Integrating Crash Severity in Roadside Safety Quantitative Analysis
Assessing Partial Proportional Odds Models

CARLOS ROQUE
JOÃO LOURENÇO CARDOSO
Laboratório Nacional de Engenharia Civil,
Departamento de Transportes. Núcleo de Planeamento, Tráfego e Segurança

Robust knowledge of the underlying factors involved in run-off-road (ROR) crash occurrences and resulting injuries is a prerequisite for the development of sound methods to support roadside cost-efficient design and redesign and related asset management/road operations decisions. Over recent years, the understanding of ROR crashes on Portuguese roads has significantly increased due to roadside safety research carried out by the Laboratório Nacional de Engenharia Civil, emphasizing the importance of this type of crash in the overall interurban safety picture.

In this paper investigations of ROR crash injury severity on Portuguese freeways are reported, exploring the application of a partial proportional odds (PPO) model to study the contributors influencing ROR crash severities. The PPO model allows the covariates that do not meet the proportional odds assumption to have diverse effects at different severity levels.

This study is based on a detailed data set of ROR crashes that occurred on Portuguese freeways during the years 2009 and 2010.

Several variables in seasonal attributes, roadway and roadside attributes, crash characteristics and driver information were tested. Specifically, the use of the partial proportional formulation allows a superior identification of the varying effect of several variables on ROR crash injury severity. Furthermore it includes the effect of traversing ditches, which previously was masked when fitting the unordered framework models.

A comparison between the application of PPO models and mixed logit models for ROR crash severity evaluation is also included here. The study shows that the PPO model is a viable method for analyzing ROR crash severities.

INTRODUCTION

Robust knowledge of the underlying factors involved in run-off-road (ROR) crash occurrences and resulting injuries is a prerequisite for the development of sound methods to support roadside cost-efficient design and redesign and related asset management/road operations decisions. In Portugal, single-vehicle run-off-road (ROR) crashes result in ten thousand crashes with roadside features every year and account for approximately half of all freeway fatalities. Portuguese crash data (2007-2010) indicate that roadside geometry – including slopes, embankments, and ditches – contributes to more than half of all ROR accidents involving serious injury or death (*1*).

Over recent years, the understanding of ROR crashes on Portuguese roads has significantly increased due to roadside safety research carried out by the *Laboratório Nacional de Engenharia Civil*, emphasizing the importance of this type of crash in the overall interurban safety picture. This allowed the development of a computer-aided procedure for supporting cost-effective decisions with regard to roadside safety alternative interventions. As already described in previous research by the authors, the procedure is based on cost–benefit analysis and makes extensive use of dedicated ROR crash prediction models (*2, 3*).

Currently, one limitation of this procedure is the non-consideration of the probability of different severity level outcomes conditioned on crash occurrence. This may impact crash cost estimations used when choosing amongst relevant alternatives. Hence, research towards the improvement of methods for considering crash severity has been carried out, attempting to integrate crash severity models in the said procedure in order to estimate the expected number of injuries at different severity levels and thus improve the estimation of ROR crash costs.

When studying crash severity injury outcomes, the most common study approaches may be grouped into unordered framework models and ordered framework models. The former includes, for example, the multinomial and mixed logit models already mentioned. Multinomial logit models are traditional discrete outcome models that consider three or more outcomes and do not explicitly consider the ordering that may be present in these outcomes. Mixed logit models are a more recent development for the analysis of discrete data that addresses the limitations of the multinomial logit (susceptibility to correlation of unobserved effects from one injury-severity level to the next) by allowing for heterogeneous effects and correlation in unobserved factors (*4, 5*).

The latter framework includes ordered probit or logit models, among others. In these models, the discrete injury severity levels are assumed to be associated with an underlying continuous latent variable (z) that is used as a basis for modeling the ordinal ranking of data. This unobserved variable is typically specified as a linear function for each crash observation, such that $z = \beta X + \varepsilon$, where X is a vector of variables determining the discrete ordering for each crash observation, β is a vector of unknown parameters to be estimated, and ε is a random disturbance term (*4, 5*). In this framework, crash injury severity outcomes are reported as an ordinal scale variable (such as no injury, minor injury, severe injury, and fatal injury). The ordered framework models explicitly recognize the inherent ordering within the outcome variable (as the severities become increasingly severe from no injury, to minor injury, to severe injury, to fatality) whilst in the non-ordered analyses it is completely ignored.

