Published in final edited form as:

Saf Sci. 2010 February 1; 48(2): 186–194. doi:10.1016/j.ssci.2009.07.010.

Investigation of Key Factors for Accident Severity at Railroad Grade Crossings by Using a Logit Model

Shou-Ren Hu¹, Chin-Shang Li², and Chi-Kang Lee³

Shou-Ren Hu: shouren@mail.ncku.edu.tw; Chin-Shang Li: cssli@ucdavis.edu; Chi-Kang Lee: leeck@mail.stut.edu.tw

¹ Assistant Professor, Department of Transportation and Communication Management Science, National Cheng Kung University, Tainan City 701, Taiwan.

² Associate Professor, Department of Public Health Sciences, Division of Biostatistics, University of California, Davis, CA, USA.

³ Professor, Department of Marketing and Logistics Management, Southern Taiwan University of Technology, Tainan County 710, Taiwan.

Abstract

Although several studies have used logit or probit models and their variants to fit data of accident severity on roadway segments, few have investigated accident severity at a railroad grade crossing (RGC). Compared to accident risk analysis in terms of accident frequency and severity of a highway system, investigation of the factors contributing to traffic accidents at an RGC may be more complicated because of additional highway–railway interactions. Because the proportional odds assumption was violated while fitting cumulative logit modeled by the proportional odds models with stepwise variable selection to ordinal accident severity data collected at 592 RGCs in Taiwan, as suggested by Strokes et al. (2000, p. 249) a generalized logit model with stepwise variable selection was used instead to identify explanatory variables (factors or covariates) that were significantly associated with the severity of collisions. Hence, the fitted model was used to predict the level of accident severity, given a set of values in the explanatory variables. Number of daily trains, highway separation, number of daily trucks, obstacle detection device, and approaching crossing markings significantly affected levels of accident severity at an RGC (p -value = 0.0009, 0.0008, 0.0112, 0.0017, and 0.0003, respectively). Finally, marginal effect analysis on the number of daily trains and law enforcement camera was conducted to evaluate the effect of the number of daily trains and presence of a law enforcement camera on the potential accident severity.

Keywords

Accident severity; generalized logit; marginal effect; proportional odds; railroad grade crossing

1. INTRODUCTION

Railroad grade crossings (RGCs), which allow traffic interchange between railroads and highways, are crucial spatial locations for transportation safety and efficiency in a surface

*Corresponding author: Shou-Ren Hu, PhD, Department of Transportation and Communication Management Science, National Cheng Kung University, Tainan City 701, Taiwan, Tel. 886-6-2757575 ext. 53203/Fax 886-6-2753882, shouren@mail.ncku.edu.tw.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

transportation network. Traffic accidents are on the rise at RGCs in Taiwan. From 2001 to 2005, 410 accidents occurred at approximately 630 RGCs in Taiwan, resulting in 125 fatalities and 153 injuries (Chang et al., 2006). In other words, from 2001 to 2005, 13.016 accidents and 3.968 fatalities occurred per year per 100 RGCs in Taiwan. These accidents have caused enormous property loss and killed and injured several people. Therefore, a model that can identify the key factors contributing to the likelihood and impact of a traffic accident at an RGC is required to improve the safety at RGCs.

Accident risk can be broken down into the product of likelihood (or frequency) and impact (or severity) (Ayyub, 2003). Accident frequency is the number of accidents observed in a specific time period and accident severity is the average level of impact of an accident. The level of impact of an accident can be property damage only or the extent of injury sustained by the most severely injured occupant(s) in the vehicle (Chang and Mannering, 1999). Therefore, the accident risk at an RGC is the level of impact during a period of time. In other words, we need to simultaneously address accident frequency and severity to correctly assess the accident risk of an RGC. For instance, in the US, the Federal Highway Administration (FHWA) requires each state to develop and implement a highway safety-improvement program (HSIP), which consists of three components: planning, implementation, and evaluation. The process for improving the safety and operations at highway–RGCs consists of the same three components and may be considered part of a state's HSIP (Ogden, 2007). The FHWA has developed a systematic method of defining and elaborating hazard indices and accident formulae to identify areas that most need safety or operational improvements, or both. One priority of this policy is to reduce the number and severity of collisions, or both, at highway–RGCs. Therefore, while assessing the safety level of an RGC, evaluating the level of accident severity is equally important to predict accident frequency in a given time period.

