Modeling traffic fatalities in Europe

George Yannis , Constantinos Antoniou and Eleonora Papadimitriou
National Technical University of Athens

ABSTRACT

The objective of this paper is to provide a parsimonious model for linking motorization level with the decreasing fatality rates observed across EU countries during the last three decades. Earlier models used to describe the - at the time - increasing relationship between motorization and traffic fatalities were adjusted in order to describe the decreasing relationship observed in the last three decades. A macroscopic analysis of road-safety in Europe at the country level (16 EU countries) is proposed through the application of non-linear models correlating fatalities and vehicles for the period between 1970 and 2002. Given the time series nature of road safety data, these models result in autocorrelated residuals, thus violating at least one of the assumptions of non-linear regression. Autoregressive forms of the considered models that overcome these limitations and provide superior predictive capabilities are also considered. An autoregressive log-transformed model seems to outperform the base autoregressive non-linear model in this respect. The use of these models allowed for the identification of the best and worst performing countries in terms of traffic fatalities. The proposed models may be proved useful for assessing the road safety performance of the examined countries, as well as for obtaining some insight on the current and future trends of less developed countries.

MODEL DEVELOPMENT

Aggregate fatality, population and vehicle data from European countries between 1970 and 2002 have been used. The data have been obtained primarily from IRTAD (International Road Traffic and Accident Database). Official representatives of the countries with missing data were contacted directly, and several responses with additional data were incorporated to the database. In the end, out of the 25 countries of the enlarged EU, sufficiently complete data have been available for 16 of them, for which this model has been applied. Fatalities data refer to the 30-day definition of fatality for all countries, i.e. include all persons who died within 30 days of being involved in a traffic accident. Naturally, data collected over a long period of time, across different countries and with different data collection techniques are expected to not be uniform. This macroscopic analysis does not explicitly consider these differences, as well as the differences in terms of different road network, vehicle fleet characteristics and driver behavior in the various countries considered.

CONCLUSIONS

Using fatality rate and vehicle ownership data from 16 EU countries for a period of 33 years (1970-2002) several models were developed, fitted, validated and compared, including simple non-linear models, their log-transformations and the related autoregressive models. The autoregressive versions of the models were proved to overcome the correlation of the residuals and also exhibit superior predictive properties. For a couple of countries (Italy and the Netherlands), however, the autoregressive model performed poorer than the base non-linear model. Log-transformed versions of the model also suffer from correlated residuals, and with the exception of few cases (especially Finland, Greece, and Hungary) have better or similar predictive capabilities than the non-linear models. The autoregressive log-transformed models also overcome the issues with the correlated residuals and provide superior predictive performance.

However, the estimated coefficients of the AR log-transformed model for five of the 16 countries are sometimes questionable (in terms of magnitudes and signs), suggesting that this model should be applied with caution, taking into account the particularities of the case examined. The autoregressive non-linear models therefore seem to be a more robust choice for prediction of macroscopic road safety trends, as they provide desirable predictive properties, satisfy the assumptions of the model (e.g. uncorrelated residuals) and provide intuitive model parameters (in terms of magnitude and sign). Of course, it is recognized that these models provide average rates of progress over the period of analysis and not necessarily the "current" rate of progress (let alone future trends).

The results of the presented models can be used to evaluate the road safety performance of various countries, identifying poor performers, as well as traffic safety leaders. Indeed, as exhibited in the previous section, the model accurately determines the poor performers among the considered countries (Greece, Portugal, Cyprus), as well as those countries that are leading in terms of their road safety performance (United Kingdom, Finland, Netherlands, Denmark). At individual country level, given estimates of a country's expected performance, the actual road safety performance of that country over the past few years may be assessed. Moreover, by applying the models, the expected road safety situation in a country in a "do-nothing" scenario is described, so that the potential impact of adopted road safety strategies may be assessed at macroscopic level (e.g. target setting). Furthermore, the study of more advanced (in terms of road safety and in general) countries may be applied to predict the future evolution of less developed or successful (in terms of traffic safety) countries. However, it is stressed that the use of the developed models for prediction should be limited within the currently applied domain, as their applicability in ranges for which data is not available cannot be verified.

