



MODELLING ROAD TRAFFIC SAFETY INDICES BY MEANS OF REGRESSION WITH PANEL DATA

KATARZYNA BRZOWSKA-RUP 
 MARZENA NOWAKOWSKA 

ABSTRACT

Although the occurrence of road accidents and the number of road accident casualties in almost all Polish voivodeships has decreased over the last few years, the rate of this change varies considerably from region to region. To provide a better understanding of such a tendency, panel data regression models are proposed to conduct this pilot research which evaluates the relative performance of Polish regions in terms of their road traffic safety. Panel data are multi-dimensional data which involve measurements over time. In the research, a voivodeship is a unit analysed at a group level, whereas a year is a unit analysed at a time level. A two-way error component regression model has been applied to survey the impact of regressors, the group effects, and time effects on a dependent variable. The analysis has been conducted using data acquired from the Statistics Poland Local Data Bank website, as well as from the General Directorate for National Roads and Motorways. The panel data from 16 regions in Poland and the 2012–2018 period have been investigated. The examined models refer to road traffic safety indices defined based on the following characteristics: the number of road accidents, the number road fatalities, and the number of people injured. The results of all the three models indicate a negative effect as regards the GDP per capita, (car) motorisation rate, the indicator of government expenditure for current maintenance of national roads, and the road length per capita. A positive association has been found between the truck motorisation rate and the indicator of local government expenditure on roads. The impact of the region's urbanisation indicators on road safety is ambiguous as, on the one hand, its increase causes a reduction in the road accident and accident injury indices, but, on the other hand, it produces a rise in the accident fatality index. In the models, the significance of time effects has been identified; a decreasing time trend suggests a general improvement in road safety from year to year. Most of the group effects have turned out to be highly significant. However, the effects differ as regards both the road accident and the accident injury indices in magnitude and direction.

Marzena Nowakowska

Kielce University of Technology, Poland
 Corresponding author:
 e-mail: spimn@tu.kielce.pl
 ORCID 0000-0002-6934-523X

Katarzyna Brzowska-Rup

Kielce University of Technology, Poland
 e-mail: krup@tu.kielce.pl
 ORCID 0000-0003-1231-8027

KEY WORDS

longitudinal data, road accidents, road accident casualties, fixed effects models

10.2478/emj-2020-0026

INTRODUCTION

Road accidents and their consequences are a result of failures in the road transportation system. Their output is represented in the form of registered data describing road traffic crashes and their casual-

ties. They are the basis for exploratory data analyses focused on road traffic safety diagnosis and improvement, carried out on a range of granularity levels. The research of aggregated data is somewhat frequently undertaken and commonly concentrated either on

Brzowska-Rup, K., & Nowakowska, M. (2020). Modelling road traffic safety indices by means of regression with panel data. *Engineering Management in Production and Services*, 12(4), 40-51. doi: 10.2478/emj-2020-0026

time series analyses (Dupont et al., 2014) or cross-sectional analyses (Hauer, 2010; Söderlund & Zwi, 1995). The latter is commonly applied to the aggregated data, which have a panel structure (Wachnicka et al., 2018). Panel data, however, provide more information, more variability, less collinearity among the variables, and more degrees of freedom (Baltagi, 2001). Therefore, employing models that consider such a data structure admits an exploration of more issues than cross-sectional or time-series data alone (Kennedy, 2018).

Panel data are also called longitudinal data or cross-sectional time-series data. They have observations on the same organisational units across time. As regards road traffic safety analyses, a year is a period commonly chosen, but the possibilities of cross-section units are extensive, starting, for example, from a set of selected road network sections (small organisational units), then cities to regions, countries (big organisational units) and the like.

In Poland, the occurrence of road accidents and road accident casualties in almost all 16 voivodeships has decreased over the past few years, which is illustrated in Fig. 1. The figure presents the time series of the total number of road traffic accidents and the total number of accident fatalities. Each voivodeship is given an identifier as follows (CS stands for Cross Section): CS1 — Dolnośląskie, CS2 — Kujawsko-Pomorskie, CS3 — Lubelskie, CS4 — Lubuskie, CS5 — Łódzkie, CS6 — Małopolskie, CS7 — Opolskie, CS8 — Podkarpackie, CS9 — Podlaskie, CS10 — Pomorskie, CS11 — Śląskie, CS12 — Świętokrzyskie, CS13 — Warmińsko-Mazurskie, CS14 — Wielkopolskie, CS15 — Zachodniopomorskie, CS16 — Mazowieckie. The last identifier has intentionally been attributed to the largest voivodeship, both in terms of area and population, thus making it the reference category at the further modelling stage.

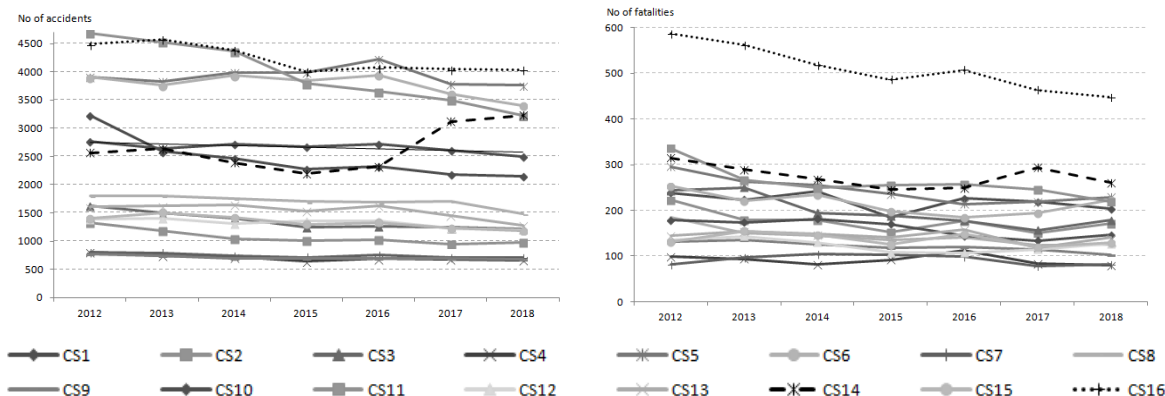


Fig. 1. The dynamics of road traffic accidents (left) and the accident fatalities (right) per voivodeship in Poland in 2012–2018

Source: authors' own elaboration

Like in other countries, in Poland, numbers of road traffic accidents, accident fatalities and injuries are the main fundamentals of road safety measures at a macro-level analysis. According to Fig. 1, a general trend of these measures is descending for the whole country. However, the patterns of change vary considerably from region to region. Certain disturbances can be observed for some regions, such as the Wielkopolskie Voivodeship (CS14) where a rapid (by 34.8%) growth of the number of accidents occurred in 2017. There is also a discrepancy between the road accident fatalities for the Mazowieckie Voivodeship (CS16) and the remaining regions (its large population and wealth can partly account for this). Thus, considering the character of the aggregated information delivered for each region within a certain period, it is justified to estimate a panel data model. These types of models enable the description of the influence of time and unit variations as well as exogenous factors on the endogenous variables, i.e., measures of road safety.