In a previous study, multinomial and mixed logit models were developed to explain ROR crash severity and detect unforgiving roadside contributors (*6*). The empirical findings showed the contribution of critical slopes and vehicle rollover towards increased probability of fatal injuries and highlighted the importance of introducing the "forgiving roadside" concept in Portuguese road design standards, namely to mitigate ROR crash severity on Portuguese freeways.

This study considers the ordered nature of crash injury severity. Thus, ordered framework models were used to examine the effect of various contributing factors to driver injury severity levels in ROR crashes on Portuguese freeways. These models represent the outcome process under consideration using a single latent propensity. Thus, the outcome probabilities are determined by partitioning the uni-dimensional propensity into as many categories as the outcome variable alternatives through a set of thresholds (*7*). However, it is important to keep in mind that these models are intrinsically case specific because they are limited to and constrained by the available data, which may be improved over time.

The main focus of this study is to investigate ROR crash injury severity, to study the contributors influencing ROR crash severities on freeways.

Accordingly, the modeling approach is mainly explanatory (based on past observations) rather than predictive (predicting new values for the future). Furthermore, a partial proportional odds (PPO) models is used, a statistical technique not yet found in reported ROR crash research. In the literature several unordered framework models were found: the multinomial and mixed logit models (*6, 8, 9, 10, 11*), the nested logit models (*11, 12, 13*) and the latent class logit model (*14*). In addition, only two ordered framework models were found: ordered probit models, by Renski et al. (*15*) and Kockelman and Kweon (*16*).

The PPO model allows the covariates that meet the proportional odds assumption to affect different crash severity levels with the same magnitude. At the same time, the covariates that do not meet the proportional odds assumption can have diverse effects at different severity levels (*17*). Thus, this model ensures minimal complexity of the analysis framework while allowing some flexibility from the multinomial and mixed logit models.

A comparison between the application of PPO models and mixed logit models for ROR crash severity evaluation is also included here. In the final section, measures to be taken into consideration in supporting decisions on roadside safety design in Portugal are discussed based on the empirical findings.

METHODOLOGY

Crash severity models focus on the estimation of the probability of a crash resulting in one or more fatalities, severe injuries, minor injuries or property damage only (PDO) given the occurrence of the crash. Savolainen *et al.* (*4*) and Mannering and Bath (*18*) extensively reviewed the numerous methodological techniques applied in studying crash severity data. The most common options found in the literature when studying crash severity injuries can be grouped into unordered framework models (like multinomial logit (MNL), nested logit, probit and mixed logit models) and ordered framework models (including ordered probit or logit models, generalized ordered models and PPO models). In this paper, by focusing the attention on ordered models, the ordered nature of crash injury severity is favored, a characteristic that cannot be ignored completely.

Occasionally, it is more realistic to assume that the explanatory variables may vary across crashes; therefore, some researchers have used fixed parameters models (like Kockelman and Kweon (*16*)) others have used random parameters or mixed effects models (e.g. Roque et al. (*6*) and Wu et al. (*11*). Random parameters models have the advantage of allowing the explanatory variables to take into account the individual differences among injury severity levels in different crashes.

In this study, two models were estimated. A PPO model was estimated using R (version 3.2.5) (*19*). “VGAM” (*20*) R package was used. The freeware BIOGEME software (*21*) was used for mixed logit model estimation.

Partial Proportional Odds Model

To Savolainen *et al.* (*4*) crash severity is ordinal in nature and recognizing this feature it is important to select the appropriate analysis tool, justifying the use of ordered framework models. In this study, driver injury severity is categorized into three levels of increasing severity and coded as: 1 = no injury, 2 = minor injury, 3 = severe or fatal injury.

On the one hand, traditional ordered logit models require data that adhere to the proportional odds assumption between different severity levels, i.e., the effect of an explanatory variable will be uniform for all levels of the outcome variable (e.g., the deployment of an airbag may decrease the probability of a fatality and also increase the probability of no injury, or vice versa) (*23*) Imposing such restriction can lead to inconsistent parameter estimation (*7*).On the other hand, while mixed logit models completely ignore the sequential order of injury severity levels (in this case the deployment of an airbag may, e.g., decrease the probability of a fatality or increase the probability of no injury). In fact, in crash severity analysis, it is not logical to assume that the proportional odds assumption will be satisfied by all explanatory variables (in reality the deployment of an airbag may decrease the probability of both fatality and no injury (because the airbag itself may cause some minor injuries) nor to ignore the ordered nature of crash injury severity. The PPO model allows certain individual explanatory variables to affect each level of the response variable differently, while other independent predictors may adhere to the proportional odds assumption, if they are found not to violate this assumption based on relevant statistical tests (e.g., Wald test) (*24, 25*).