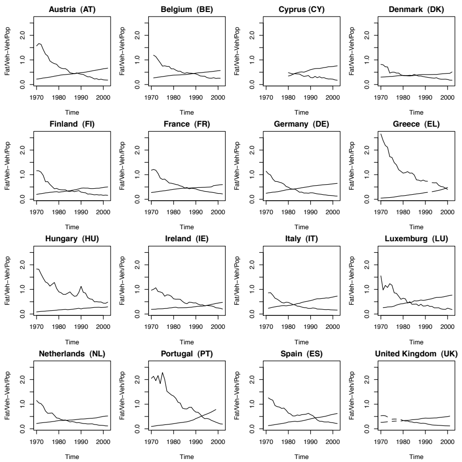

FIGURE 1. Presentation of the data set: fatalities per vehicle (decreasing trend) and vehicles per population (increasing trend)

TABLE 1. Non-linear model estimation results (top: base, bottom: after correcting for correlation)

	Coefficient α Estimate	Coefficient α Standard error	Coefficient α t-test	Coefficient β Estimate	Coefficient β Standard error	Coefficient β t-test
AT	0,099	0,007	14,962	-1,962	0,054	-36,252
BE	0,080	0,006	13,215	-2,068	0,069	-30,091
CY	0,219	0,018	12,262	-0,770	0,108	-7,158
DK	0,012	0,004	3,204	-3,477	0,291	-11,958
FI	0,026	0,006	4,597	-2,475	0,162	-15,263
FR	0,083	0,006	13,151	-2,153	0,073	-29,698
DE	0,070	0,006	12,469	-2,012	0,070	-28,597
EL	0,288	0,016	18,252	-0,711	0,023	-31,058
HU	0,172	0,028	6,260	-0,984	0,082	-11,987
IE	0,035	0,008	4,540	-2,075	0,151	-13,762
IT	0,081	0,006	14,078	-1,677	0,060	-27,834
LU	0,156	0,018	8,626	-1,542	0,104	-14,815
NL	0,017	0,002	8,384	-2,844	0,091	-31,123
PT	0,290	0,039	7,398	-0,956	0,075	-12,753
ES	0,212	0,017	12,716	-0,876	0,049	-17,784
UK	0,030	0,003	11,403	-2,210	0,076	-28,933

	Coefficient α Estimate	Coefficient α Standard error	Coefficient α t-test	Coefficient β Estimate	Coefficient β Standard error	Coefficient β t-test	Coefficient φ Estimate	Coefficient φ Standard error	Coefficient φ t-test
AT	0,090	0,010	9,303	-2,051	0,096	-21,484	0,3387	0,1255	2,699
BE	0,077	0,012	6,215	-2,111	0,158	-13,396	0,4487	0,197	2,277
CY	0,214	0,027	7,994	-0,815	0,180	-4,524	0,2047	0,2905	0,705
DK	0,015	0,010	1,550	-3,227	0,617	-5,228	0,5686	0,1695	3,355
FI	0,021	0,009	2,209	-2,687	0,370	-7,273	0,4647	0,1429	3,252
FR	0,068	0,016	4,329	-2,382	0,251	-9,494	0,5339	0,1798	2,970
DE	0,069	0,011	6,215	-2,034	0,153	-13,329	0,5282	0,1752	3,015
EL	0,294	0,025	11,740	-0,701	0,037	-19,013	0,3005	0,2131	1,410
HU	0,155	0,062	2,516	-1,045	0,218	-4,794	0,5825	0,1784	3,265
IE	0,034	0,015	2,295	-2,119	0,309	-6,865	0,6081	0,152	4,001
IT	0,071	0,009	8,243	-1,818	0,116	-15,687	0,3571	0,1819	1,964
LU	0,131	0,015	8,521	-1,757	0,114	-15,438	0,2458	0,1454	1,691
NL	0,015	0,003	4,986	-2,969	0,163	-18,200	0,3247	0,1364	2,380
PT	0,219	0,068	3,230	-1,154	0,196	-5,893	0,5303	0,1314	4,037
ES	0,135	0,054	2,473	-1,306	0,418	-3,122	0,7992	0,0688	11,619
UK	0,025	0,006	4,244	-2,374	0,230	-10,333	0,5916	0,2184	2,709

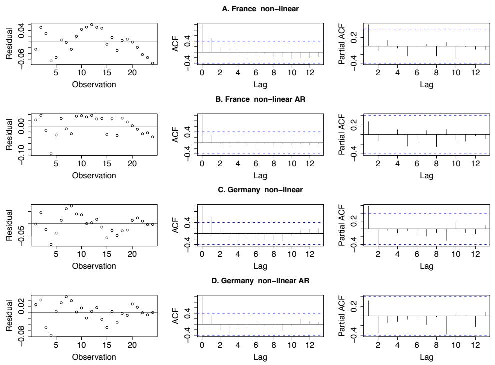

FIGURE 2. Model diagnostics for non-linear and AR non-linear models (France, Germany)

TABLE 2. Log-transformed model estimation results (top: base, bottom: autoregressive model)