In the past two decades, several studies have focused on developing highway accident frequency models to model crashes on roadway segments (Miaou et al., 1992; Miaou and Lum, 1993; Miaou, 1994; Shankar et al., 1997; Vogt and Bared, 1998; Harwood et al., 2000; Shankar et al., 2004; Qin et al., 2004; Zhang and Ivan, 2005; Lord et al., 2005; El-Basyouny and Sayed, 2006) or crashes on intersections (Vogt and Bared, 1998; Vogt, 1999; Bauer and Harwood, 2000; Heydecker and Wu, 2001; Miaou and Lord, 2003; Oh et al., 2004; Persaud et al., 1999). Some studies have aimed to develop forecasting models to predict crashes at the planning level (Hedayeghi et al., 2003; de Guevara et al., 2004; Lovegrove and Sayed, 2006). Because of the nonnegative and count data nature of accident frequency for a time period, the Poisson or negative binomial (NB) models and their variants are developed to model crash or accident frequency, or both, at spatial locations on highways. Although several efforts have been made to investigate traffic accidents on highways, there is little research on modeling accidents at highway–RGCs (Gitelman and Hakkert, 1997; Austin and Carson, 2002; Saccomanno et al., 2004; Miranda-Moreno et al., 2005; Park and Saccomanno, 2007; Saccomanno et al., 2007; Miranda-Moreno et al., 2009). Models that capture traffic accidents at RGCs are mostly based on statistical inference.

Studies focusing on modeling accident severity of highway systems have defined various response variables in terms of the levels of impact of a collision. Accident severity levels of highways are usually classified as (1) property damage only, (2) minor injury, (3) severe injury, and (4) fatality. Because of the nature of accident severity data, ordinal or discrete choice models and relevant inference approaches are employed to determine the categorical characteristics of accident severity on highways. Different levels of injury in motor vehicle accidents are modeled either by ordered logit and probit models (O'Donnell and Connor, 1996; Kockelman and Kweon, 2002; Kweon and Kockelman, 2003) or multinomial logit models (Shankar and Mannering, 1996; Carson and Mannering, 2001; Lee and Mannering, 2002; Khorashadi et al., 2005). Alternatively, nested logit models are applied to model different

levels of accident severity under a nested framework of either considering property damage only or possible injury (Shankar et al., 1996; Lee and Mannering, 2002) or additional vehicle occupants (Chang and Mannering, 1999). Other variants of logit models also capture the different degrees of severity of collision at highways. For instance, Yamamoto and Shankar (2004) used a bivariate ordered-response probit model to simultaneously capture different levels of accident severity and most severely injured passengers. Yau (2004) analyzed the key factors affecting the severity of single-vehicle traffic accidents by a logistic regression model with stepwise variable selection. Milton et al. (2007) used a mixed logit model to analyze highway accident severity in which the model parameters can vary randomly across roadway segments to account for unobserved effects related to different highway and driver characteristics. To account for the uncertainty associated with the accident frequency/severity model parameters and some of the key variables such as vehicle occupancy levels and severity weighting factors, Miranda-Moreno et al. (2009) proposed a Bayesian multinomial model to estimate the severity levels of each individual involved in an accident. More importantly, compared to the traditional accident frequency or consequence models, the risk-based Bayesian accident risk analysis framework which integrates both accident frequency and its expected consequences can provide more objective and effective measures of total risk for hotspot identification.