If j denotes the crash severity level (1 to 3) and J represents the number of severity levels (here $J = 3$), then the form of the PPO model is as follows (*26*):

$$\Pr(Y_i > j) = \frac{\exp[\alpha_j + (X_i\beta + T_i\gamma)]}{1 + \exp[\alpha_j + (X_i\beta + T_i\gamma)]}, j = 1,2, \ldots, J - 1 \tag{1}$$

where Y_i represents the observed severity for crash i; plus, γ and β are the vectors of parameter estimations that do and do not violate the parallel line assumption, respectively. The corresponding vectors of explanatory variables that do and do not violate this assumption are T_i and X_i, respectively; and α_j is the cutoff term for the thresholds in the model.

In order to determine which predictor variables will belong to the subset q that rejects the proportional odds assumption, each variable was analyzed individually using a Wald test of proportional odds. This test takes the multinomial response variable and dichotomizes it based on cumulative probability, using $P(Y_i \geq j)$ and $P(Y_i < j)$ for each crash severity level j. This method determines whether the effect of a variable will remain the same across all "cuts" of j (*17*).

Special care must be exercised when interpreting the coefficients of intermediate categories in PPO models. The sign of β does not always determine the direction of the effect of the intermediate outcomes. Thus, marginal effects were used in this study for interpretation of the variables (*27*). The marginal effects estimated for an explanatory variable, measure how changes in the explanatory variable affect the outcome variable.

Mixed Logit Model

According to Train (*5*), a mixed logit model is derived from the multinomial logit model by allowing β_j to be random across i individuals in the severity function:

$$T_{ij} = \beta_{ij} X_{ij} + \varepsilon_{ij} \quad \text{with } \beta_i \sim f(\beta \mid \theta), \tag{2}$$

where β_j is a vector of coefficients to be estimated for outcome j, X_{ij} is a vector of exogenous (or explanatory) variables, θ are the parameters of the distribution of β_{ij} over the population, such as the mean and variance of β_{ij} and ε_{ij} is the error term that is independent and identically distributed (*iid* extreme value property) and does not depend on underlying parameters or data characteristics.

As mentioned, the mixed logit is a generalization of the multinomial structure that allows the parameter vector β_j to vary across each driver or most severely injured occupant. The injury outcome-specific constants and each element of β_j may be either fixed or randomly distributed over all parameters with fixed means, allowing for heterogeneity in effects. A mixing distribution is introduced to the model formulation, resulting in injury severity probabilities as follows (*5*):

$$P_{ij} = \int_x \frac{e^{\beta_i X_{ij}}}{\sum_x e^{\beta_i X_{ix}}} f\left(\beta \mid \varphi\right) d\beta \tag{3}$$

where $f(\beta\ \varphi)$ is a density function of β and φ is a vector of parameters which describe the density function, with all other terms conforming to previous definitions (*30*). The injury severity outcome probability is then simply a mixture of logits (*5*). The distribution is flexible in that β can also be fixed, and when all parameters are fixed the model reduces to the standard MNL formulation. In those instances, where β is allowed to vary, the model is in the open form, and the probability of an observation having a particular outcome can be calculated through integration (*4*).

In this particular case, the parameters vary across the roadway segment population according to a normal distribution (less well-fitting distributions were considered but discarded, such as the log-normal and uniform). Estimation can be done by solving the integral with Monte Carlo simulation. Efficiency has been increased using simulation with Halton draws, an efficient estimation technique for random parameters models (*5, 22*).

Goodness-of-Fit Statistics

The models' performance was evaluated using several well-known statistics: Pseudo R^2 measure $R^2 = 1-(\ln L/\ln L_0)$; the McFadden adjusted-$R^2 = 1 - [(\ln L\text{-}p) / \ln L_0)$; Akaike's information criterion AIC $= -2 \ln L + 2p$; and Bayesian Information Criterion $BIC = -2 \ln L - p \,.\, \ln n$. Where $\ln L$ and $\ln L_0$ are the log likelihood of the fitted and intercept-only models; p is the number of parameters used in each model; and n is the sample size.

Pseudo R^2 coincides with an interpretation of linear model R^2 (*29*). The McFadden adjusted-R^2 statistic was chosen to measure the explanatory power of the models fitted based on the sample data (*27*). AIC and BIC are two measures to evaluate and compare the quality of the models estimated.. AIC and BIC are estimated by considering simultaneously goodness of fit and complexity of the model (*25*). BIC is more appropriate for measuring goodness-of-fit for explanatory power; whilst AIC is more appropriate for measuring predictive accuracy (*28*) and hence predictive power.