	Coefficient α Estimate	Coefficient α Standard error	Coefficient α t-test	Coefficient β Estimate	Coefficient β Standard error	Coefficient β t-test
AT	-2,395	0,057	-42,122	-2,031	0,057	-35,857
BE	-2,521	0,074	-33,904	-2,056	0,077	-26,896
CY	-1,555	0,083	-18,642	-0,818	0,120	-6,808
DK	-4,004	0,278	-14,423	-3,047	0,273	-11,160
FI	-2,985	0,188	-15,884	-1,916	0,166	-11,528
FR	-2,565	0,091	-28,050	-2,236	0,100	-22,443
DE	-2,715	0,068	-40,134	-2,056	0,073	-28,224
EL	-1,210	0,048	-25,150	-0,694	0,024	-28,465
HU	-1,647	0,172	-9,569	-0,919	0,096	-9,567
IE	-3,353	0,224	-14,995	-2,073	0,164	-12,682
IT	-2,395	0,051	-46,635	-1,558	0,054	-29,118
LU	-1,994	0,071	-27,923	-1,671	0,082	-20,433
NL	-4,187	0,088	-47,684	-2,928	0,078	-37,472
PT	-1,340	0,078	-17,150	-1,012	0,053	-19,184
ES	-1,558	0,094	-16,500	-0,878	0,070	-12,540
UK	-3,566	0,117	-30,529	-2,248	0,112	-20,050

	Coefficient α Estimate	Coefficient α Standard error	Coefficient α t-test	Coefficient β Estimate	Coefficient β Standard error	Coefficient β t-test	Coefficient φ Estimate	Coefficient φ Standard error	Coefficient φ t-test
AT	-2,452	0,097	-25,263	-2,100	0,103	-20,420	0,451	0,155	2,910
BE	-2,509	0,159	-15,772	-2,045	0,173	-11,850	0,539	0,188	2,862
CY	-1,581	0,109	-14,456	-0,865	0,167	-5,191	0,079	0,300	0,263
DK	-3,526	0,662	-5,324	-2,540	0,675	-3,765	0,688	0,150	4,594
FI	-1,871	1,036	-1,805	0,149	0,713	0,209	0,940	0,040	23,539
FR	-2,954	0,571	-5,177	-2,697	0,711	-3,795	0,748	0,185	4,035
DE	-2,567	1,269	-2,022	0,738	1,270	0,581	0,961	0,020	49,141
EL	-1,184	0,087	-13,568	-0,680	0,045	-14,979	0,431	0,205	2,097
HU	-1,977	0,508	-3,896	-1,110	0,304	-3,653	0,709	0,165	4,290
IE	-2,817	5,391	-0,522	-0,523	0,582	-0,898	0,980	0,066	14,977
IT	-2,360	0,126	-18,768	-1,521	0,149	-10,193	0,686	0,165	4,165
LU	-2,053	0,063	-32,483	-1,760	0,075	-23,359	-0,033	0,184	-0,178
NL	-4,219	0,134	-31,500	-2,963	0,122	-24,254	0,358	0,177	2,030
PT	-1,437	0,145	-9,879	-1,094	0,109	-10,010	0,548	0,158	3,472
ES	-1,948	0,687	-2,837	-1,216	0,760	-1,601	0,849	0,191	4,451
UK	0,123	2,283	0,054	0,044	0,766	0,057	1,040	0,040	26,168

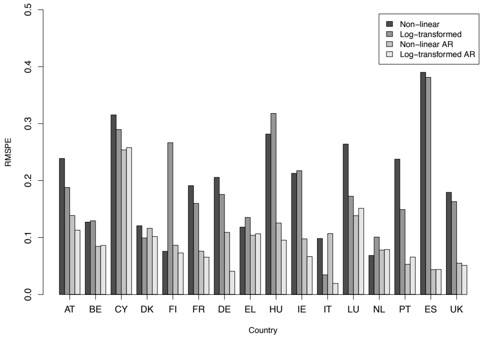

FIGURE 3. Summary goodness-of-fit statistic (RMSPE)

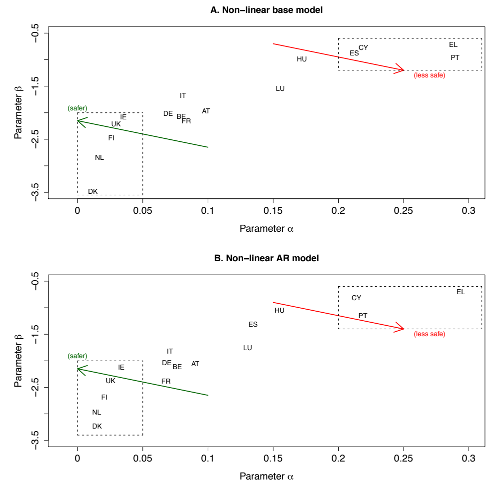

FIGURE 4. Interpretation of parameters (non-linear AR model)

ACKNOWLEDGMENT

This paper is based on the work carried out in the SAFETYNET research project, partly financed by the European Commission. The authors would like to thank Dr. Ruth Bergel from INRETS for useful suggestions in earlier drafts of this work.