ROAD TRAFFIC FATALITIES ON A GLOBAL SCALE: CHALLENGES AT THE MACRO-LEVEL INFORMATION

Paraskevas Nikolaou*[a], Loukas Dimitriou[a], Ph.D., Constantinos Antoniou[b], Ph.D.

[a] Lab. for Transport Engineering, Department of Civil and Environmental Engineering, University of Cyprus75 Kallipoleos Str., P.O. Box 20537, 1678 Nicosia, Cyprus. Email: nikolaou.paraskevas@ucy.ac.cy

[b] Professor, Chair of Transportation Systems Engineering, Technical University of Munich, Department of Civil, Geo and Environmental Engineering, Arcisstrasse 21, D-80333, Munich, Germany. Email: c.antoniou@tum.de

(* Corresponding Author)

Extended Abstract:

Road traffic fatalities appear to have a correlation with the countries' societal development that, in turn, is affected by economic and demographic factors (e.g. GDP, Income per capita, population, number of registered vehicles etc.) (Dimitriou et al., 2017; Antoniou et al., 2016). However, the analysis of macro-level information is a constituent sine qua non for the implementation of a technically sound methodological framework. The results of this analysis may support national and regional policymakers towards the improvement of road safety policies.

The objective of this research is to provide evidence on causality, utilizing evidence from global statistics. The present paper make use of an extensive dataset of national statistics from 105 member countries of the United Nations (UN-105) for the year 2016 in order to quantitatively expose possible, statistically significant, explanatory relationships. In detail, 12 socio-economic and demographic variables were collected and used for fulfilling the current study's scope. However, despite the direct relationships between the socio-economic and demographic variables with road traffic fatalities it is also considered that within the dataset there also exists a latent part of the information that has a relationship with road traffic fatalities. This concern extends previous research of Dimitriou et al., (2017), where they identified these latent elements and their relationship with road traffic fatalities. In order to identify these latent constructs and possible data inflations (e.g. collinearity) a suitable statistical method, namely Factor Analysis, was implemented. Factor Analysis was able to identify the existence of 2 possible latent constructs inside the dataset. Both factors have a structure of three variables, i.e., Factor 1: "Population", "Total Registered Vehicles" and "GDP" and; Factor 2: "Birth Rate", "Individuals Using the Internet" and "Income Group". As a result, the factors were named as "Demography" and "Socio-Economy", respectively.

In order to incorporate the effect of this latent information (which has emerged from the Factor Analysis) for explaining road safety levels and also for validating these effects, two suitable methodological approaches of Structural Equation Modeling were implemented, namely Covariance-Based Structural Equation Modeling (CB-SEM) and Partial Least Squares Structural Equation Modeling (PLS-SEM). Both SEMs where used in a confirmatory manner for validating the direct and indirect relationships of the observed and unobserved variables with road traffic fatalities.

Notwithstanding, the fact that the resulted models appeared to provide adequate information on the factors that have a strong relationship with road traffic fatalities, the analysis of macro-level information for global scale phenomena requires a further investigation due to the challenges of heterogeneity within this information. Therefore, the need for studying the phenomenon of road traffic fatalities in homogeneous sets of countries led to the implementation of a comprehensive analysis of clusters. In

detail, two types of hierarchical cluster analysis were developed, namely, agglomerative clustering and divisive clustering. However, for the development of the hierarchical cluster analysis, it was needed to use the 'appropriate' variables. In order to do that we tested the correlation of the variables with road traffic fatalities through correlation analysis. The correlation analysis showed that the variables "Population", "Total Number of Registered Vehicles", "Land Area", "Road Network Length" and "GDP" have the most significant relationship with road traffic fatalities. In the agglomerative hierarchical clustering four methods were tested (i.e. complete, average, single and Ward) and the method with values of agglomerative coefficient closer to one was suggested as the strongest clustering structure. Ward's method identified as the strongest clustering structure of the four methods assessed. The divisive coefficient of the divisive hierarchical clustering was also tested and found to be less than the agglomerative coefficients. Therefore, the agglomerative hierarchical cluster analysis (Ward's method) was implemented. However, at this point, the question raised was the identification of the optimal number of clusters. In order to answer this question, we developed Elbow method, which identified five clusters as the optimum number of clusters from the agglomerative hierarchical clustering. Cluster 1 and 2 included 86 and 14 countries, respectively. Cluster 3, included three countries and Cluster 4 and 5 included only one country each. Thus, due to the insignificant number of countries the Clusters, three to five were excluded from the following implementations of this study and only Clusters one and two were included.

For analyzing the phenomenon of road traffic fatalities in these two sets of countries the entire methodological framework was implemented starting from the Factor Analysis and ending with GoF structural equation models (CB-SEM and PLS-SEM). The resulted Structural Equation Models of both homogeneous sets of countries were compared with the initial models for identifying similarities and dissimilarities of the effects that this macro-level information may have on road traffic fatalities.

The implementation of the proposed methodological approaches provided useful findings on the challenges that arise from the analysis of macro-level information for the investigation of global phenomena, likewise road traffic fatalities. Finally, the analysis of homogeneous sets of countries will open different directions for analyzing road traffic fatalities also in a national-level with a particular focus on socio-economic and demographic conditions.

Keywords: Global road traffic fatalities, Covariance - Based Structural Equation Modeling, Partial Least Squares - Structural Equation Modeling, Cluster Analysis, Socio-Economic and Demographic Macro Information.

References

C. Antoniou, Yannis G., Papadimitriou E. and Lassarre S. Relating traffic fatalities to GDP in Europe on the long term.

L. Dimitriou, Nikolaou P. and Antoniou C. Policy-Driven Investigation of Sectoral Latent Information Regarding Global Road Fatalities. Transportation Research Procedia, Vol. 22, 2017, pp. 685-694.

World Health Organization. Global Status Report on Road Safety 2018. https://www.who.int/violence_injury_prevention/road_safety_status/2018/en/. Accessed July 10, 2018.