DATA

In this study police reported ROR crashes that occurred on Portuguese freeways during a two-year period (2009 and 2010) are analyzed. Data were obtained from the national accident database maintained by the National Road Safety Authority (ANSR) which manages the Portuguese road accidents database, a main source of evidence for this study. However, information on roadside features is lacking in that database. It was thus necessary to collect additional information from the original accident reports. This data was provided by the Guarda Nacional Republicana (GNR), which is a police force responsible for maintaining security and public order as well as protecting and defending the population and their property.

This dataset comprises 580 km of dual carriageway freeway segments situated in various regions across Portugal (Figure 1). All segments have full access control, two lanes per carriageway and paved shoulders (with widths of less than 2.5 m and 4.0 m for left and right shoulders, respectively). Access to and from the freeway is only possible through interchange ramps.

Only single-vehicle ROR crashes involving roadside features were used in this study.

Table 1 shows the variables that proved to be relevant for explaining crash severities and their observed distributions across different severity levels. Information related to ROR crashes including injury severity levels, seasonal attributes (winter, peak hour), roadside attributes (obstacles, barrier, ditch), roadway attributes (right curve), accident information (persons involved, right encroach, rollover, car, speed limit) and driver information (age, gender) was included in the models. The total frequency of crashes in different categories and the proportions of different injury severity levels for each category are also included in Table 1. In the case of continuous variables, the mean and standard deviation parameters are included. ROR crashes with missing information on the accident, driver or vehicle characteristics were removed before the statistical analysis, which resulted in a total of 764 crashes selected, out of 840 registered ROR crashes on dual carriageway roads. Correlation analyses were conducted for all independent variables considered in the study as a first step to identify correlated variables.

FIGURE 1 Google Street View still images of typical Portuguese freeway cross sections.

TABLE 1 Descriptive Statistics of the Significant Variables in the Models

Variable		Description	Total No. of Crashes	PDO (%)	Minor Injury (%)	Severe and Fatal Injury (%)
ROR crashes			764	16.4	76.7	6.9
Categorical variables						
Seasonal attributes						
	Winter	Winter (December, January or February)	225	19.6	74.7	5.8
	Peak hour	Evening period (18.00 to 20.00 pm)	33	12.1	72.7	15.2
Roadside attributes						
	Barrier	Collision with metallic safety barrier as first harmful event	313	19.2	73.5	7.3
	Ditch	Traversing/colliding with ditch as first harmful event	22	18.2	81.8	0.0
Roadway attributes						
	Right curve	Horizontal curve to the right (vs. straight segment or left curve)	94	10.6	85.1	4.3
Accident information						
	Right encroach	Leaving the road to the right side of the carriageway (vs. leaving the road to the left side of the carriageway)	422	13.5	79.1	7.3
	Rollover	Rollover	280	10.4	79.3	10.4
	Car	Passenger car involved	579	18.1	75.6	6.2
Driver information						
	Age	Driver under 32 years old	288	16.3	79.9	3.8
	Female	Gender (female=1)	282	8.9	86.5	4.6
Continuous variables			**Mean**	**sd**	**Min**	**Max**
Roadside attributes						
	Obstacles	Number of obstacles hit in a ROR crash	1.5	0.7	0	4
Roadway attributes						
	Speed limit	Segment speed limit	119.8	2.3	90	120
Accident information						
	Persons involved	Number of involved persons	1.6	0.9	1	7
	Speed limit	Segment speed limit	119.8	2.3	90	120

MODELING RESULTS

Partial proportional odds model

The first step in the development of the model was the examination of the parallel regression assumption to determine if the PPO model is the appropriate ordered-response model to use. As mentioned earlier, in this study, a Wald test was employed to examine if any variable violates the parallel regression assumption. The results of the Wald test demonstrated that only one variable (female) violated this assumption, hence justifying the development of the PPO model. PPO models with both logit and probit functions were fitted with this variable changing across equations while other variables were forced to have their effects meet the parallel-lines assumption. The PPO model with a logit function performed better than that with a probit function (AIC = 940.44 vs. 949.71; pseudo R^2 = 0.126 vs. 0.1171).