The objective of the research is to address the following issues:

- possible existing differences in the levels of road safety among individual regions of Poland;
- possible changes in safety levels with time and the nature of these changes;
- presumable relationship between selected factors, particularly road expenditure, characterising Polish regions, and the level of road traffic safety.

The contribution of the work is that:

- endogenous variables are the measures of road safety calculated in relation to road length and not to population because road length is a more stable characteristic for a voivodeship than the number of inhabitants;

- unlike in other research, motorisation rates by motor vehicle type are analysed, considering cars, trucks, and motorcycles;
- expenditure on roads is investigated, in particular including structural expenditure on national roads.

The study is divided into several sections. Following the Introduction, there is a literature review of how panel data models are applied in the road traffic safety investigation. Then, the empirical model is explained. The next part of the work presents the data to be analysed and a set of potential explanatory variables along with explained variables, which are certain measures indicating the level of road traffic safety. Then, the results and the corresponding discussion are provided regarding full and backward selection regression models. Finally, the main conclusions are presented.

1. LITERATURE REVIEW

Extensive research has been carried out in the field of road accidents, and it has been done in different ways. The synthesis of various approaches developed in the road safety analyses, their principles, and the characteristics of developed estimation techniques are discussed for example by Laaraj and Jawab (2018), Muthusamy et al. (2015), Antoniou et al. (2016), Badura (2017) and other authors indicated in the list of references of this article. This study reviewed the selected literature covering different issues related to road safety analysis with the application of panel data models.

Numerous studies have been devoted to analysing road traffic safety based on aggregate data (Besharati, 2020). However, there are limited sources available, in which researchers make use of the attractiveness of the panel data structure. Yet, in this respect, a variety of possibilities are considered regarding the explained variables that characterise a cross-section unit in a time unit, such as the Gross Domestic Product (GDP) per capita, population density, health care issues, motorisation rate, and road infrastructure. The number of road accidents or the number of road accident casualties (sometimes in relation to population) are road safety measures estimated as output variables. Different types of panel data models are employed, with the number of cross-section (organisation) units N and the number of time units T adapted to the needs of research.

Fixed effect panel data models with annual dummies (to capture the common trend in all the provinces) were estimated independently for urban roads in total within a province and urban roads within the capitals of the provinces in the study by Castro-Nuño et al. (2018). A total number of urban road traffic accidents and a number of urban road traffic accident fatalities per accident were modelled using the $N = 50$ Spanish provinces and $T = 11$ -year time units (2003–2013) data structure. The general conclusion was that in the case of cities in Spanish provinces, a wider urban spread inevitably led to more severe traffic accidents, whereas road fatality was lower in urban areas with denser populations (thus, more densely concentrated).

Annual death counts and road accident injuries were investigated with the use of one-way fixed-effect models, with time trend considered, for the road network as well as for the interurban part of the network in Spain (Albalate, 2013). In this respect, the panel with $N = 50$ Spanish provinces and $T = 21$ -year time units (1990–2010) was employed. The impact on road safety was evaluated in relation to selected characteristics of Spanish provinces and, in particular, to recent road infrastructure spending together with the main regulatory changes introduced. The results demonstrate that both regulations and road infrastructure spending influence road safety. In particular, road maintenance expenses produce a significant safety benefit in terms of reducing road accident fatalities and injuries.

The number of road traffic accidents and the number of road traffic accident casualties (injuries and fatalities) were the output variables in two-way fixed effect panel data models estimated for $N = 846$ road control stations of the Spanish highway network and $T = 5$ -year time units (2008–2012) to explore the role of Public Private Partnerships (PPP) along with road infrastructure and demographic characteristics (Albalate et al., 2019). The quality of road design was indicated as the most relevant aspect influencing road safety outcomes. Nevertheless, evidence was found suggesting that privately operated highways were positively correlated with better road safety outcomes for roads of similar quality.

The number of fatalities per 100 000 population (as a measure of safety) was estimated in two-way fixed effect panel data models for $N = 30$ provinces in Iran and $T = 11$ -year time units (2005–2015) (Besharati et al., 2020). The results revealed that the fatality rates were positively associated with certain exposure proxies, but negatively related to the varia-

bles representing the level of urbanisation. The increase in the number of speed cameras turned out to be connected with the reduction of fatality rates. The differences between the Iranian provinces as well as the time decreasing trends were identified with respect to the discussed road safety output.

In the study where the cross-sectional units had a greater level of granularity, a variety of panel data models were investigated after taking a natural logarithm of all variables in the model (Yaseen, 2018); the causality of road traffic fatalities per million people was investigated in the panel of 30 OECD countries and $T = 21$ -time units (1995–2015). The long-run regression results indicated a significant role of health expenditure, trade openness, and research and development engagement in the reduction of road traffic fatalities.

Kweon and Kockelman (2005) analysed how speed limit changes on high-speed roadways affected total safety. They defined the panel consisting of $T = 4$ -years and over 63 000 homogenous road segments. Interestingly, results indicated that speed limit changes did not influence fatal crash rates. Fatal and non-fatal crash rates decreased for road design elements, such as wider shoulders and more gradual curves. However, when traffic levels rose, non-fatal rates remained constant but fatal rates decreased.

