WILEY | Hindawi

Research Article

Investigation into Interactions between Accident Consequences and Traffic Signs: A Bayesian Bivariate Tobit Quantile Regression Approach

Xuecai Xu [1,2] **and Željko Šarić** [3]

[1] School of Civil Engineering and Mechanics, Huazhong University of Science and Technology, Wuhan, China
[2] School of Civil and Environmental Engineering, Nanyang Technological University, Singapore
[3] Department of Traffic Accident Expertise, Faculty of Transport and Traffic Sciences, University of Zagreb, Croatia

Correspondence should be addressed to Xuecai Xu; xuecai_xu@hust.edu.cn

This study intended to investigate the interactions between accident severity levels and traffic signs in state roads located in Croatia and explore the correlation between accident severity levels and heterogeneity attributed to unobserved factors. The data from 460 state roads between 2012 and 2016 were collected from Traffic Accident Database System maintained by the Republic of Croatia Ministry of the Interior. To address the correlation and heterogeneity, Bayesian bivariate Tobit quantile regression models were proposed, in which the bivariate framework addressed the correlation of residuals with Bayesian approach, while the Tobit quantile regression model accommodated the heterogeneity due to unobserved factors. By comparing the Bayesian bivariate Tobit quantile and mean regression models, the proposed quantile models showed priority to mean model. Results revealed that (1) low visibility and the number of invalid traffic signs per km increased the accident rate of material damage, death, or injury; (2) average speed limit exhibited a close relation with accident rate; and (3) the number of mandatory signs was more likely to reduce the accident rate of material damage, while the number of warning signs was significant for accident rate of death or injury.

1. Introduction

With the development of economy and society, a variety of roadside signs, e.g., advertising signs, neon lights, and gaudy billboards, are emerging and dominating the visual landscape in many urban and suburban areas, which makes the roadside environment complicated. Although there are some studies focusing on the impact of such development, limited studies have investigated the relationship between the signs and traffic safety in Croatia. Among these, traffic signs, primary means of communication between the road authorities and traffic participants, play an important role as a principal causative factor by affecting drivers' reaction time and road safety.

Currently a number of studies concentrate on the traffic sign detection, e.g., the color, the size, and lettering of road signs [1, 2], and although there are some studies [3–5]

attempting to establish the relationship between accidents and traffic signs, merely one type of specific traffic signs, e.g., warning signs and speed limit signs, was investigated, which might not completely reflect the effectiveness of traffic signs and safety problem due to different categories of traffic signs.

During the last decade a variety of different approaches and perspectives [6–8] have been presented in safety evaluation. Studies have indicated that heterogeneity issue can be addressed through finite mixture regression models [9] and random parameter model [10, 11], in which the heterogeneity from the data or locations caused by unobserved factors was accommodated, and the estimation results and the statistical inferences were improved. However, the aforementioned models belong to mean regression, in which the model assumptions cannot be easily extended to noncentral locations and do not always complement the nature real-world data, especially in the case of homoscedasticity [12]. A more

appropriate and more complete view is required to capture the distributional properties with a broader spectrum than only mean and variance.

In recent years, quantile regression (QR) has attracted increasing attention in various fields, e.g., sociology, economics, finance, and medical science [12–14], whereas the application of quantile regression in the transportation field is still at the initial stage [12, 15–18]. The main merit of quantile regression is to offer a more complete view and a highly comprehensive analysis of the relationship between variables from a broad spectrum. Compared with the mean regression, quantile regression does not require the data to follow a specific distribution but estimate multiple variations from several regression curves for different percentage points of the distribution, which may reflect different effects at different quantiles of the response variable. Moreover, quantile regression is more robust against outliers because the estimation results may be less sensitive to outliers and multimodality [19]. In particular, quantile regression can handle the heterogeneity issue for the data collected from different sources at different locations and different times without many assumptions [12, 15, 16], which is helpful to describe the unobserved factors at different roads more clearly.

However, when the dependent variables take on more than one type of results, the conventional QR may be inadequate. To overcome the drawbacks, a Bayesian approach to bivariate/multivariate quantile regression can be used to handle it appropriately. The main feature of Bayesian approach is concerned with generating the posterior distribution of the parameters and provides a more complete view of the uncertainty in the estimation of unknown parameters, especially after the confounding effects of nuisance parameters are removed. Yu and Moyeed [20], Waldmann and Kneib [21], and Xu et al. [22] have demonstrated that this method is very competitive in analyzing crash frequency/severity involving spatial correlation or heterogeneity issue. Recently, the extension has been expanded to Bayesian network and its transferability.

In order to capture the impact of covariates on two types of censored accidents simultaneously, a new Bayesian bivariate Tobit QR is proposed in this study. Therefore, the purpose of this study is to gain insight into the relationship between accident consequences and a series of traffic signs systematically located at state roads using Bayesian bivariate Tobit QR models, in which the accident consequences can be estimated appropriately, while addressing the heterogeneity at different roads.