Only statistically significant explanatory variables were considered in the final specification model. A minimum confidence level of 85% was considered as criterion. Altogether, 13 parameters were calibrated, through which the potential effects of different factors related to the categories listed above were identified. It is important to point out that most parameters were statistically significant with *p-value* below 5% (i.e., confidence levels above 95%), with three exceptions where *p-values* ranged between 5% and 10% and one case where *p-values* went up to 15%. As previously mentioned, the aim of this paper was to detect unforgiving roadside contributors through a retrospective severity analysis of run-off-road crash data. Therefore, the models are used for explanatory purposes (within the range of values observed, only), where lower *p-values* are acceptable (*27*). The PPO model estimated for different crash injury severity levels is given in Table 2. The estimated PPO model had one beta coefficient for each variable, two gamma coefficients for the variable violating parallel-lines assumption, and three alpha coefficients reflecting the cut-off points. Insignificant parameter estimates are not included in Table 2.

Mixed Logit Model

The coefficients and standard errors for predictors in the mixed logit model developed for different injury severity levels are shown in Table 3.

Minor injury was set as the baseline severity level for the mixed logit model; the Alternative Specific Constant (ASC) was defined accordingly. To improve the numerical stability, the number of Halton draws to evaluate the log-likelihood function was 1000.

Comparison of Models

The same dataset was used to fit the two models, which were PPO and mixed logit, to make a comparison between their performances. The log-likelihood values at convergence, AIC and BIC values were used to compare the performance of the two models used in this study. AIC and BIC are both measures of unexplained variations in the data with a penalty for model complexity. Therefore, models with lower values provide a relatively better fit (*31*). FIGURE 2 shows such comparison based on the AIC and BIC values.

TABLE 2 PPO model for ROR crash injury severities in Portuguese freeways

Variables	Coefficient	S.E.	z
Beta			
Winter	-0.4803	0.1975	0.0150
Peak hour	0.7240	0.4557	0.1121
Obstacles	0.3300	0.1274	0.0096
Ditch	-0.9270	0.5348	0.0830
Persons involved	-0.8220	0.1072	<0.001
Right encroach	0.4175	0.1855	0.0244
Rollover	0.9298	0.1987	<0.001
Car	-0.3612	0.2156	0.0938
Age	-0.4628	0.1877	0.0137
Gamma_1			
Female	1.1045	0.254	<0.001
Gamma_2			
Female	-0.5920	0.3394	0.0811
Alpha			
Constant 1	2.3465	0.3456	<0.001
Constant 2	-1.9681	0.3513	<0.001
Summary statistics			
Number of observations		764	
Log likelihood at convergence		-457.222	
Adjusted-ρ^2		0.101	
Pseudo R^2		0.126	
AIC		940.443	
Bayesian Information Criteria (*BIC*)		1000.745	

Figure 2 shows that the PPO model has the lowest AIC and BIC values in this study, compared to mixed logit. This shows that the PPO model performed slightly better than the mixed logit model and it can be considered as a viable method in ROR crash injury severity modeling. Moreover, the results of the PPO model had plausible signs for all predictors, and the overall model fit was better than that of the mixed logit model. The McFadden's pseudo R-square of 0.126 is good considering the large amount of variance in the injury severity data. Based on the LL, AIC, and BIC, the PPO model provides a better fit than the mixed logit model in analyzing ROR crash injury severity data.

TABLE 3 Mixed Logit Model for ROR Crash Injury Severities in Portuguese Freeways

Severity level	Variables	Coefficient	t-test	p-value
PDO	Constant	6.890	11.06	<0.001
	Winter	0.502	1.85	0.06
	Barrier	0.428	1.67	0.09
	Right encroach	-0.738	-2.84	<0.001
	Car	0.683	2.04	0.04
	Persons involved	1.010	7.06	<0.001
Std. dev. of parameter (*Persons involved*)		0.905	2.37	0.02
	Female	-0.884	-3.38	<0.001
Minor	Night	-0.504	-2.08	0.04
injury	Right curve	0.646	1.66	0.10
	Age	-0.017	-2.08	0.04
Fatal and	Constant	14.600	3.73	<0.001
severe	Speed limit	-0.079	-2.40	0.02
injury	Rollover	0.646	1.59	0.11
	Persons involved	1.010	7.06	<0.001
Number of observations		764		
Log likelihood at convergence		-464.926		
Adjusted-ρ^2		0.429		
Pseudo R^2		0.446		
AIC		955.852		
Bayesian Information Criteria (*BIC*)		1016.153		

NOTE: The attribute *persons involved* was restricted to be equal across PDO and fatal and severe injury severity levels.

