile_{rural}$ 14.49 % $speeding_{rural}$	13.62 $ha_{rural}/100km$ 7.13 $hb_{rural}/100km$ 4.81 % $mobile_{rural}$ 8.97 % $speeding_{rural}$
Highway	**1**	32.39 % $speeding_{highway}$	13.06 % $speeding_{highway}$	3.98 % $speeding_{highway}$
	2	12.34 % $mobile_{highway}$	3.73 % $mobile_{highway}$	0.74 % $mobile_{highway}$
	3	3.40 $ha_{highway}/100km$ 1.67 $hb_{highway}/100km$	1.74 $ha_{highway}/100km$ 1.02 $hb_{highway}/100km$	0.98 $ha_{highway}/100km$ 0.49 $hb_{highway}/100km$
	4	2.80 $ha_{highway}/100km$ 1.61 $hb_{highway}/100km$ 5.40 % $mobile_{highway}$ 29.31 % $speeding_{highway}$	1.91 $ha_{highway}/100km$ 1.05 $hb_{highway}/100km$ 5.61 % $mobile_{highway}$ 13.08 % $speeding_{highway}$	1.24 $ha_{highway}/100km$ 0.50 $hb_{highway}/100km$ 3.92 % $mobile_{highway}$ 7.01 % $speeding_{highway}$
Overall	**1**	17.12 % $speeding_{urban}$ 21.25 % $speeding_{rural}$ 24.24 % $speeding_{highway}$	12.50 % $speeding_{urban}$ 14.41 % $speeding_{rural}$ 14.26 % $speeding_{highway}$	8.37 % $speeding_{urban}$ 8.48 % $speeding_{rural}$ 9.72 % $speeding_{highway}$
	2	17.07 % $mobile_{urban}$ 13.30 % $mobile_{rural}$ 9.75 % $mobile_{highway}$	7.22 % $mobile_{urban}$ 5.99 % $mobile_{rural}$ 4.37 % $mobile_{highway}$	3.89 % $mobile_{urban}$ 2.85 % $mobile_{rural}$ 2.05 % $mobile_{highway}$
	3	36.94 $ha_{urban}/100km$ 19.26 $ha_{rural}/100km$ 3.12 $ha_{highway}/100km$ 12.42 $hb_{urban}/100km$ 9.33 $hb_{rural}/100km$ 1.44 $hb_{highway}/100km$	30.09 $ha_{urban}/100km$ 16.26 $ha_{rural}/100km$ 1.76 $ha_{highway}/100km$ 10.34 $hb_{urban}/100km$ 7.36 $hb_{rural}/100km$ 0.95 $hb_{highway}/100km$	17.13 $ha_{urban}/100km$ 8.46 $ha_{rural}/100km$ 1.32 $ha_{highway}/100km$ 7.87 $hb_{urban}/100km$ 4.85 $hb_{rural}/100km$ 0.87 $hb_{highway}/100km$
	4	-	-	-

CONCLUSION

This paper provides an innovative solid framework for benchmarking driving efficiency based on Data Envelopment Analysis (DEA). Efficiency is examined in terms of speed limit violation, driving distraction, aggressiveness and safety on urban, rural and highway road and in an overall model. Drivers were divided into three categories and results indicated considerable differences in driving characteristics between inefficient, weakly efficient and most efficient drivers with the difference of the two latter to be less significant. Concerning aggressiveness, harsh braking events appeared to be 2-3 times less than harsh acceleration events in all models indicating a higher significance of this attribute for a driver to be characterized as aggressive. The same is observed for harsh acceleration events in overall safety models of all road types where a) the number of harsh events does not considerably influence the overall driving profile of a less risky driver (it is approximately the same especially for efficiency classes 2 and 3) and b) percentage of speeding and mobile usage was identified as key factors for safety efficiency index estimation. Weakly efficient drivers in urban and rural road have a lower average number of harsh acceleration events and in average, the same driving characteristics for the rest of the attributes investigated. In terms of aggressiveness and safety, drivers of the same class were found to have similar values of harsh events. Finally, for non-efficient drivers of urban and rural road, it was found that all driving attributes were reduced compared to model 1, 2 and 3 probably due to the interaction among variables.