2. Literature Review

There have been a variety of studies about the relationship between accidents and traffic signs during the last thirty years, but they are investigated from different perspectives. Initially Loo [23] discussed the role of primary personality factors in the perception of traffic signs, driver violations, and accidents. The findings indicated that the fast decision-time component of personality factors carried the relationships with accidents and the ability to perceive embedded traffic signs in the verbal message and symbolic message types. As

for the warning signs, Jørgensen and Wentzel-Larsen [24] optimized the use of warning signs in traffic. It was concluded that warning signs would increase safety and have a greater positive impact on total driving costs than on accident costs. Similarly, Carson and Mannering [3] investigated the effect of ice warning signs on ice accident frequencies and severities. It was found that an ice warning sign would be one way of understanding the impact of warning signs, and ice warning signs were more likely to be placed at locations with high numbers of ice-related accidents. From the freeway perspective, Wu and Wang [5] explored the impact of speed limit signs on freeway work zones. The speed limit model was established by considering the relevant factors of driver's field of vision and transportation distance of visual cognition. The results showed the proposed speed limit model within warning area played an important role for freeway traffic safety, while Xuan and Kanafani [25] evaluated the effectiveness of accident information on freeway changeable message signs (CMS) by comparing different aggregate analysis methodologies. The findings showed that CMS accident messages had no significant effect on driver diversion but visible congestion is an important factor. On the other side, if too many signs are installed, the effect may be opposite. Strawderman et al. [26] investigated the effect of signs on driver behavior and accident frequency in school zones. It was found that sign saturation had a significant effect on vehicle speed, compliance, and accident frequency. All of the literatures above have not indeed built up the actual relationship between and traffic signs except the latest study by Babić et al. [27, 28]. They analyzed traffic signs in terms of traffic safety in Croatia and found out that improper installation and maintenance of traffic signs can affect the traffic safety, which is the base of this study.

The quantile regression (QR) is introduced to deal with the relationship between the quantile locations and covariates. QR was initially proposed by Koenker and Bassett [29] to specify conditional quantiles as functions of predictors. In a pioneering study, Hewson [30] examined the potential role of QR in modeling the speed data and demonstrated the potential benefits of using QR methods, providing more interest than the conditional mean. From the perspective of discrete variables, Qin et al. [15] identified crash-prone locations with quantile regression. The flexibility of estimating trends at different quantiles was offered, and the data with heterogeneity were tackled. The findings suggest that QR yields a sensible and much more refined subset of risk-prone locations. Following that, Qin and Reyes [16] and Qin [12] modeled crash frequencies with quantile regression. QR tackles heterogeneous crash data and offers a complete view of how the covariates affect the responsible variable from the full range of the distribution, which is beneficial for data with heavy tails, heteroscedasticity, and multimodality. The results illustrate that QR estimates are more informative than conditional means. Similarly, Wu et al. [17] analyzed crash data using quantile regression for counts. The results revealed more detailed information on the marginal effect of covariates change across the conditional distribution of the response variable. They also provided more robust and accurate predictions on crash counts. After that, Liu et al.

[19] extended to the train derailment severity using zero-truncated negative binomial regression and QR and provided insights into train derailment severity under various operational conditions and by different accident causes. To deal with the drinking and driving fatality issue, Ying et al. [31] explored the impact of drinking and driving laws on varying severity rates by employing QR. The empirical results showed that QR reflects more detailed information, and different fatality rates (low quantiles and high quantiles) depend on the specific conditions in various regions. From the perspective of identifying accident blackspots in a transportation network, Washington et al. [18] applied QR to model equivalent property damage only (PDO) crashes. The proposed method identified covariate effects on various quantiles of the population and performed better than traditional negative binomial model.

However, QR models above all include one dependent variable, and when the dependent variables take on more than one type of results, the bivariate/multivariate QR models may be adequate [32, 33], and if the dependent variables are latent, bivariate/multivariate Tobit QR models can be used to address them [34]. Furthermore, for all of models abovementioned, the sample $X_1,........,X_n$ is extracted from a population with an unknown but fixed parameter θ, which is obtained from the observed random sample, but for the Bayesian approach, which is fundamentally different, θ is regarded as a random variable and its variation can be described by a probability distribution. The main feature of Bayesian approach is concerned with generating the posterior distribution of the parameters and provides a more complete view of the uncertainty in the estimation of unknown parameters, especially after the confounding effects of nuisance parameters are removed [35]. A detailed introduction to Bayesian analysis can be found in Leung and Yu [36] and Mokatrin [35].

To take advantage of the Bayesian approach and multivariate QR modeling, the Bayesian bivariate Tobit QR models were proposed integrating the latent QR model with multivariate modeling into the Bayesian framework so as to accommodate the issues above.

3. Data Description

The dataset integrated the Traffic Accident Database System from 2012 to 2016 with the traffic signs maintained by the Republic of Croatia Ministry of the Interior. Four main components from the Traffic Accident Database System were included: the accident consequence, the traffic signs profiles, the accident environment, and roadway characteristics.

About 135 state roads were elaborately selected from Croatia in five years. As shown in Figure 1, totally there were 665 state roads, but after removing some roads without data and processing, 460 valid state roads were determined for the analysis. In Croatia, the consequences of accidents are typically categorized as death, injury, or material damage. In our sample, the death cases only accounted for 2.1%. Given that the two adjacent accident categories were quite similar, merging the death and injury categories was not expected to substantially affect the inference. Consequently,

FIGURE 1: Selected state road in Croatia.

the dependent variables in the proposed model were bivariate accident severity levels in which the response of interest referred to death and injury (DI), and material damage (MD) was treated as the contrast. As required by the Tobit QR model, dependent variables should be continuous; thus accident rate (expressed as accidents per million vehicles miles traveled (MVMT)) is introduced to represent MD and DI. The way that accident rate is considered is because it incorporates the effect of volume and road length; it is more adequate to measure the accident risk faced and perceived by individual drivers compared with accident frequency, which is highly related to the traffic volume. Thus, the variables road length, traffic volume, and the number of accidents cannot be included in the independent variables to avoid the estimation bias.

By aggregating the traffic signs and accident environment profiles, the explanatory variables reflecting the traffic signs including classification, types, the number of different traffic signs, and the proportion of all traffic signs, the environmental factors (i.e., visibility, weather conditions), the roadway and traffic characteristics (i.e., road type, speed limit, road length, and annual average daily traffic (AADT)), and demographic characteristics of passengers (i.e., average, minimum and maximum age) were extracted. To be consistent with other aggregated variables, speed limit in this study was averaged by segment length; thus average speed limit was employed to represent the mobility.