In contrast to the count data nature of accident frequency, accident severity levels, which are usually classified into different degrees of impact of an accident, are ordinal categories. It is more appropriate to use the so-called cumulative logit, which can be modeled by the proportional odds models (McCullagh, 1990), for ordinal categorical outcomes, but the p -value of a score test for proportional odds assumption is 0.01, which is less than the commonly used significance level 0.05 and, hence, indicates there was no proportionality when fitting this model with stepwise variable selection to the traffic accident severity data at RGCs in Taiwan by using PROC LOGISTIC with LINK=CUMLOGIT in SAS V9.1 (SAS Institute, Inc., Cary, NC) (see Section 2 for detailed description of the traffic accident severity data). Therefore, as suggested by Strokes et al. (2000, p. 249), in practice it would be best to treat accident severity data as nominal and fit a model to generalized logit. Consequently, the generalized logit model with stepwise variable selection was used to explore the key factors that may be responsible for different degrees of accident severity at RGCs. We used the traffic accident severity data and other traffic exposure and railroad/highway-related data in Taiwan to identify the key factors significantly associated with the severity of accidents at an RGC. Elasticity analysis in terms of the marginal effects of specific explanatory variables (e.g., number of daily trains and law enforcement camera) was conducted to evaluate the effect of the number of daily trains and presence of a law enforcement camera on accident severity. The highway-related and railway-related characteristics that were identified as contributing to accident severity can be further explored by field investigation and effective countermeasures need to be taken accordingly. The results and corresponding applications may help both government agencies and transportation system users improve safety at an RGC.

The remainder of this paper is organized as follows. Section 2 details the descriptive statistics and data characteristics for the accident severity data and other railroad and highway features. Section 3 provides background information on the generalized logit model. Section 4 gives results of data analysis and their interpretations. Section 5 discusses the implications of using generalized logit models to analyze the accident severity data at RGCs in Taiwan. Finally, Section 6 summarizes findings, draws conclusions, and gives directions for future research.

2. DATA DESCRIPTION

The railway police agency and the Taiwan Rail Administration (TRA), the Ministry of Transportation and Communications (MOTC), routinely record and maintain traffic accident

data at RGCs in Taiwan. However, the government offices do not regularly collect railway- and highway-related characteristics. The only available historical accident data set for modeling traffic accidents at RGCs were collected from 1995 through 1997, in which attributes associated with railway and highway features and traffic exposure with regard to the number of daily trains and average annual daily traffic (AADT) were recorded (MOTC, 1998). The traffic accident data consist of accident frequency and fatality or injury in the three-year time period. From 1995 to 1997, no accidents occurred at 399 (67.4%) of 592 RGCs; 321 accidents occurred at the remaining 193 (32.6%) RGCs.

Impacts of fatality or injury are our primary concern and were classified into four accident severity levels designated 0, 1, 2, and 3: higher the severity level, more severe the accident. For a specific RGC, we defined accident severity by an equivalent injury divided by accident frequency, that is, the average equivalent injury per accident in a given time period. The equivalent injury captures the equivalent impact of a fatality relative to an injury. It is adopted from the Institute of Transportation (MOTC, 1998) in Taiwan, where 1 injury is equivalent to 0.368 fatalities. In other words, the severity of a fatality is equivalent to 2.72 injuries. Accordingly, a severity of 0 means no injuries; 1 means 2.72 equivalent injuries; 2 means 5.44 equivalent injuries; and 3 means 8.16 or more equivalent injuries. In the historical accident data set, a traffic accident severity of 3 occurred at three RGCs only. Therefore, to investigate the explanatory variables significantly associated with accident severity, we created an ordinal three-level response variable by combining the accident severities of 2 and 3. Let Y be an ordinal three-level response variable that is 0 if accident severity is 0; 1 if accident severity is 1; 2 if accident severity is 2 or 3. Table 1 gives the observed accident severity frequency and frequency of $Y = y$, $y = 0, 1, 2$.

The original dataset consisted of 35 explanatory variables, including railway features (number of daily trains, train speed, and number of tracks), highway features (highway type, highway authority, highway width, highway grade, highway lanes, highway separation, annual averaged daily traffic, and daily trucks), crossing features (crossing type, crossing width, crossing length, crossing angle, crossing surface, height restricted gate, obstacle detection device, warning bell and flashing light, emergence call, and trolley wire), traffic control (guidance sign 31, guidance sign 32, guidance sign 33, guidance sign 34, warning sign 25, warning sign 26, warning sign 27, warning sign 28, warning sign 29, approaching crossing markings, stop line markings, police patrol, and law enforcement camera), and others (land use of the area adjacent to an RGC).