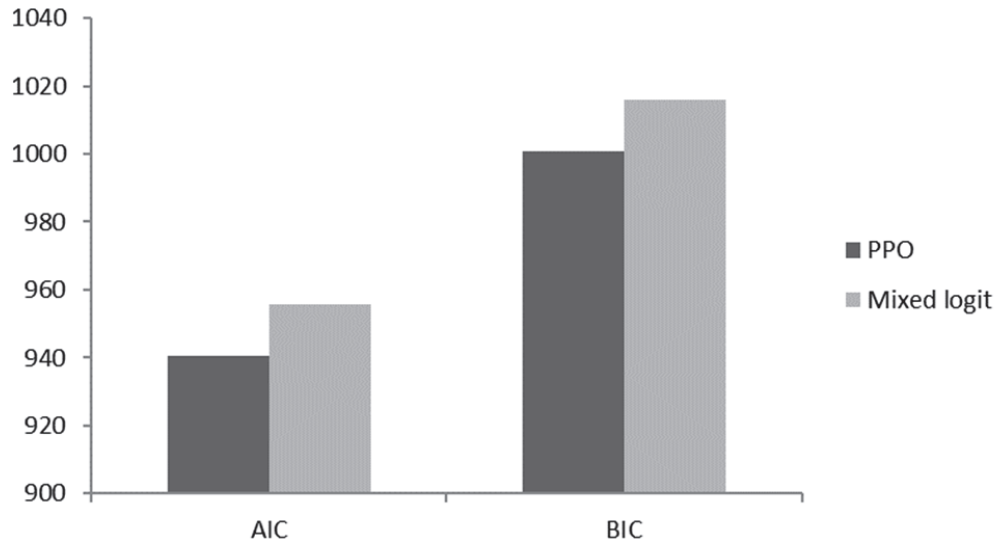

FIGURE 2 Comparison of PPO and mixed logit models

SUMMARY AND DISCUSSION

This paper describes the use of a PPO model to study the role of contributors influencing ROR crash severities on freeways and compared this model with a mixed logit model. The study was based on a two-year detailed infrastructure and accident dataset from Portuguese freeways (2009 and 2010). The crashes were categorized into three different levels based on driver's injury severity. Models were then estimated for the two methodological approaches (PPO and mixed logit models). Marginal effects of the PPO model were also computed to complement the analysis.

In all models, plausible signs were estimated for the coefficients of the variables. The PPO model performed the best out of the two models, based on log-likelihood at convergence, AIC and BIC values.

A PPO model allows the predictors that meet proportional odds assumption to take the same coefficient for all injury severity levels and other predictors to vary between injury severity levels, ensuring no potential loss in accuracy of prediction (*25*). The PPO models are clearly a viable method for modeling ROR crash injury severities.

Several variables in seasonal attributes, roadway and roadside attributes, crash characteristics and driver information were identified as significant predictors influencing the driver injury severity level in ROR crashes. The marginal effects of the parameters for PPO model provide valuable insight on the contribution factors for ROR crash injury severity. Table 4 shows the marginal effects and standard errors reported by the PPO model for different crash injury severity levels. Table 4 shows that, when involved in a ROR crash, the probability of occurrence of an occupant fatality or severe injury is higher for crashes: involving vehicle rollover; with vehicles leaving the road to the right side of the carriageway; and occurring at peak hours. Similarly, the probability of occurrence of severe injuries or fatalities in a ROR crash is lower: during winter; on ditches; with higher occupancy vehicles (higher number of persons involved); if passenger cars are involved; with younger population (below 32); and if female drivers are involved.

TABLE 4 Marginal Effects and Standard Errors for Different ROR Crash Injury Severity Levels

Variables	Crash injury severity					
	PDO		Minor injury		Severe injury + Fatal	
	M.E.	S.E.	M.E.	S.E.	M.E.	S.E.
Winter	0.0564	0.0329	-0.0276	0.0493	-0.0288	0.0219
Peak hour	-0.0850	0.0496	0.0416	0.0743	0.0434	0.0330
Obstacles	-0.0387	0.0226	0.019	0.0339	0.0198	0.0150
Ditch	0.1088	0.0635	-0.0533	0.0951	-0.0555	0.0423
Persons involved	0.0965	0.0563	-0.0472	0.0843	-0.0492	0.0375
Right encroach	-0.0490	0.0286	0.024	0.0428	0.025	0.019
Rollover	-0.1091	0.0637	0.0534	0.0954	0.0557	0.0424
Car	0.0424	0.0247	-0.0207	0.0371	-0.0216	0.0165
Age	0.0543	0.0317	-0.0266	0.0475	-0.0277	0.0211
Female	-0.1296	0.0757	0.1651	0.0632	-0.0355	0.027

These results are in line with several previous findings reported in the literature on ROR crash severity. This is the case for rollover and number of persons involved, which were found to increase the propensity for severe and fatal injury ROR crashes, just as in (6, 10, *14, 32*) for rollover, and (6, 12) for the latter factor.