Annual data for 1997–2013 of 51 US states were analysed by Ahangari, Atkinson-Palombo, and Garrick (2017). The results of their research showed that vehicle miles travelled, vehicles per capita, and infant mortality rates (as a proxy of health care quality) have the strongest positive impact on traffic fatality rates. The authors also found that the states with a higher urban density and more walking were associated with lower traffic fatality rates. Some suggestions were made regarding the use of multimodal transportation for the reduction of fatality rates.

2. RESEARCH METHODS

Panel data combine the characteristics of cross-sectional data and time series; they contain information about N objects (groups, units, spatial elements) registered in T time units for each object. The data set defined in this way consists of $N \cdot T$ observations. Panel data prove better at identifying and measuring effects that are simply not detectable in pure cross-section or pure time-series data (Baltagi, 2005). Panel regression models are derived from the multiple linear regression model of the form:

$$y_{it} = \alpha + \sum_{k=1}^K \beta_k x_{kit} + \varepsilon_{it} \quad (1)$$

and they have the generalised form, in which the above model (1) is a two-way model, in which the random component ε_{it} is decomposed into three components: unit (μ_i) and time (θ_t) effects account for both unit-specific (but time-invariant) and time-specific (but unit-invariant) and the purely random component v_{it} :

$$y_{it} = \alpha + \sum_{k=1}^K \beta_k x_{kit} + \mu_i + \theta_t + v_{it} \quad (2)$$

The meaning of the symbols in equation (2) is as follows:

- i, t indices denoting the object (subject, group, unit, section element) $i = 1, \dots, N$, and time period $t = 1, \dots, T$, respectively;
- α , structural parameters (constant coefficients)
- β as in the classic multiple linear regression model, the β vector (vector of slopes) determines the effect of the exogenous variables X_k on the endogenous variable Y ;
- x_{kit} the k -th explanatory variable;
- K the number of exogenous variables;
- μ_i individual effect resulting from the observation belonging to the i -th group also referred to as group effect;
- θ_t time-specific effect;
- v_{it} the random component of the model, $v_{it} \sim IID(0, \sigma_v^2)$.

The panel data model assumes that all coefficients are constant. Group effects μ_i reflect the individual characteristics of units that are constant over time for a particular entity (they are not subject to change over time). Time effects θ_t remain constant for all objects at a given time. The presence of effects of both types (group and time) in the regression equation (2) defines a two-way panel data model, while the presence of only one type of effects defines a one-way panel data model. In each of these cases, they can also be random-effects or fixed-effects models.

In the fixed-effects model, the differences between objects and periods are expressed by assigning this information to a component specific to the object or to a period, which is incorporated in the equation by coding — creating zero-one variables for

a group or time (dummy variables). The differences between group and time units are captured by the intercepts in the model, which means that each dummy variable (except the reference one) is described by its specific constant. The correlation between individual errors and exogenous variables is allowed (Park, 2011).

The estimated panel data model requires verification which will confirm its accuracy and suitability. The first commonly used tool is the F test; rejecting the null hypothesis implies that the combined influence of individual effects on the endogenous variable is significant.

When applying the Breusch–Pagan test, it can be verified whether the variance of the individual effects component (group or time) is zero; rejection of the null hypothesis allows to conclude that the model with specified group or time effects is better than the model in which these effects have not been specified. In turn, in the Hausman test, rejection of the null hypothesis means the possibility of a correlation between exogenous variables and random effects, which implies the legitimacy of building a fixed-effects model. Then, the LSDV (Least Squares Dummy Variables) method is used to estimate this type of model.

In the study Excel and GRETL computer programs were used to pre-process data and conduct all the calculations.

3. DATA DESCRIPTION

In the proposed panel data models, a voivodeship (region) is the entity (organisation) unit and a year is the time unit, to which all the variables refer. There are 16 regions in Poland and the analysed period covers 2012–2018. No data are missing as all entities have measurements in the whole period. The same entities are observed for each period. Thus, a well-organised balanced fixed panel data set (Park, 2011) is subject to the analysis. It consists of 112 records ($N \cdot T = 16 \cdot 7$). The majority of data were acquired from the Statistics Poland Local Data Bank (SPLDB) website (<https://bdl.stat.gov.pl/BDL/dane/podgrup/temat>, 19-22.02.2020). However, some specific information on expenditure on national roads was kindly provided by the General Directorate for National Roads and Motorways, Poland, at the request of the authors.

A variety of indices are used to measure road traffic safety. The level of fatality is commonly represented as the number of fatalities per 100 000 people

(IRTAD, 2014), but it has some limitations in highly populated and poorly motorised regions. Another indicator is fatalities per distance travelled. Yet, this is not always easily available. Instead, fatalities per 10 000 registered vehicles are utilised, which, though, may be misleading when traffic levels are different (IRTAD, 2009).

Considering the multi-faceted approach to the phenomenon under study, certain indicators have been proposed as measures for the level of road traffic safety. They are relative measures defined as the endogenous (output) variables, arising from the concept of a variety of measures used in the comparison of traffic accident data between countries or between regions (Farchi et al., 2006):

- RA100KM — Road Accidents calculated as the number of road traffic accidents per 100 road kilometres. Data source: SPLDB;
- RAFR100KM — Road Accident Fatalities calculated as the number of deaths due to road traffic accidents (according to the Vienna Convention's international criterion) per 100 road kilometres. Data source: SPLDB;
- RAI100KM — Road Accident Injury calculated as the number of injuries due to road traffic accidents per 100 road kilometres. Data source: SPLDB.

Despite the fact that demographic indicators are commonly used in the literature, which means that a threat variable (such as the number of accidents, fatalities, or injuries) is related to the fixed number of human population unit, another approach is adopted in this work.

A fixed road length unit has been proposed as the reference, which is connected with the fact that the variability of this reference is smaller than that of the population reference. This is particularly important when data have a panel structure.

Selected characteristics on demography, economy, and road infrastructure were considered as exogenous (input) variables used to diagnose their influence on the road traffic safety expressed by measures calculated from the data on road traffic accidents and their severity.

Variables were selected characterising sources that may influence existing road safety conditions: socio-economic features, road-traffic conditions, expenditure on national roads made by the General Directorate for National Roads and Motorways and by the local governments.

The first group of variables characterises the socio-economic growth of the region; they can be

considered as a stimulus for the growth of motorisation, road traffic and exposure.