To identify the covariates that might have different levels of impact on accident severity, we first began with a univariable analysis of each independent variable to examine the logical relationship between the outcome variable and potential independent variables. We then selected variables for the multivariable analysis. Any variable whose univariable test had a p -value of log-likelihood ratio test < 0.25 was a candidate for the multivariable model along with all variable of known importance of accident severity at RGCs discussed in the references. We used the level of 0.25 as a screening criterion for variable selection based on the work by Mickey and Greenland (1989). As a result, 27 independent variables among totally 35 were selected for modeling purposes:

- Railway feature: daily trains, train speed, track number;
- Highway feature: highway type, highway width, highway grade, highway lanes, highway separation, annual averaged daily traffic, daily trucks;
- Crossing feature: crossing type, crossing width, crossing length, crossing angle, crossing surface, obstacle detection device, warning bell and flashing light, emergence call, trolley wire;

- Traffic controls: guidance sign 31, guidance sign 33, guidance sign 34, approaching crossing markings, stop line markings, police patrol, law enforcement camera;
- Others: land use.

We further analyzed the logical relationships between the outcome variable and potential independent variables to be included in the model. For the 27 independent variables selected for later modeling purposes, in the railway feature, number of daily trains, train speed, and number of track(s) were selected into the modeling process. Since the more daily trains, the higher traffic exposure, causing more traffic accidents with different accident severity levels. The higher the train speed, the more severe the accident consequence if an accident occurs at the RGC. Similarly, the more the rail tracks, the higher the train frequency, resulting in higher possibility of traffic accidents and/or more severe consequence. For the covariates associated with the highway feature, highway type, highway width, highway grade, highway lanes, highway separation, annual averaged daily traffic, daily trucks were chosen assuming to have different impacts on traffic accidents at RGCs. Five highway types are classified: (1) provincial highway, (2) county road, (3) urban street, (4) rural road, and (5) farm road. The higher functional levels of the highway type, the more severe traffic accidents might be observed at RGCs, since higher traffic exposure is generally found at these highway types. Similarly, the wider the highway, the more severe accidents might be observed if an accident(s) happens at the RGC area. The deeper the highway grade, the more dangerous situations might occur since the driver might be constrained by insufficient vertical sight distance. Highways are generally divided with a median or markings separation. Traffic accidents, which are happened at highways with a median separation, are generally less severe than highways with traffic markings separation or without separation. In addition, highways with more daily traffic might incur different degrees of impact on accident severity if an accident happens at the RGC area, and more truck traffic might have higher probability of more severe accidents. In terms of the elements associated with the crossing feature, crossing type, crossing width, crossing length, crossing angle, crossing surface, obstacle detection device, warning bell and flashing light, emergence call, trolley wire might have different degrees of impact on accident severity at RGCs. Finally, for the traffic control feature, different traffic control devices such as traffic signage and markings might have different effects on reducing more severe traffic accidents. In addition, special traffic controls such as police patrol, law enforcement camera are expected to have effects on reducing traffic accidents and more severe accidents at the RGC area. Different land uses around the RGC area might incur different degrees of traffic exposure, resulting in different severity levels of traffic accidents.

After the 27 independent variables that potentially have an impact on the outcome variable (accident severity) were selected as candidates for the multivariable model, we then included these candidate variables via the stepwise selection procedure and by trying these different combinations of significance level for entering candidate variables, and significance level for removing candidate variables: (0.15, 0.20), (0.15, 25), (0.15, 0.30), (0.15, 35), (0.15, 0.40), (0.15, 45), (0.15, 0.50), (0.15, 55), (0.20, 25), (0.20, 0.30), (0.20, 35), (0.20, 0.40), (0.20, 45), (0.20, 0.50), (0.20, 55), (0.25, 0.30), (0.25, 35), (0.25, 0.40), (0.25, 45), (0.25, 0.50), and (0.25, 55) while including train speed, highway type and crossing angle in every model. We finally obtained the best model according to the Pearson goodness-of-fit criterion. The best model consists of 12 independent variables. Among the 12 independent variables, 5 independent variables are statistically associated with the ordinal outcomes of the accident severity at RGCs. See the details in a later numerical analysis section.