Classification of traffic signs in terms of their function and retroreflective material is the major assortment. In Croatian regulations, traffic signs are, by their function, classified into the following: A-warning signs, B-mandatory signs, C-information signs, D-directional signs, E-additional panels and K-traffic equipment, while in Europe retroreflective materials are divided into materials of Classes I, II, and III. Materials of Classes I and II use spherical or prismatic retroreflective sheeting, while materials of Class III use exclusively prismatic retroreflective sheeting [27, 28]. Both classifications with valid and invalid numbers and proportions were collected for the five years. Figure 2 shows one segment of state road D1 with valid and invalid signs. The blue marks

TABLE 1: Summary of the parameters.

Factor	Attribute	Count (proportion)	
State roads	2012	131(19.7%)	
	2013	131(19.7%)	
	2014	131(19.7%)	
	2015	131(19.7%)	
	2016	141(21.2%)	
Accident consequence	Death or injured	8613(40.7%)	
	Material damage	12562(59.3%)	
Traffic signs	Class I	381540(65.3%)	
	Class II	161079(27.5%)	
	Class III	41664(7.2%)	
Environmental factors		Accidents (proportion)	
Visibility	Day	13513(67.5%)	
	Low visibility for MD	4063(20.3%)	
	Low visibility for DI	2439(12.2%)	
Weather	Clear sky	12902(64.5%)	
	Cloudy	4043(20.2%)	
	Bad weather (rain/fog/snow/other)	3039(15.3%)	
	Range	Mean	S.D.
Accident rate			
Material damage (MD)	Min: 0.022; Max: 19.56	0.78	1.22
Death or injured (DI)	Min: 0.017; Max: 7.95	0.56	0.80
Traffic Signs			
Number of warning signs	Min: 6; Max: 2342	262.47	32.86
Number of mandatory signs	Min: 10; Max: 4457	367.39	57.06
Number of information signs	Min: 4; Max: 4224	40.66	56.50
Number of directional signs	Min: 2; Max: 973	73.12	11.63
Number of additional panels	Min: 0; Max: 942	87.73	13.18
Amount of traffic equipment	Min: 0; Max: 2339	230.83	30.57
Number of traffic signs in Class I	Min: 27; Max: 10436	932.86	133.95
Number of traffic signs in Class II	Min: 3; Max: 4030	393.84	55.93
Number of traffic signs in Class III	Min: 0; Max: 890	101.87	12.73
Number of valid traffic signs per km	Min: 5.39; Max: 72.23	18.32	8.02
Number of invalid traffic signs per km	Min: 0.28; Max: 15.97	4.31	2.26
Traffic characteristics			
Speed limit, km/hr	Min: 28; Max: 98	62.54	10.38
Road length, km	Min: 1.5; Max: 644.6	66.73	8.92
AADT	Min: 229; Max: 14637	3546.61	227.99
Driver's characteristics			
Average age for MD	Min: 20; Max: 68	42.3	6.12
Average age for DI	Min: 19; Max: 65	40.7	6.69

represent technically correct signs while the red marks are technically defective signs.

The environmental factors collected include visibility and weather conditions. In terms of visibility degree, the number of MD and DI was categorized for high visibility (i.e., daytime) and low visibility (including night, twilight, and dawn). Similarly, the atmospheric conditions were collected according to the weather variation, and three types were classified, clear sky, cloudy, and bad weather conditions (including rain, fog, snow, and salty road).

The variables used for model development are displayed in Table 1, with the proportions of the categorical variables in upper part before and the descriptive statistics of the continuous variables in the following.

4. Methodology

4.1. Bayesian Tobit Quantile Regression (TQR). The Tobit model was presented by James Tobin firstly [37], which is referred to for addressing the range of dependent variable in
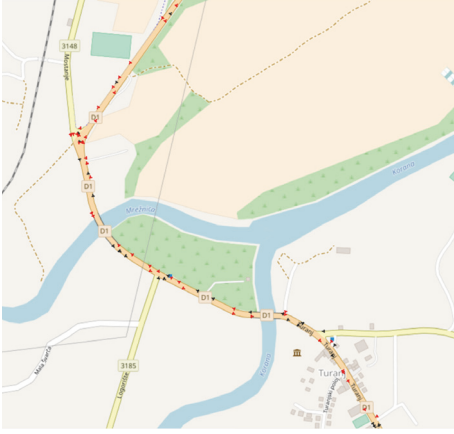
FIGURE 2: State road D1 with traffic signs.

regression model censored in some way. Censoring stands for a data limitation that can result in a data clustering at a lower threshold (left-censored) or upper threshold (right-censored) or both. Censored data differ from truncated data because the latter provide only unlimited values, whereas censored data also provide limited data information. As for the accident rate, the data can be considered as left-censored at zero (zero crash per million vehicle miles traveled) since not all the state roads experienced accidents during observation period. Thus, the Tobit model could be described as outlined:

$$y_i^* = x_i'\beta + \varepsilon_i, \quad i = 1, \ldots n, \ \varepsilon_i \stackrel{iid}{\sim} N\left(0, \sigma^2\right) \qquad (1)$$

where x_i is a vector of explanatory variables, β is a vector of regression coefficients, ε_i follows an identically independent normal distribution with mean zero and variance σ^2. The observed accident rate is assumed to be related to the latent value by the following:

$$y_i = \begin{cases} y_i^*, & if \ \ y_i^* > 0, \\ 0, & otherwise \end{cases} \qquad (2)$$

Then, consider the p-th quantile regression model for y_i^*:

$$y_i^* = x_i'\beta_p + \varepsilon_i \qquad (3)$$

The p-th quantile regression model of y_i^* can be described as

$$Q_{y_i^*|x_i,\beta_p}\left(p\right) = x_i\beta_p \qquad (4)$$

where $Q_{y_i^*|x_i,\beta_p}(\cdot)$ represents the conditional quantile function y_i^*. If p=0.5, $Q_y(0.5)$ is the conditional median, the value that splits the conditional distribution of the outcome variable into two parts with equal probability.