- **GDPPC** — Gross Domestic Product Per Capita (in PLN) is the broadest quantitative measure of a nation's total economic activity. It represents the monetary value of all goods and services produced within specific geographic boundaries over a given period. Data source: SPLDB.
- **RUI** — Region Urbanisation Indicator is the number of people that live in urban areas in relation to the total number of the region inhabitants. In Poland, an urban area is a locality that has been granted a city charter. The level of urbanisation influences the magnitude of traffic generation effects. Data source: SPLDB.

Another category of variables refers to a region's motorisation indicators, defined using the number of motorised vehicles for the region. Vehicle type was taken into account as this attribute is strongly connected with road traffic safety.

- **CMR** — Car Motorisation Rate is the number of passenger cars per 1 000 inhabitants. Data source: SPLDB.
- **TMR** — Truck Motorisation Rate is the number of trucks (light, medium, and heavy) per 1 000 inhabitants. Data source: SPLDB.
- **MMR** — Motorcycle Motorisation Rate is the number of motorcycles per 1 000 inhabitants. Data source: SPLDB.

The remaining input variables relate to road infrastructure, which is one of the most important road safety components. A well-developed and modern road network, with the appropriate density of highways, expressways and express roads, is a precondition for a properly functioning national economy. Thus, in Poland, the intensive modernisation and new road investments have significantly accelerated, especially since the accession to the EU. The scope of these activities and amounts spent depend on the classification of Polish roads. According to relevant

Tab. 1. Characteristics of Polish public road categories

ROAD CATEGORY	ROAD CLASS	ROAD OWNER	ROAD MANAGER (ADMINISTRATOR)
National roads	Motorway, expressway, main road of accelerated traffic	The Treasury	General Director of National Roads and Motorways, Concessionaire
Voivodeship roads	Main road of accelerated traffic, main road	Voivodeship self-government	Voivodeship Board
County roads	Main road of accelerated traffic, main road, collector road	County self-government	County Board
Communal roads	Main road of accelerated traffic, main road, collector road, local road, local access road	Communal self-government	Head of the Commune (Mayor, Mayor of the City)

Source: authors' own elaboration

law regulations, there are four public road categories (<https://www.lexlege.pl>, 11.04.2020): national, voivodeship, county and communal, as presented in Table 1, where information is ordered from the highest to the lowest road category. Each road category is determined by technical conditions and operational requirements; all highways and expressways are national roads in Poland. The road owner covers all expenses related to construction, road network maintenance and repairs. However, based on mutual agreements and cooperation, funds may be transferred to cover expenses for roads of other categories. A competent road manager is responsible for the implementation of the tasks related to these direct road expenses.

The following exogenous variables were used to describe the Polish road infrastructure in the analysis.

- **DCR** — Dual Carriageway Ratio is the percentage of the length of two-way roads in the total length of public roads. Such roads are designed to meet higher standards than one-way roads, separated by a central reservation for traffic travelling in opposite directions. Dual carriageways (among other benefits) improve road traffic safety over single carriageways. Data source: SPLDB.
- **RLPC** — Road Length Per Capita is the indicator of the length of public roads in kilometres to the number of voivodeship inhabitants.
- **SGTERK** — Self-Government Total Expenditure per Road Kilometre (per one kilometre of public roads), in thousands of PLN. The variable represents the total expenditure of voivodeships, and counties and communes (belonging to the voivodeships) per one kilometre of public roads (national, voivodeship, county, and communal). Data source: SPLDB.

In the analysed period (2012–2018), the national roads accounted for only about 7% of the public road network. However, they carry more than 60% of total

traffic, there were 24.5%–28% of the total number of accidents recorded on these roads, which resulted in as many as 36%–39% of the total number of road accident fatalities. Raising road standards, especially as regards the national roads network, is said to improve the situation. Therefore, the information concerning major and current expenditure on national roads made by the General Directorate for National Roads and Motorways (GDNRM or GD) has been considered. The following variables were included in the analysis to evaluate how the outlays for the most important category roads were effective in terms of general road traffic safety.

- GDICERK — General Directorate Investment Construction Expenditure per one kilometre of national roads, in thousands of PLN. Data source: GDNRM.
- GDRRERK — General Directorate expenditure on Road network Repairs per one kilometre of national roads, in thousands of PLN. Data source: GDNRM.
- GDCRMERK — General Directorate expenditure on Current Road network Maintenance per one kilometre of national roads, in thousands of PLN. Data source: GDNRM.

4. MODELLING RESULTS

In the modelling procedures, the last category of the cross-section and time unit variables is the reference. This means that the Mazowieckie Voivodeship is the reference for cross-section, as is the year 2018 for time. The robust standard errors technique was employed to obtain unbiased standard errors of OLS coefficients under heteroscedasticity. Table 2 presents the outcome of the modelling. The endogenous variable names were used to identify related models further on in the study.

The first part of Table 2 contains the results of consecutive stages of research leading to the final forms of the models. Using the F test, the hypothesis of individual effects was verified. The p-values for individual models indicate the rejection of the null hypothesis, which implies the validity of using the panel data model for each considered endogenous variable. The results of Breusch–Pagan and Hausman tests justify the choice of fixed-effects models.

For each endogenous variable, two separate models were constructed: the full model, which includes all the proposed exogenous variables, and the re-estimated model with variables obtained in the

backward selection procedure, in which the significance level for entering a variable into the model was set to 10%. As can be seen, the sign and magnitude of all the variables significant in the backward selection models are consistent with those of the respective full models, indicating generally stable results.

For each model, the elimination of insignificant variables increased the adjusted R^2 value, thus confirming the validity of the model final form obtained in the backward selection procedure.

Adjusted R^2 exceeding 95% and the graphic illustration of predicted and observed values in Fig. 2a and 2c, respectively, indicate that models for RA100KM and RAI100KM variables are high-quality models. A slightly worse fit (adjusted R^2 does not exceed 80%) was obtained in the case of the RAF100 variable, which is due to the relatively high variability of the feature within individual voivodeships (Fig. 2b).

The differences in the values of the analysed endogenous variables by voivodeships are presented in the form of box plots in Fig. 3. Relating them to the obtained results, it may be concluded that, in principle, the structure of mutual relationships identified in the models reflects the patterns illustrated in the plots.