3. STATISTICAL METHODS

As detailed in Sections 1 and 2, we will use the generalized logit model, instead of the cumulative logit model, for the ordinal 3-level outcomes. We now introduce the generalized

logit model to investigate the relationship between the three-level response variable Y and explanatory variables. Let $\mathbf{x}=(x_1, x_2, \dots, x_{16})$ be a 16-variable vector corresponding to an RGC. Here, x_1 denotes the number of daily trains. x_2 denotes train speed. (x_3, x_4, x_5, x_6) is used to denote highway type, which is (0, 0, 0, 0) for farm road; (1, 0, 0, 0) for county road; (0, 1, 0, 0) for provincial highway; (0, 0, 1, 0) for rural road; (0, 0, 0, 1) for urban street. x_7 is a binary variable for highway width, which is 0 for highway width less or equal than 3.5 meters; 1 otherwise. (x_8, x_9) is used to denote highway separation, which is (0, 0) for no separation; (1, 0) for marking separation; (0, 1) for highways with a median separation. x_{10} denotes daily trucks/100. x_{11} denotes crossing angle. x_{12} is a binary variable for obstacle detection device, which is 1 for an RGC equipped with the device and 0 for being not equipped. x_{13} is a binary variable for guidance signed 34, which is 1 for an RGC equipped with the sign and 0 for being not equipped. x_{14} is a binary variable for approaching crossing markings, which is 1 for an RGC warned by the markings and 0 for being not warned. x_{15} is a binary variable for stop line markings, which is 1 for an RGC warned by the markings and 0 for being not warned. x_{16} is a binary variable for law enforcement camera, which is 1 for an RGC equipped with the device and 0 for being not equipped. Let $\pi_j(\mathbf{x}) = P(Y=j;\mathbf{x})$ be the probability that $Y=j, j=0, 1, 2$.

Note that $\sum_{j=0}^2 \pi_j(\mathbf{x})=1$.

We define the generalized logit as follows:

$$\text{logit}[\pi_j(\mathbf{x})]=\log \frac{\pi_j(\mathbf{x})}{\pi_0(\mathbf{x})}, j=1, 2.$$

That is, we form a logit for the probability of $Y=j$ over $Y=0$. Each generalized logit is modeled by a linear predictor as follows:

$$\text{logit}[\pi_j(\mathbf{x})]=\log \frac{\pi_j(\mathbf{x})}{\pi_0(\mathbf{x})}=\alpha_j+\mathbf{x}\boldsymbol{\beta}_j, j=1, 2. \quad (1)$$

Here, α_j is an intercept parameter and $\boldsymbol{\beta}_j=(\beta_{j1}, \beta_{j2}, \dots, \beta_{j,16})^T$ is a 16-dimensional vector of regression parameters for the j th logit. By using Eq. (1) and taking $\alpha_0=0$ and $\boldsymbol{\beta}_0=\mathbf{0}$ for notational simplicity, we can express the probability $\pi_j(\mathbf{x})$ as follows:

$$\pi_j(\mathbf{x})=\frac{\exp(\alpha_j+\mathbf{x}\boldsymbol{\beta}_j)}{\sum_{k=0}^2 \exp(\alpha_k+\mathbf{x}\boldsymbol{\beta}_k)}, j=0, 1, 2. \quad (2)$$

Let $(\hat{\alpha}_j, \hat{\boldsymbol{\beta}}_j)$ be the maximum likelihood estimate (MLE) of $(\alpha_j, \boldsymbol{\beta}_j)$, $j=0, 1, 2$, with $(\hat{\alpha}_0, \hat{\boldsymbol{\beta}}_0)=(0, \mathbf{0})$ because $(\alpha_0, \boldsymbol{\beta}_0)$ was taken as $(0, \mathbf{0})$ for simplicity of notation, where $\hat{\boldsymbol{\beta}}_j=(\hat{\beta}_{j1}, \hat{\beta}_{j2}, \dots, \hat{\beta}_{j,16})^T$. In addition, the deviance and Pearson goodness-of-fit statistics were used to assess whether the generalized logit model was adequate: higher the p -value, better the fit of the generalized logit model. Analyses were performed by using the PROC LOGISTIC with LINK=GLOGIT in SAS V9.1 (SAS Institute, Inc., Cary, NC).