The TQR estimates $\widehat{\beta}_p$ can be expressed as the following problem:

$$\widehat{\beta}_p = \arg\min \sum_{i=1}^{n} \rho_p\left(y_i^* - x_i\beta_p\right) \qquad (5)$$

where $\rho_p(x) = x(p - I(x < 0))$ and where $I(\cdot)$ denotes the indicator function. For any $p \in (0,1)$ the loss function ρ_p assigns a weight of p to positive residuals and a weight of (*1-p*) to negative residuals. It can be found out that the p-th regression quantile coincides with the maximum likelihood estimate under independent asymmetric Laplace distribution (ALD) for the unobserved error terms, which is needed for the specification of the likelihood in the Bayesian framework.

To implement the Bayesian inference, the three-parameter ALD with a skewness parameter is employed to model the quantile of interest as Yu and Zhang (2005):

$$f\left(x \mid \mu, \sigma, p\right) = \frac{p\left(1 - p\right)}{\sigma} \exp\left\{-\rho_p\left(\frac{x - \mu}{\sigma}\right)\right\} \qquad (6)$$

where ρ_p is the same function in (4), μ denotes the location, and $\delta = 1/\sigma$ denotes the precision. Minimizing (4) is equivalent to maximizing the regression likelihood of (5) utilizing ALD errors with $\mu = x_i\beta_p$.

By forming a ALD with $\mu = x_i\beta_p$, specifying the quantile of interest p, and putting priors on the model parameters β and σ, the resulting posterior distribution can be expressed as follows:

$$\psi\left(\beta, \sigma \mid y, x, p\right) \propto \pi\left(\beta, \sigma\right) \prod_{i=1}^{n} ALD\left(y_i \mid x_i\beta, \sigma, p\right) \qquad (7)$$

where $\pi(\beta, \sigma)$ is the joint prior on the regression parameters. The inference about the model parameters follows Bayesian procedures and the model estimation can be referred to in Yue and Hong [33].

4.2. Bayesian Bivariate Tobit Quantile Regression. In order to capture the impact of covariates on two dependent variables simultaneously, the Bayesian TQR that relies on a correlated ALD is extended to the bivariate case; i.e., the vector y_i with independent entries y_{i1} and y_{i2} is used to model the regression at different quantile levels p_1 and p_2. Therefore, the bivariate Gaussian distribution is employed to express the model:

$$y_i \mid \eta_i, \delta_1^2, \delta_2^2, w_i, p$$

$$\sim MVN\left(\begin{pmatrix} \eta_{i1} + \xi_1 w_{i1} \\ \eta_{i2} + \xi_2 w_{i2} \end{pmatrix}, \begin{pmatrix} \dfrac{\sigma_1^2 w_{i1}}{\delta_1^2} & 0 \\ 0 & \dfrac{\sigma_2^2 w_{i2}}{\delta_2^2} \end{pmatrix}\right) \qquad (8)$$

where η_i is the predictor, $w_i = (w_{i1}, w_{i2})'$ are weights, which follow an exponential distribution with rate δ^2 (i.e., the precision of the ALD), and $\xi_k = (1 - 2p_k)/p_k(1 - p_k)$, $\sigma_k^2 = 2/p_k(1 - p_k)$ for k=(1,2). If w_{ik} is assumed to be distributed exponentially with rate δ_k^2, the marginal distributions for y_{ik} are similar to those from previous sections; i.e., the estimators of the maximized likelihood are the minimizers of the function ρ_p.

In order to compute the regression parameters, a full Bayesian inference using the MCMC method is implemented to construct the previous model. Specifically, the variability of

accident rate of MD and DI is considered, which may explain some of between-roads variability. Noninformative priors are assigned for model parameters:

$$\beta \sim Normal\,(0, 1000)$$

$$\mu_1 \sim Normal\left(0, \sigma_1^2\right) \qquad (9)$$

$$\mu_2 \sim Normal\left(0, \sigma_2^2\right)$$

where μ_1 and μ_2 denote the random intercepts at the accident rate of MD and DI equations, respectively, and both variance components σ_1^2 and σ_2^2 are assigned inverse-gamma prior distributions with scale and shape parameters of 0.01, i.e.,

$$\sigma_1^2 \sim igamma\,(0.01, 0.01)$$
$$\qquad (10)$$
$$\sigma_2^2 \sim igamma\,(0.01, 0.01)$$

Different from 95% confidence interval of maximum likelihood estimation, the results present Bayesian credible interval (BCI) as a probability statement about the parameter itself; i.e., a 95% BCI contains the true parameter value with ~95% certainty. If the 95% BCI of the posterior mean does not include 0, it implies that this effect is statistically significant at the 95% level. In this study, all estimation and computations are performed with software STATA 15. More relevant estimation details can be referred to in Alhusseini and Georgescu [38].

For model comparison, as provided by many other studies under the Bayesian framework, the Deviance Information Criterion (DIC) is used to compare the models mentioned above:

$$DIC = D\left(\overline{\theta}\right) + 2p_D = \overline{D} + p_D \qquad (11)$$

where $D(\overline{\theta})$ is the deviance evaluated at $\overline{\theta}$, the posterior mean of the parameter of interest, p_D is the effective number of parameters in the model, and $\overline{D}$ is the posterior mean of the deviance statistic $D(\overline{\theta})$. The lower the DIC, the better the model fits. Generally speaking, differences in DIC of more than 10 definitely rule out the model with the higher DIC; differences between 5 and 10 are considered substantial, while the difference less than 5 indicate that the models are not statistically different.