5. DISCUSSION OF THE RESULTS

The discussion of estimation results presented below refers to the final models.

GDPPC — a standard measure of economic well-being was found to be statistically significant in the case of the RAF100KM model but insignificant for the RA100KM and RAI100KM models. At the same time, the CMR variable turned out to be significant for the last two models discussed but insignificant for the first one. What is more, the direction of the effects' influence is the same, and the magnitude of their influence is remarkably similar. It can be assumed that both exogenous variables under consideration are a proxy of voivodeship economic development and growth (thus being also correlated). The findings confirm that better economic conditions of regions are associated with lower rates of accidents, fatalities as well as injuries.

Road traffic with a high proportion of trucks may imply heightened hazard as regards road traffic safety. All the models estimated in the study confirm this relationship; both the road accident number and the road accident casualty number (fatalities and injuries)

Tab. 2. Estimation results of the panel data models for the three road traffic safety measures (endogenous variables)

MODEL NAME	RA100KM	RAF100KM	RAI100KM			
F test DF(15,85)	F = 49.55 (p-value = 0.00)	F = 5.91 (p-value = 0.00)	F = 43.04 (p-value = 0.00)			
B-P test	Asymptotic Chi-2 = 161.05 (p-value = 0.00)	Asymptotic Chi-2 = 19.56 (p-value = 0.00)	Asymptotic Chi-2 = 134.94 (p-value = 0.00)			
H a u s m a n test	Asymptotic Chi-2 = 47.25 (p-value = 0.00)	Asymptotic Chi-2 = 39.61 (p-value = 0.00)	Asymptotic Chi-2 = 61.85 (p-value = 0.00)			
PANEL DATA MODEL ESTIMATION RESULTS						
MODEL TYPE	FULL*	BACKWARD SELECTION	FULL*	BACKWARD SELECTION	FULL*	BACKWARD SELECTION
Exogenous variable	Estimator (p value)	Estimator (p value)	Estimator (p value)	Estimator (p value)	Estimator (p value)	Estimator (p value)
Intercept	3.89 (0.74)	4.28 (0.00)	0.1 (0.95)	0.49 (0.00)	6.54 (0.76)	16.94 (0.00)
GDPPC	-0.05 (0.32)		-0.01 (0.43)	-0.01 (0.01)	-0.05 (0.36)	
RUI	-19.97 (0.16)	-14.90 (0.00)	1.51 (0.59)	1.04 (0.00)	-27.46 (0.28)	-14.45 (0.00)
CMR	-0.01 (0.32)	-0.01 (0.05)	0 (0.59)		-0.02 (0.45)	-0.04 (0.00)
TMR	0.20 (0.00)	0.11 (0.00)	0.01 (0.38)	0.005 (0.00)	0.26 (0.04)	0.20 (0.00)
MMR	0.01 (0.97)		0.02 (0.39)	0.01 (0.00)	-0.03 (0.84)	
DCR	-0.22 (0.66)		-0.02 (0.79)		-0.65 (0.25)	
RLPC	136.40 (0.49)		-47.93 (0.05)	-53.82 (0.00)	296.18 (0.36)	
SGTERK	0.03 (0.01)	0.02 (0.00)	0 (0.47)		0.03 (0.00)	0.03 (0.00)
GDICERK	0 (0.65)		0 (0.35)	0.00004 (0.02)	0 (0.09)	
GDRRERK	0 (0.88)		0 (0.89)		0 (0.66)	
GDCRMERK	0 (0.71)		0 (0.26)	-0.002 (0.02)	-0.01 (0.67)	
CS1	6.51 (0.00)	4.63 (0.00)	-0.11 (0.62)	-0.19 (0.00)	9.64 (0.01)	6.80 (0.00)
CS2	0.03 (0.99)		-0.24 (0.40)	-0.28 (0.00)	0.50 (0.88)	
CS3	-3.22 (0.36)	-1.73 (0.00)	0.08 (0.88)		-4.38 (0.37)	-1.81 (0.00)
CS4	1.44 (0.52)	1.41 (0.00)	-0.26 (0.25)	-0.31 (0.00)	3.32 (0.31)	3.02 (0.00)
CS5	5.82 (0.00)	5.14 (0.00)	-0.15 (0.35)	-0.20 (0.00)	8.14 (0.00)	7.46 (0.00)
CS6	2.54 (0.35)	2.83 (0.00)	-0.07 (0.88)	-0.19 (0.00)	3.35 (0.40)	3.41 (0.00)
CS7	1.35 (0.63)	1.09 (0.00)	0.18 (0.68)		2.00 (0.66)	3.65 (0.00)
CS8	-0.70 (0.88)		-0.22 (0.76)	-0.30 (0.00)	-0.06 (0.99)	
CS9	-0.94 (0.71)		0.01 (0.98)		-1.60 (0.66)	-1.61 (0.00)
CS10	7.72 (0.00)	6.96 (0.00)	-0.25 (0.13)	-0.34 (0.00)	10.92 (0.00)	8.86 (0.00)
CS11	12.33 (0.00)	9.19 (0.00)	-0.35 (0.38)	-0.42 (0.00)	17.02 (0.01)	11.09 (0.00)
CS12	-6.48 (0.06)	-2.74 (0.00)	0.05 (0.93)		-9.03 (0.09)	-4.66 (0.00)
CS13	3.33 (0.20)	3.45 (0.00)	-0.07 (0.80)	-0.11 (0.00)	4.57 (0.19)	3.58 (0.00)
CS14	-2.01 (0.38)	-1.05 (0.00)	-0.19 (0.60)	-0.25 (0.00)	-2.19 (0.50)	
CS15	3.52 (0.10)	3.09 (0.00)	-0.35 (0.17)	-0.35 (0.00)	5.10 (0.14)	2.89 (0.00)
Year 2012	2.80 (0.15)	2.71 (0.00)	0.15 (0.48)	0.16 (0.00)	3.25 (0.34)	0.51 (0.02)
Year 2013	2.29 (0.18)	2.25 (0.00)	0.13 (0.53)	0.13 (0.00)	2.59 (0.38)	0.34 (0.01)
Year 2014	1.72 (0.25)	1.75 (0.00)	0.11 (0.55)	0.11 (0.00)	1.83 (0.48)	
Year 2015	1.32 (0.25)	1.34 (0.00)	0.04 (0.77)	0.05 (0.03)	1.39 (0.48)	
Year 2016	1.33 (0.09)	1.36 (0.00)	0.04 (0.73)	0.04 (0.05)	1.59 (0.23)	0.65 (0.00)
Year 2017	0.73 (0.10)	0.71 (0.00)	-0.01 (0.84)		0.95 (0.19)	0.40 (0.00)
Adjusted R ²	0.9740	0.9744	0.7759	0.7960	0.9727	0.9732
AIC	133.633	124.946	-329.891	-346.322	194.218	184.226

* The 10% significant effects for full models are marked by a grey background.