4. RESULTS AND DISCUSSION

The deviance and Pearson goodness-of-fit statistics, 564.4895 (deviance/degrees of freedom = 564.4895/1036 = 0.5449) and 974.9986, (Pearson residual/degrees of freedom = 974.9986/1036 = 0.9411), respectively, indicated that the fitted generalized logit model had

adequate goodness of fit (p -value = 1 for deviance and 0.9119 for Pearson). By fitting the generalized logit model with Wald-type chi-squared test, there was a significant association between the severity of collisions and the number of daily trains, highway separation, daily trucks, obstacle detection device, and approaching crossing markings, as seen in Table 2 (p -value = 0.0009, 0.0008, 0.0112, 0.0017, and 0.0003, respectively).

We can interpret the effects of the aforementioned explanatory variables on accident severity in terms of odds ratios. Table 3 presents the results of the odds ratio estimation. The odds were $\exp(\beta_{21}) = \exp(0.012) = 1.012$, 95% CI (1.000, 1.023), times statistically significantly higher of $Y=2$ than $Y=0$ if the number of daily trains was increased by 1. The odds were $\exp(\beta_{11}) = \exp(0.012) = 1.012$, 95% CI (1.005, 1.020), times statistically significantly higher of $Y=1$ than $Y=0$ if the number of daily trains was increased by 1. Similarly, it was $\exp(\beta_{2,10}) = \exp(0.013) = 1.013$, 95% CI (0.925, 1.109), times more likely to have $Y=2$ than $Y=0$ if the daily truck was increased by 100 trucks (i.e., daily trucks/100 increased by 1 truck). It was $\exp(\beta_{1,10}) = \exp(0.082) = 1.085$, 95% CI (1.026, 1.148), times statistically significantly more likely to have $Y=1$ than $Y=0$ if daily trucks/100 was increased by 1 truck.

Traffic exposure variables such as the number of daily trains and the number of daily trucks had positive effects on the occurrence of more severe accidents. Similarly, highways with marking separation had also positive effects on more severe accidents; highways with marking separation were more likely to observe more severe accidents than highways without separation. The highway separated by traffic markings had positive effects on the accident severity. Specifically, the odds were $\exp(\beta_{28}) = \exp(1.858) = 6.411$, 95% CI (2.460, 16.707), times statistically significantly higher of $Y=2$ than $Y=0$ if the highway was separated by markings. Similarly, the odds were $\exp(\beta_{18}) = \exp(0.676) = 1.966$, 95% CI (1.015, 3.808), times statistically significantly higher of $Y=1$ than $Y=0$ if the highway was separated by markings. However, even though the divided highways had positive effects on the accident severity, the 95% confidence interval of odds ratio contains 1, indicating no statistical difference between severity levels (i.e. $Y=1$ versus $Y=0$). Specifically, the odds were $\exp(\beta_{19}) = \exp(-0.266) = 0.767$, 95% CI (0.175, 3.360), times higher of $Y=1$ than $Y=0$ if the highway was separated by a median.

The obstacle detection device had positive effects on the accident severity. Specifically, the odds were $\exp(\beta_{2,12}) = \exp(2.761) = 15.818$, 95% CI (3.471, 72.094), times statistically significantly higher of $Y=2$ than $Y=0$ if the obstacle detection device was equipped. Similarly, the odds were $\exp(\beta_{1,12}) = \exp(0.834) = 2.303$, 95% CI (0.263, 20.156), times higher of $Y=1$ than $Y=0$ if the obstacle detection device was equipped.

Finally, approaching crossing markings had different effects on the accident severity. Specifically, the odds were $\exp(\beta_{2,14}) = \exp(0.588) = 1.8$, 95% CI (0.736, 4.401), times higher of $Y=2$ than $Y=0$ if the crossing was warned by approaching markings. However, the odds were $\exp(\beta_{1,14}) = \exp(-1.089) = 0.337$, 95% CI (0.184, 0.615), times statistically significantly lower of $Y=1$ than $Y=0$ if the crossing was warned by approaching markings.