5. Results

Before the proposed model was run, the correlation test was conducted, which shows visibility in daytime is highly related to low visibility for MD and DI, clear sky, cloudy and bad weather conditions, and number of traffic signs in Classes I, II and III, while the low visibility for MD and DI is highly related to clear sky as well as cloudy and bad weather conditions; thus they are not adopted at the same time as the independent variables. The total number of traffic signs here was equivalent to the sum of the following three categories: warning sign, mandatory signs, information signs, directional signs, additional panels, and traffic equipment in

terms of function; signs in Classes I, II, and III in terms of retroreflective material; and valid and invalid signs; thus there existed interactions among the three categories because they belong to each other; e.g., functional signs and retroreflective signs include both valid and invalid ones. Therefore, the first category was selected originally, while the proportions of signs in Classes I, II, and III were considered (but not significant), and valid/invalid signs were introduced as the number of valid/invalid signs per km so as to avoid the correlation. Other variables, such as number of directional signs, number of information signs, number of additional panels, and average age of drivers, did not show up in the results because they are not significant for the crash rate of MD and DI.

As stated in the modeling, Bayesian quantile estimation method is employed to evaluate the relationship between crash rate and traffic signs, and confidence intervals are calculated for each estimated mean. Table 2 gives the estimated results for Bayesian bivariate quantile and bivariate regression models and 95% confidence intervals for statistically significant variables at the 25th, 50th, 75th, 90th, and 95th percentile of crash rate distribution. It can be found that a broader and complete view of the variables with different crash rates is revealed; that is to say, rather than assuming the coefficients are fixed across all the arterials, some or all of them are allowed to vary to account for heterogeneity attributed to unobserved factors.

In order to demonstrate the proposed model, the Bayesian bivariate Tobit mean regression model was performed. DIC is used to compare the results in Table 2. Obviously DIC values of the proposed model are smaller, especially at 0.75 quantile, than that of Bayesian bivariate Tobit regression model, which indicates that the proposed model is better than the Bayesian bivariate Tobit regression model, according to the DIC rule: the smaller the better; and same trend occurs in variances σ_1 and σ_2.

As shown from Table 2, low visibility for MD and DI, average speed limit, number of warning signs and mandatory signs, and number of invalid traffic signs per km are significant factors influencing accident rate. Moreover, average speed limit is not significant from quantile 0.90 for both models. However, the closer examination of the magnitude of the estimated coefficients reveals some similarities and differences between quantiles. First, the similarity is that all the influencing variables are of significance before quantile 0.75, but not beyond quantile 0.75; that is to say, the impact of all the variables may not be even, and some are more likely to influence crash rate in the low tails and others in the high tails. This may be due to different accident consequences from different influencing factors. This indicates that certain influencing factors would lead to specific accident consequences and need to be considered separately. Secondly, the difference is that significant variables may reveal different impacts on crash rate at different percentiles; e.g., the low visibility is significant all through the quantiles, while average speed limit is only significant at certain quantiles. This indicates that, from the variables considered in the model, low visibility may be the most important variable influencing crash rate; thus

Table 2: Estimation results for Bayesian bivariate Tobit quantile/mean regression models.

Variable	Quantile					Mean Regression
	0.25	0.50	0.75	0.90	0.95	
Accident Rate of MD						
Low visibility	0.013*	0.020*	0.026*	0.081*	0.188*	0.024*
	(5.51)	(7.20)	(3.78)	(2.46)	(6.89)	(4.32)
Average speed limit	-0.005*	-0.007*	-0.016*	-0.019	-0.037	-0.009*
	(-2.71)	(-3.51)	(-3.08)	(-0.78)	(-1.81)	(-1.72)
No. of mandatory signs	-0.002*	-0.004*	-0.004*	-0.009*	-0.010*	-0.001*
	(-5.18)	(-7.17)	(-3.23)	(-1.96)	(-1.99)	(-5.16)
No. of invalid traffic signs per km	0.006*	0.015*	0.028*	0.065*	0.048	0.155*
	(2.60)	(5.49)	(4.28)	(2.06)	(1.87)	(6.18)
Constant	0.586*	0.832*	1.397*	1.474	3.215	0.808*
	(4.66)	(5.67)	(3.96)	(0.86)	(2.28)	(2.16)
Accident Rate of DI						
Low visibility	0.007*	0.015*	0.032*	0.093*	0.183*	0.019*
	(4.15)	(5.48)	(5.62)	(2.67)	(6.95)	(4.88)
Average speed limit	-0.004*	-0.007*	-0.008*	-0.020	-0.032	-0.012*
	(-3.11)	(-3.26)	(-1.98)	(-1.79)	(-1.56)	(-3.26)
No. of warning signs	-0.001*	-0.001*	-0.001*	-0.003*	-0.003*	-0.001*
	(-3.45)	(-4.30)	(-3.55)	(-3.15)	(-2.55)	(-5.48)
No. of invalid traffic signs per km	0.005*	0.011*	0.026*	0.072*	0.087*	0.102*
	(3.41)	(4.02)	(4.77)	(3.72)	(3.35)	(6.31)
Constant	0.385*	0.665*	0.721*	1.385*	2.332	0.963*
	(4.29)	(4.49)	(2.39)	(1.96)	(1.66)	(3.99)
Goodness-of-fit						
No. of Observations		460				
σ_1	1.123	0.982	0.973	0.984	1.128	1.137*
σ_2	0.724	0.652	0.638	0.658	0.730	0.731*
Correlation	0.758	0.766	0.791	0.764	0.753	0.765
DIC	1783.452	1635.683	1583.364	1642.538	1827.693	1858.634

Note. *Significant at the 5% level.

more attention should be paid to it by roadway planners or management departments.