Female drivers were found to have lower probabilities of PDO, severe and fatal and injury ROR crashes. On the one hand this agrees with findings from Wu et al. (10) and Xie, et al. (*14*); on the other it partially differs from the results obtained by Schneider et al. (9), who found that female drivers are more likely to be injured in ROR crashes.

In addition to rollover and number of persons involved mentioned above, there are several factors for which this study found partly similar findings to those of previous research by the authors (6). These are the role of winter in ROR crashes, the involvement of vehicles leaving the road to the right side of the carriageway, the involvement of passenger cars and driver age.

This study also adds some new insight into the effect that some variables have on ROR crash severities for the case of freeways with "unforgiving" roadsides. Specifically, the use of the partial proportional formulation allows an improved identification of the varying effect that several variables have on ROR crash injury severity, and includes the effect of traversing ditches and the influence of the number of obstacles hit in a ROR crash, which were previously masked, when the unordered framework models were used.

In this study, the number of obstacles hit in a ROR crash was found to decrease the propensity for PDO ROR crashes. This is reasonable, as high kinetic energy may be involved in these crashes, more areas of a vehicle are damaged or more impacts are sustained in the same area of an errant vehicle. This study also shows that traversing a ditch tends to increase the chance of a PDO ROR crash. This appears sensible, as ditches on freeways are not especially "aggressive" (despite not being tolerant to errant vehicles, as well) and they are associated with cut embankments or carriageways leveled with the nearby terrain.

Findings from this type of studies are relevant for setting up preventing measures at the design stage and also in operation management. In the former case, one may expect that applying traversable ditches designs may improve considerably the safety of the road stretches where they are constructed. Additionally, it is more appropriate to improve embankment characteristics, rather than to address a few roadside obstacles. In the latter case, it may be hypothesized that enforcement should be stricter and more intense when cars are expected to carry less passengers (not in holiday periods) and outside of peak hour periods.

The procedure developed in SAFESIDE (described in (3)) does not fully take into consideration the probability of occurrence of crashes with different severity levels conditioned on crash occurrence. By estimating the probability of crash occurrence at different severity levels (using mixed logit models or PPO models), these crash severity models can be integrated in the said procedure, enabling the distribution of the estimated expected number of crashes by three severity levels and thus allowing the development of better crash cost calculations. Cost-benefit estimates tuned to the Portuguese freeway crash characteristics will positively support roadside safety decisions adapted to the country's context and contribute to the progressive construction of an efficient safe traffic system.

ACKNOWLEDGMENTS

The authors would like to thank the *Guarda Nacional Republicana* for its cooperation in providing data on crashes and roadside characteristics used in this research.