Source: authors' own elaboration

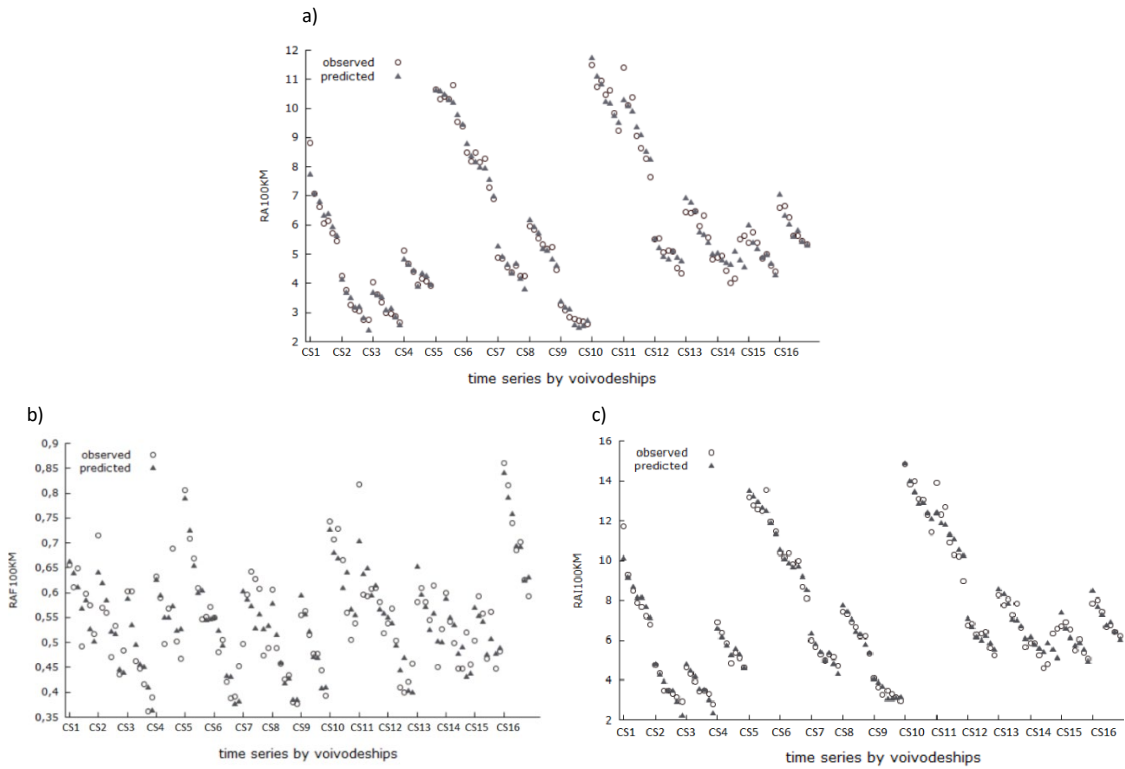


Fig.2. Illustration of the goodness of fit for the panel data models for the variables: RA100 (a), RAF100 (b), RAI100 (c)
Source: authors' own elaboration

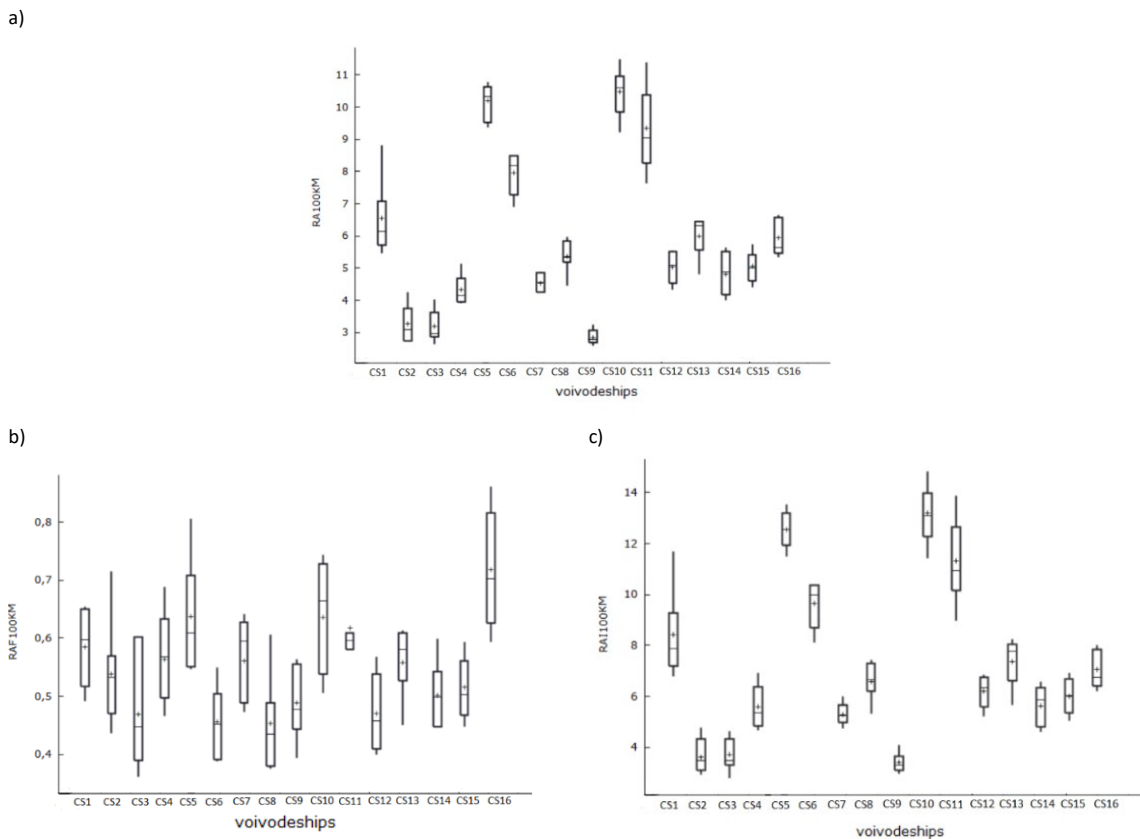


Fig. 3. Illustration of the goodness of fit for the panel data models for the variables: RA100 (a), RAF100 (b), RAI100 (c)
Source: authors' own elaboration