Figure 1 illustrates the estimation results for all the variables except the number of warning signs and mandatory signs because the means are kept relatively constant. The solid line denotes the mean for the 0.25, 0.50, 0.75, and 0.90 quantiles, which are enveloped by two dashed lines representing a 95% Bayesian credible interval.

As shown from Table 2, low visibility for MD and DI is significantly and positively related to the crash rate. As demonstrated in Figures 3(a) and 3(b), although the impact of low visibility for MD and DI on crash rate is similar and smooth before quantile 0.5, the trend is still going up, but after quantile 0.5, especially after quantile 0.75, the rising trend is becoming larger, and the impact of low visibility for DI is higher than that for MD. This indicates that the low visibility has different impact on different accident consequences; i.e., low visibility has almost the same effect on MD and DI at low tails whereas it has more effect on DI at high tails. The transition implies that if visibility is weak, both MD and DI may occur, but if visibility is worse, it may turn to DI

accidents. Therefore, visibility should be taken care of when the state road safety is considered.

Average speed limit is associated with crash rate at all quantiles from 0.25 to 0.75 in Table 2, and the trend of effect increases, implying that the increase of accident rate at higher quantile is larger than that at lower quantile. This indicates that the higher the average speed limit, the larger the chances of the accident occurrence.

Since one purpose of this study is to find out the relationship between accident consequences and traffic signs, the results reveal that the numbers of warning signs and mandatory signs are both negatively concerned with accident rate. The negative relation represents that more warning signs and mandatory signs would reduce the accident rate, which makes sense. These two types of signs are of importance because the mandatory signs are very necessary for the road users, while the warning signs are usually installed at dangerous locations, thus easily causing accidents.

The number of invalid traffic signs per km is the significant factor for both MD and DI till quantile 0.90 and has positive relation with accident rate from Table 2, but there
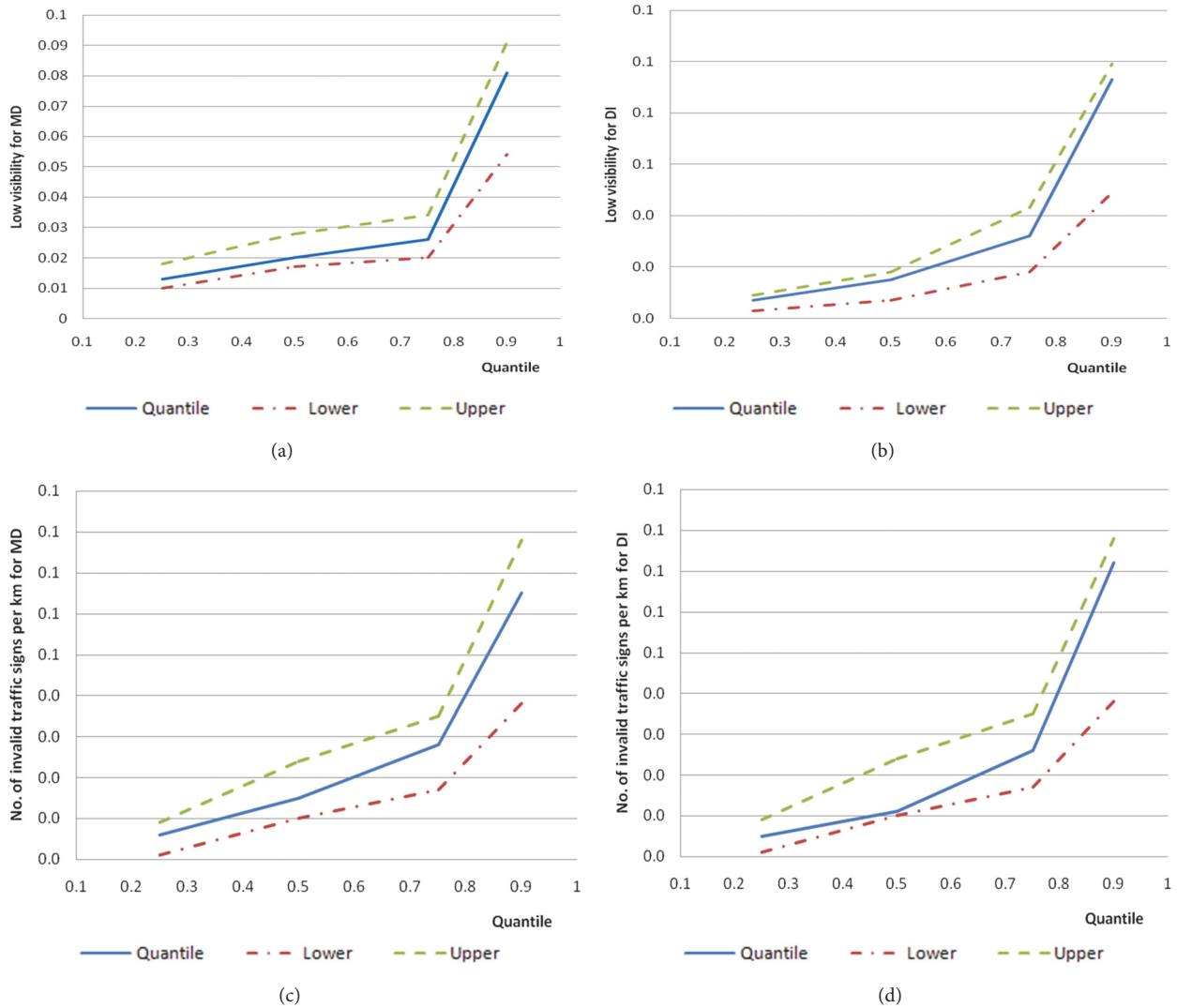
(a)

(b)

(c)

(d)

Figure 3: Quantile plots for variable means.

exists a clear turning point at quantile 0.75 in Figures 3(c) and 3(d). For all types of traffic signs listed above, every year there may be various reasons causing the signs to be invalid, like the sun light, rain washing, erosion, etc., the more invalid traffic signs per km, and the higher chances for drivers to run into accidents because not all the drivers are familiar with the roadways in the whole country, and appropriate instructions at certain locations are definitely necessary. The turning point at quantile 0.75 illustrates that the effect reaches its peak till quantile 0.9, and after that the impact starts to go down.