REFERENCES

1. Roque, C., Cardoso, J.L. *Definição de Cenários Tipo para Acidentes envolvendo a Área Adjacente à Faixa de Rodagem*, 4º Relatório Safeside. (Run-off-road crash scenario evaluation. 4th Safeside report. – in Portuguese) Fundação para a Ciência e Tecnologia, Laboratório Nacional de Engenharia Civil, Lisboa 2012.
2. Roque, C., Cardoso, J.L. Investigating the relationship between run-off-the-road crash frequency and traffic flow through different functional forms, *Accident Analysis & Prevention 63, 2014, pp. 121-132.*
3. Roque, C., Cardoso, J.L. SAFESIDE: a computer-aided procedure for integrating benefits and costs in roadside safety intervention decision making, *Safety Science 74, 2015, pp. 195-205.*
4. Savolainen, P., Mannering, F., Lord, D., Quddus, M. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. *Accident Analysis and Prevention 43* (*5*)*, 2011, pp. 1666–1676.*
5. Train, K. *Discrete Choice Methods with Simulation*. Cambridge University Press, New York, NY, 2009.
6. Roque, C., Moura, F., Cardoso, J.L. Detecting unforgiving roadside contributors through the severity analysis of ran-off-road crashes. *Accident Analysis & Prevention 80, 2015, pp. 262–273.*
7. Eluru, N., Yasmin, S. A note on generalized ordered outcome models. *Analytic Methods in Accident Research 8, 2015, pp. 1–6.*
8. Kim, J.K., Ulfarsson, G.F, Kim, S., Shankar, V.N. Driver-injury severity in single-vehicle crashes in California: a mixed logit analysis of heterogeneity due to age and gender. *Accident Analysis and Prevention 50, 2013, pp. 1073–1081.*
9. Schneider, W., Savolainen, P., Zimmerman, K. Driver injury severity resulting from single-vehicle crashes along horizontal curves on rural two-lane highways. *Transportation Research Record: Journal of the Transportation Research Board 2102, 2009, pp. 85–92.*
10. Wu, Q., Chen, F., Zhang, G., Liu, X. C., Wang, H., Bogus, S. M. Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways. *Accident Analysis and Prevention 72, 2014, pp. 105–115.*
11. Wu, Q., Zhang, G., Zhu, X., Liu, X.C., Tarefder, R.. Analysis of driver injury severity in single-vehicle crashes on rural and urban roadways. *Accident Analysis and Prevention*, *2016, pp. 94, 35–45.*
12. Holdridge, J., Shankar, V., Ulfarsson, G. The crash severity impacts of fixed roadside objects. *Journal of Safety Research 36* (*2*)*, 2005, pp. 139–147.*
13. Lee, J., Mannering, F. Impact of roadside features on the frequency and severity of run-off-roadway accidents: an empirical analysis. *Accident Analysis and Prevention 34, 2002, pp.149–161.*
14. Xie, Y., Zhao, K., Huynh, N. Analysis of driver injury severity in rural single- vehicle crashes. *Accident Analysis and Prevention 47, 2012, pp. 36–44.*
15. Renski, H.,Khattak, A.,Council, F. Effect of speed limit increases on crash injury severity: analysis of single-vehicle crashes on North Carolina Interstate Highways. *Transportation Research Record 1665, 1999, pp. 100–108.*
16. Kockelman, K.M., Kweon, Y.J. Driver injury severity: an application of ordered probit models. *Accident Analysis and Prevention 34* (*3*)*, 2002, pp. 313–321.*
17. Mooradian, J., Ivan, J.N., Ravishanker, N., Hu, S. Analysis of driver and passenger crash injury severity using partial proportional odds models. *Accident Analysis & Prevention 58, 2013, pp. 53–58.*

18. Mannering, F.L., Bhat, C.R. Analytic methods in accident research: Methodological frontier and future directions. *Analytic Methods in Accident Research 1, 2014, pp. 1–22.*
19. R Development Core Team. R: A Language and Environment for Statistical Computing. *R Foundation for Statistical Computing*, Vienna, Austria, ISBN 3-900051-07-0 http://www.R-project.org/, 2011.
20. Yee, T.W. The VGAM package for categorical data analysis. *Journal of Statistical Software 32* (*10*), *2010, 1–34.*
21. Bierlaire, M. BIOGEME: A free package for the estimation of discrete choice models. *Proceedings of the 3rd Swiss Transportation Research Conference*, Ascona, Switzerland, 2003.
22. Bhat, C. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B 37*(*9*), *2003, pp. 837–855.*
23. Wang, X., Abdel-Aty, M. Analysis of left-turn crash injury severity by conflicting pattern using partial proportional odds models. *Accident Analysis and Prevention 40* (*8*), *2008, pp. 1674–1682.*
24. Pour-Rouholamina, M., Zhouba, H. Analysis of driver injury severity in wrong-way driving crashes on controlled-access highways. *Accident Analysis and Prevention 94, 2016, pp. 80–88.*
25. Sasidharan, L., Menendez, M. Partial proportional odds model – an alternate choice for analysing pedestrian crash injury severities. *Accident Analysis and Prevention 72, 2014, pp. 330–340.*
26. Williams, R. Generalized ordered logit/partial proportional odds models for ordinal dependent variables. *Stata J. 6* (*1*), *2006, pp. 58–82.*
27. Washington, S., Karlaftis, M., Mannering, F.L. *Statistical and Econometric Methods for Transportation Data Analysis*. Second Edition. Chapman and Hall/CRC. ISBN: 978-1-4200-8285-2.0, 2011.
28. Shmueli, G. To explain or to predict? *Statistical Science 25* (*3*), *2010, pp. 289–310.*
29. Cameron, A., Trivedi, P. *Regression Analysis of Count Data*. Cambridge University Press, Cambridge, England, 1998.
30. Milton, J.C., Shankar, V.N., Mannering, F.L. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis & Prevention 40* (*1*), *2008, pp. 260–266.*
31. Finch, W.H., Bolin, J.E., Kelley, K. *Multilevel Modeling Using R*. CRC Press, Florida, USA, 2014.
32. Hu, W., Donnell, E. T. Median barrier crash severity: some new insights. *Accident Analysis and Prevention 42, 2010, pp. 1697–1704.*