increase along with the TMR value (increase). The last indicator of motorisation rate MMR proved to be significant and positively associated with the fatality rate in the RAF100KM model. This is consistent with the fact that motorcyclists, who are vulnerable road users, are highly likely to be fatal accident casualties. Rising motorisation rates for trucks and motorcycles lead to increased traffic fatalities while rising motorisation rate for cars does not.

The urbanisation indicator (RUI) is inversely related to the accident rate and injury rate while positively related to the increase in the fatality rate. The higher the urbanisation rate, the smaller the number of accidents and injuries per road kilometre, but the higher the number of fatalities (thus, more serious consequences). On the one hand, the differences may result from the fact that a much larger percentage of unprotected road users prone to exposure (pedestrians, cyclists) occur in cities (accidents involving such participants reach up to $\frac{1}{4}$ of the total number of accidents). On the other hand, the relationship between RUI and RA100KM and RAI100KM may become reverse, which may be connected with the progressive improvement of rural road infrastructure. The results seem to be surprising; therefore, further research is necessary.

The DCR variable turned out not to be significant in any of the models. Still, its role as a specific measure of road infrastructure development, which is an essential element of the overall development of the region, is represented in the models by the GDPPC or CMR variables.

The indicator of the length of public roads to the number of voivodeship inhabitants (RLPC) affects accident fatality rates in the RAF100KM model, but in the remaining models, it turns out to be insignificant. Extending the roadway network can be associated with the improvement of road quality as a consequence of the investment or modernisation processes. This aspect of increasing the length of roads and relating it to the population density can account for the decline in road hazard, expressed in the reduction of the number of fatalities. Yet, the result should be interpreted carefully since other factors, featuring the heterogeneity of the RLPC variable, such as intersection density or the impact of the road classification (national, voivodeship, etc.), were not included in the model.

Road expenditures investigated in the study work on distinct levels. Local authorities finance regional (voivodeship, county and commune) roads and cover some of the costs directly related to national roads.

These total amounts are provided by the SGTERK variable. However, in the case of national (major) roads, expenses have been differentiated according to their structure by considering such variables as GDICERK, GDRRERK and GDCRMERK. While it is presumed that an increase in road expenditure improves road safety outcomes, the obtained results fail to confirm this. Such a result may be due to several factors, such as not considering the time delay of SGTERK, GDICERK, GDRRERK variables (the expected effect of the investment requires time) and the fact that the structure of expenditure of local administrations has been left out of consideration. Only in the RAF100KM model, the GDCRMERK variable is statistically significant, and has a negative effect on the endogenous variable. The increasing expenditure on the current maintenance of national roads immediately translates into a reduction in the number of fatalities on these roads (this was also reflected in the overall picture of road network safety).

Individual intercepts are included to identify individual-specific group and time characteristics. These intercepts are called fixed effects.

The specification of the two-way panel data model made it possible to confirm the correctness of the inclusion of time effects in the research. The significance of these effects in the models RA100KM, RAF100KM, RAI100KM along with the identified, decreasing time trend suggest that new advances in vehicle technology, national transportation safety policies, educational initiatives, rising public awareness, and other dedicated activities have changed over time, with safety generally improving year to year.

Also, most of the cross-sectional effects are highly significant; in the total number of 15 dummy variables, 12 appeared in the RA100KM and RAI100KM backward models and 11 in the RAF100KM backward model. This means that region specificity factors (such as: geographical conditions, road user features, business-oriented activity, educational and environmental aspects, and also administrative policies) not included in the models, might significantly affect the differences between voivodeships in the considered measures of the level of road traffic safety. However, the significance is not uniform. There are eight cross-sections (CS1, CS4, CS5, CS6, CS10, CS11, CS13, CS15) evidently distinguished from the reference cross-section (Mazowieckie) by the three rates: accident, fatality, and injury. These effects are positive in both RA100KM and RAI100KM

models, but not of similar magnitude. Only for three voivodeships (Lubelskie — CS3, Świętokrzyskie — CS12, and Wielkopolskie — CS14) smaller values of RA100KM were identified; these regions are also characterised by smaller values of the RAI100KM indicator or its lack of significance. In relation to all the region effects included in the RA100KM and RAI100KM models, the Śląskie Voivodeship (CS11) performs worst (relatively large positive influence) while the Świętokrzyskie Voivodeship (CS12) performs best (relatively large negative influence). Although the above results may arise from the specifics of the voivodeships, completely different factors may define these specifics (historical, economic and social conditions). The RA100KM results show that the Mazowieckie Voivodeship is doing worse than 11 Polish regions in terms of accident fatality rate (statistically significant negative values for respective cross-section parameters). This somewhat surprising outcome might be a consequence of the fact that both the Region Capital City of Warsaw and the Mazowieckie Voivodeship have been treated as one.

CONCLUSIONS

The objective of the research was to investigate the relative road safety performance of the Polish voivodeships, with a special focus on expenditures on roads. The study set out to identify factors significantly affecting the measures of the road safety level expressed in terms of accident rate as well as fatality and injury rates. The two-way panel data models with fixed effects were built for the annual data from 2012 to 2018 and for 16 Polish regions.

The panel model results suggest a varied impact of motorisation rates in terms of trucks and passenger cars on road safety. Motorcycles have proved to be positively relevant only to fatality rate. It has been found that the effect related to greater exposure to accident rate and injury rate is stronger for trucks than for passenger cars.