6. Conclusions

This study investigated the interactions between accident rate and traffic signs in state roads in Croatia through the analysis of accident consequences, a series of traffic signs, traffic characteristics, environmental factors, and driver's characteristics, and the Bayesian bivariate Tobit quantile regression models were proposed.

Some key findings are achieved from this analysis. To our knowledge, it is the first attempt to analyze the interactions between accident rate and traffic signs by employing Bayesian bivariate Tobit quantile regression models, in which the Tobit quantile regression model offers a more complete view and a highly comprehensive analysis of the relationship between accident rate and traffic signs and accommodates the heterogeneity attributed to unobserved factors, while bivariate framework addresses the correlation between accident severity levels with Bayesian approach; the dataset collected from states roads, Croatia, from 2012 to 2016 is employed to illustrate the proposed models; by involving the visibility, the present study demonstrates that the low visibility causes a relatively higher risk of MD and DI. It is noteworthy that average speed limit corresponds to accident rate. The number of mandatory signs and the number of warning signs are more likely to reduce the accident rate. The number of invalid traffic signs per km are important for accident rate; thus regular maintenance should be kept for a safer roadway environment.

Some weakness still exists in this study. Although the proposed model is concerned with the heterogeneity issue, the role of spatial correlation of severity levels in different state roads has not been strongly addressed. Thus, in the future the spatial correlation, such as Bayesian hierarchical approach, could be attempted. Besides, since the results of the study were based on the dataset from Croatia, it is worthwhile to try out different data sources to confirm the findings and transferability of this study in future studies.

In practice, according to our results, in order to reduce the accident rate, from the environmental aspect it is suggested that the streetlight be adjusted during certain time periods, especially at dawn, twilight, and night, because during those periods the visibility is weak, and the drivers' eyesight is ambiguous, thus leading to the accidents easily. What is more, because the range covered by each streetlamp is limited, the streetlamps should be installed at appropriate distances so as to increase the visibility of the roadways all along and keep the light continuous, hence decreasing the collision risk. As for the state roadways in Croatia, the average speed limit plays an important role in determining the accident severity levels, so before the speed limit signs are installed, it should be decided carefully after a variety of investigations according to traffic volume, roadway characteristics, and other influencing factors. Since the number of mandatory signs and warning signs are significant to accident rate, the two types of traffic signs should be installed more at appropriate intervals correspondingly, compared to other types of traffic signs. Particularly, because severe accidents usually occur at dangerous locations, warning signs should be set up at certain intervals before the dangerous locations so as to remind the drivers in advance, or double warnings should be set up at some specific places. As for the total traffic signs, roadway operation/management departments should maintain and change them regularly in case of bad conditions so as to keep the signs valid all the time and avoid the collision risk as much as possible.

Data Availability

(1) Accident data from Traffic Accident Database System: the data that support the findings of this study are maintained by Republic of Croatia Ministry of the Interior, which may be available upon request. (2) Traffic signs data: the data that support the findings of this study, are available at University of Zagreb, Croatia. Please contact Professor Darko Babic (Email: dbabic@)fpz.hr) for more information.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] G. Villalón-Sepúlveda, M. Torres-Torriti, and M. Flores-Calero, "Traffic sign detection system for locating road intersections and roundabouts: The chilean case," *Sensors*, vol. 17, no. 6, 2017.

[2] J. L. Vilchez, "Representativity and univocity of traffic signs and their effect on trajectory movement in a driving simulation task: Warning signs," *Traffic Injury Prevention*, vol. 18, no. 5, pp. 544–549, 2017.

[3] J. Carson and F. Mannering, "The effect of ice warning signs on ice-accident frequencies and severities," *Accident Analysis & Prevention*, vol. 33, no. 1, pp. 99–109, 2001.

[4] R. J. Porter, K. M. Mahoney, and J. M. Mason, "Seemingly unrelated regression model of car speeds and speed deviations in freeway work zones," *Transportation Research Record*, vol. 2023, pp. 44–51, 2007.

[5] Q. Wu and J. J. Wang, "Exploration on speed-limit signs of traffic accident section on freeways," *Applied Mechanics and Materials*, vol. 361-363, pp. 2349–2352, 2013.

[6] D. Lord and F. Mannering, "The statistical analysis of crash-frequency data: a review and assessment of methodological alternatives," *Transportation Research Part A: Policy and Practice*, vol. 44, no. 5, pp. 291–305, 2010.

[7] P. T. Savolainen, F. L. Mannering, D. Lord, and M. A. Quddus, "The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1666–1676, 2011.

[8] F. L. Mannering and C. R. Bhat, "Analytic methods in accident research: methodological frontier and future directions," *Analytic Methods in Accident Research*, vol. 1, pp. 1–22, 2014.

[9] E. S. Park and D. Lord, "Multivariate poisson-lognormal models for jointly modeling crash frequency by severity," *Transportation Research Record*, no. 2019, pp. 1–6, 2007.

[10] P. C. Anastasopoulos and F. L. Mannering, "A note on modeling vehicle accident frequencies with random-parameters count models," *Accident Analysis & Prevention*, vol. 41, no. 1, pp. 153–159, 2009.