The results of the present research could be exploited by a smartphone app to provide feedback to drivers on their driving characteristics that need further improvement in order to improve their overall and per road type driving efficiency. Results could also be used for insurance pricing based on driving usage and characteristics.

Future research should center to larger samples of drivers and trips. Other limitations should be addressed including DEA's sensitivity to outliers and that drivers with zero input attributes should be eliminated from the sample. Finally, results of per trip and per driver analysis of each driver should be compared to derive more information.

REFERENCES

1. Young, Mark S., Stewart A. Birrell, and Neville A. Stanton. Safe driving in a green world: A review of driver performance benchmarks and technologies to support 'smart'driving. Applied ergonomics Vol. 42, No. 4, 2011, pp. 533-539. https://doi.org/10.1016/j.apergo.2010.08.012
2. Tselentis, Dimitrios I., George Yannis, and Eleni I. Vlahogianni. Innovative motor insurance schemes: a review of current practices and emerging challenges. Accident Analysis & Prevention Vol. 98, 2017, pp. 139-148. https://doi.org/10.1016/j.aap.2016.10.006
3. Haque, Md Mazharul, and Simon Washington. The impact of mobile phone distraction on the braking behaviour of young drivers: a hazard-based duration model. Transportation research part C: emerging technologies Vol. 50, 2015, pp. 13-27. https://doi.org/10.1016/j.trc.2014.07.011
4. Andersson, Gunnar, and Göran Nilsson. Speed management in Sweden: Speed, speed limits and safety. Swedish National Road and Transport Research Institute, 1997.
5. Bonsall, Peter, Ronghui Liu, and William Young. Modelling safety-related driving behaviour—impact of parameter values. Transportation Research Part A: Policy and Practice Vol. 39, No. 5, 2005, pp. 425-444. https://doi.org/10.1016/j.tra.2005.02.002
6. Thanassoulis, Emmanuel. Introduction to the theory and application of data envelopment analysis. Dordrecht: Kluwer Academic Publishers, 2001.

7. Cook, W. D., & Seiford, L. M. Data envelopment analysis (DEA)–Thirty years on. European journal of operational research, Vol. 192, No. 1, 2009, pp. 1-17. https://doi.org/10.1016/j.ejor.2008.01.032

8. Emrouznejad, A., Parker, B. R., & Tavares, G. Evaluation of research in efficiency and productivity: A survey and analysis of the first 30 years of scholarly literature in DEA. Socio-economic planning sciences, Vol. 42, No 3, 2008, pp. 151-157. https://doi.org/10.1016/j.seps.2007.07.002

9. Karlaftis, M.G., Gleason, J.M., Barnum, D.T. Bibliography of Urban Transit Data Envelopment Analysis (DEA) Publications (2013). Available at SSRN: http://ssrn.com/abstract=1350583, http://dx.doi.org/10.2139/ssrn.1350583

10. Egilmez, G., & McAvoy, D. Benchmarking road safety of US states: A DEA-based Malmquist productivity index approach. Accident Analysis & Prevention, Vol. 53, 2013, pp. 55-64. https://doi.org/10.1016/j.aap.2012.12.038

11. Alper, D., Sinuany-Stern, Z., & Shinar, D. Evaluating the efficiency of local municipalities in providing traffic safety using the Data Envelopment Analysis. Accident Analysis & Prevention, Vol. 78, 2015, pp. 39-50. https://doi.org/10.1016/j.aap.2015.02.014

12. Charnes, A., Cooper, W. W., & Rhodes, E. Measuring the efficiency of decision making units. European journal of operational research, Vol. 2, No. 6, 1978, pp. 429-444. https://doi.org/10.1016/0377-2217(78)90138-8

13. Hollingsworth, B., Dawson, P. J., & Maniadakis, N. Efficiency measurement of health care: a review of non-parametric methods and applications. Health care management science, Vol 2, No. 3, 1999, pp. 161-172. https://doi.org/10.1023/A:1019087828488