Self-government total road expenditure turned out to be significantly and positively associated with road accident rate and injury rate, while the increase of national road maintenance expenditure significantly contributed to a reduction in the road fatality rate. These promising results require a more in-depth analysis.

It has been detected that there were unquestionable differences among voivodeships taking into account the studied endogenous variables. That con-

siderable variation in road safety could be the key factor in planning road investments and other dedicated activities, in particular, intensified police patrols. It is optimistic that in most voivodeships, the values of the examined road safety measures decreased over the considered period.

Panel data models are extremely helpful in analysing various questions related to road traffic safety policy in different regions. In particular, such models could be used in identifying fund allocation based on the relative risk exposure of the regions.

The results obtained for the road accident rate and the road accident injury rate were very similar, which suggests that the indicators related to the number of road injuries and the number of road fatalities are sufficient for modelling road safety, obviating the need to create models for the number of accidents.

ACKNOWLEDGEMENT

The authors would like to thank the General Directorate for National Roads and Motorways, Poland for their effort in preparing and providing data for the research.

LITERATURE

- Ahangari, H., Atkinson-Palombo, C., & Garrick, N.W. (2017). Automobile-dependency as a barrier to vision zero, evidence from the states in the USA. *Accident Analysis and Prevention*, 107, 77-85. doi: 10.1016/j.aap.2017.07.012
- Albalade, D., & Bel-Piñana, P. (2019). The effects of public private partnerships on road safety outcomes. *Accident Analysis and Prevention*, 128, 53-64. doi: 10.1016/j.aap.2019.03.015
- Albalade, D., Fernández, L., Yarygina, A. (2013). The road against fatalities: Infrastructure spending vs. regulation? *Accident Analysis and Prevention*, 59, 227-239. doi: 10.1016/j.aap.2013.06.008
- Antoniou, C., Yannis, G., Papadimitriou, E., & Lassarre, S. (2016). Relating traffic fatalities to GDP in Europe on the long term. *Accident Analysis and Prevention*, 92, 89-96. doi: 10.1016/j.aap.2016.03.025
- Badura, D. (2017). Urban traffic modeling and simulation. *Forum Scientiae Oeconomia*, 5(4), 85-97. doi: 10.23762/FSO_VOL5NO4_17_7
- Baltagi, B. H. (2005). *Econometric Analysis of Panel Data*. John Wiley & Sons, Ltd.
- Besharati, M., Kashani, A.T., & Washington, S. (2020). A comparative analysis of road safety across the provinces of Iran from 2005 to 2015. *International Journal of Sustainable Transportation*, 1-9. doi: 10.1080/15568318.2019.1710877

- Castro-Nuño, M., Castillo-Manzano, J. I., & Fageda, X. (2018). The role of road safety in a sustainable urban mobility: An econometric assessment of the Spanish NUTS-3 case. *International Journal Of Sustainable Transportation*, 12(3), 205-217. doi: 10.1080/15568318.2017.1346733
- Dupont, E., Commandeur, J. J. F., Lassarre, S., Bijleveld, F., Martensena, H., Antoniou, C., Papadimitriou, E., Yannis, G., Hermans, E., Pérez, K., Santamarina-Rubio, E., Usami, D. S., & Giustiniani, G. (2014). Latent risk and trend models for the evolution of annual fatality numbers in 30 European countries. *Accident Analysis and Prevention*, 71, 327-369. doi: 10.1016/j.aap.2014.06.009
- Farchi, S., Molino, N., Rossi, P. G., Borgia, P., Krzyzanowski, M., Dalbokova, D., Kim, R., & The European Road Accident Indicator Working Group (2006). Defining a common set of indicators to monitor road accidents in the European Union. *BMC Public Health*, 6(183), 1-12. doi: 10.1186/1471-2458-6-183
- Hauer, E. (2010). Cause, effect and regression in road safety: a case study. *Accident Analysis and Prevention*, 42(4), 1128-1135. doi:10.1016/j.aap.2009.12.027
- Kweon, Y. J., & Kockelman, K. M. (2015). Safety Effects of Speed Limit Changes: Use of Panel Models, Including Speed, Use, and Design Variables. *Transportation Research Record: Journal of the Transportation Research Board*. doi: 10.3141/1908-18
- Laaraj, N., & Jawab, F. (2018). Road accident modeling approaches: literature review. *International Colloquium on Logistics and Supply Chain Management (LOGISTIQUA)*, Tangier, 188-193. doi: 10.1109/LOGISTIQUA.2018.8428276
- Muthusamy, A. P., Rajendran, M., Ramesh, K., & Sivaprakash, P. (2015). A Review on Road Traffic Accident and Related Factors. *International Journal of Applied Engineering Research*, 10(11), 28177-28183.
- Park, H. M. (2011). *Practical Guides To Panel Data Modeling: A Step by Step Analysis Using Stata*. Tutorial Working Paper. Graduate School of International Relations, International University of Japan.
- Söderlund, N., & Zwi, A. B. (1995). Traffic-related mortality in industrialized and less developed countries. *Bull. World Health Organization*, 73(2), 175-182.
- Statistics Poland Local Data Bank. Retrieved from <https://bdl.stat.gov.pl/BDL/dane/podgrup/temat>
- Ustawa o drogach publicznych. Stan prawny aktualny na dzień: 05.06.2020. Dz.U.2020.0.470 t.j. - Ustawa z dnia 21 marca 1985 r. o drogach publicznych [Act on public roads. Legal status as at: 05.06.2020. Setup log 2020.0.470 - Act of 21 March 1985 on public roads]*. Retrieved from <https://www.lexlege.pl/ustawa-o-drogach-publicznych>
- Wachnicka, J., Kustra, W., Jamroz, K., & Budzyński, M. (2018). Road infrastructure safety management tools for national and voivodeship roads in regions (NUTS 2) in Poland. *12th International Road Safety Conference GAMBIT 2018 "ROAD INNOVATIONS FOR SAFETY - the national and regional perspective*.
- Yaseen, M. R., Ali, Q., & Khan, M. T. I. (2018). General dependencies and causality analysis of road traffic fatalities in OECD countries. *Environmental Science and Pollution Research*, 71, 327-336. doi: 10.1007/s11356-018-2146-4