[11] K. El-Basyouny and T. Sayed, "Accident prediction models with random corridor parameters," *Accident Analysis & Prevention*, vol. 41, no. 5, pp. 1118–1123, 2009.

[12] X. Qin, "Quantile effects of casual factors on crash distributions," *Transportation Research Record*, vol. 2219, pp. 40–46, 2012.

[13] Y. Wang, X.-N. Feng, and X.-Y. Song, "Bayesian quantile structural equation models," *Structural Equation Modeling: A Multidisciplinary Journal*, vol. 23, no. 2, pp. 246–258, 2016.

[14] R. Koenker, "Quantile regression: 40 years on," *Annual Review of Economics*, vol. 9, pp. 155–176, 2017.

[15] X. Qin, M. Ng, and P. E. Reyes, "Identifying crash-prone locations with quantile regression," *Accident Analysis & Prevention*, vol. 42, no. 6, pp. 1531–1537, 2010.

[16] X. Qin and P. E. Reyes, "Conditional Quantile Analysis for Crash Count Data," *Journal of Transportation Engineering*, vol. 137, no. 9, pp. 601–607, 2011.

[17] H. Wu, L. Gao, and Z. Zhang, "Analysis of Crash Data Using Quantile Regression for Counts," *Journal of Transportation Engineering*, vol. 140, no. 4, p. 04013025, 2014.

[18] S. Washington, M. M. Haque, J. Oh, and D. Lee, "Applying quantile regression for modeling equivalent property damage only crashes to identify accident blackspots," *Accident Analysis & Prevention*, vol. 66, pp. 136–146, 2014.

[19] X. Liu, M. R. Saat, X. Qin, and C. P. L. Barkan, "Analysis of U.S. freight-train derailment severity using zero-truncated negative binomial regression and quantile regression," *Accident Analysis & Prevention*, vol. 59, pp. 87–93, 2013.

[20] K. Yu and R. A. Moyeed, "Bayesian quantile regression," *Statistics & Probability Letters*, vol. 54, no. 4, pp. 437–447, 2001.

[21] E. Waldmann and T. Kneib, "Bayesian bivariate quantile regression," *Statistical Modelling. An International Journal*, vol. 15, no. 4, pp. 326–344, 2015.

[22] X. Xu, S. Xie, S. C. Wong, P. Xu, H. Huang, and X. Pei, "Severity of pedestrian injuries due to traffic crashes at signalized intersections in Hong Kong: a Bayesian spatial logit model," *Journal of Advanced Transportation*, vol. 50, no. 8, pp. 2015–2028, 2016.

[23] R. Loo, "Role of primary personality factors in the perception of traffic signs and driver violations and accidents," *Accident Analysis & Prevention*, vol. 11, no. 2, pp. 125–127, 1979.

[24] F. Jørgensen and T. Wentzel-Larsen, "Optimal use of warning signs in traffic," *Accident Analysis & Prevention*, vol. 31, no. 6, pp. 729–738, 1999.

[25] Y. E. Xuan and A. Kanafani, "Evaluation of the effectiveness of accident information on freeway changeable message signs: A comparison of empirical methodologies," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 158–171, 2014.

[26] L. Strawderman, M. M. Rahman, Y. Huang, and A. Nandi, "Driver behavior and accident frequency in school zones: Assessing the impact of sign saturation," *Accident Analysis & Prevention*, vol. 82, pp. 118–125, 2015.

[27] D. Babić, A. Ščukanec, and M. Fiolić, "Traffic sign analysis as a function of traffic safety on Croatian state road," in *Proceedings of the ICTTE-International Conference on Traffic and Transportation Engineering*, Belgrade, Serbia, 2014.

[28] D. Babić, D. Babić, and D. Macura, "Model for predicting traffic signs functional service life-the republic of Croatia case study," *Promet-Traffic & Transportation*, vol. 29, no. 3, pp. 343–349, 2017.

[29] R. Koenker and J. Bassett, "Regression quantiles," *Econometrica*, vol. 46, no. 1, pp. 33–50, 1978.

[30] P. Hewson, "Quantile regression provides a fuller analysis of speed data," *Accident Analysis & Prevention*, vol. 40, no. 2, pp. 502–510, 2008.

[31] Y.-H. Ying, C.-C. Wu, and K. Chang, "The effectiveness of drinking and driving policies for different alcohol-related fatalities: A quantile regression analysis," *International Journal of Environmental Research and Public Health*, vol. 10, no. 10, pp. 4628–4644, 2013.

[32] W. G. Gilchrist, *Statistical Modelling with Quantile Functions*, Chapman & Hall/CRC, Boca Raton, FL, USA, 2000.

[33] Y. R. Yue and H. G. Hong, "Bayesian Tobit quantile regression model for medical expenditure panel survey data," *Statistical Modelling. An International Journal*, vol. 12, no. 4, pp. 323–346, 2012.

[34] A. F. Galvao, C. Lamarche, and L. R. Lima, "Estimation of censored quantile regression for panel data with fixed effects," *Journal of the American Statistical Association*, vol. 108, no. 503, pp. 1075–1089, 2013.

[35] L. Mokatrin, *Bayesian approach for sample selection bias correction in regression [Dissertation]*, American University, Washington, DC, USA, 2011.

[36] S. F. Leung and S. Yu, "On the choice between sample selection and two-part models," *Journal of Econometrics*, vol. 72, no. 1-2, pp. 197–229, 1996.

[37] J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica*, vol. 26, pp. 24–36, 1958.

[38] F. H. Alhusseini and V. Georgescu, "Bayesian composite Tobit quantile regression," *Journal of Applied Statistics*, vol. 45, no. 4, pp. 727–739, 2018.