Based on the above model calibration results, it indicates that traffic control and/or management (i.e. highway separation and obstacle detection) did not prevent RGCs from the occurrence of more severe accidents. It is surprising that the analysis results of highway separation and obstacle detection device contradict our expectation that they were presumably expected to reduce the probability of the severity levels ($Y=2$ or 1) under a situation where highways were separated by traffic markings or equipped with obstacle detection device. One common characteristic found in the RGCs with more severe accidents is that these RGCs are usually located in urban areas where traffic exposure is relatively high, and compared to rural areas more traffic accidents are observed in these traffic busy areas. Thereby, the government office

has spent a lot of monetary resources to equip these RGCs with many traffic control devices. By this improvement, it might help to reduce the number of traffic accidents, however if an accident happens at an RGC in urban areas where various means of crossing protective devices are generally provided (e.g., highways with a median separation, obstacle detection device, and approaching crossing markings), a traffic accident is usually more severe than those found in the rural areas. Although the model estimation results might contradict our expectation, the empirical results really reflected the traffic accident data nature and local traffic characteristics at RGCs in Taiwan. In summary, engineering-based investment could reduce traffic accident frequency but might not have significant effects on reducing more severe accidents.

5. APPLICATIONS

The generalized logit model was used to fit the data of accident severity at RGCs in Taiwan. We now discuss the implications and interpret the generalized logit model with the accident severity data at RGCs in Taiwan. These results can be used to reduce accident severity in the future. The idea of Liao (1994) was applied to the generalized logit model to interpret the predicted probability of the severity of collisions, given a set of values in the explanatory variables, and the marginal effects of the explanatory variables on the probability of accident severity.

5.1. Predicted Probability of the Severity of a Traffic Accident by Changing Daily Trains

The predicted probability of accident severity, given a set of values in the explanatory variables, can be calculated by plugging the MLEs $(\hat{\alpha}_j, \hat{\beta}_j)$, $j=1, 2$ (Table 2), and $(\hat{\alpha}_0, \hat{\beta}_0) = (0, \mathbf{0})$ into Eq. (2) as follows:

$$\hat{\pi}_j(\mathbf{x}) = \hat{P}(Y=j;\mathbf{x}) = \frac{\exp(\hat{\alpha}_j + \mathbf{x}\hat{\beta}_j)}{\sum_{k=0}^2 \exp(\hat{\alpha}_k + \mathbf{x}\hat{\beta}_k)}, \quad j=0, 1, 2. \quad (3)$$

Thus, by using Eq. (3), we can evaluate the effect of the change of an explanatory variable of interest on the probability of severity of collisions, given that the other variables are held at their sample means or at certain levels. For example, we can obtain the predicted probabilities of different accident severities at the different numbers of daily trains while holding the train speed, number of daily trucks, and crossing angle at their sample mean, highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, and guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present.

Figure 1 shows the plots of predicted probabilities of $Y=j$, $j=0, 1, 2$ versus the number of daily trains in the true range of 18–240 trains in the accident severity data set. To clearly see the trend of changes of the predicted probabilities over the number of daily trains, we show plots of the predicted probabilities of $Y=j$, $j=0, 1, 2$ versus the number of daily trains in the hypothetical range of 18–400 trains in Figure 2. For plots of predicted probabilities of accident severity levels versus the number of daily trains in Figs. 1 and 2, we hold the train speed, number of daily trucks, and crossing angle at their sample mean and the highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present. Figure 2 shows that the probability of $Y=0$ (i.e., accident severity of 0) decreases with the increase in the number of daily trains. Moreover, the probability of $Y=1$ (i.e., accident severity of 1)

increases slowly as the number of daily trains increases. However, the probability of $Y = 2$ (i.e., accident severity of 2) increases quickly with the increase in the number of daily trains. The predicted probability results on different accident severity levels indicate that more severe accidents are observed as the number of daily trains is increasing. This is due to that more traffic conflicts might occur at the crossing area when traffic exposure with regard to the number of daily trains and/or daily trucks is increased.

5.2. Marginal Effect of Daily Trains on the Probability of Severity of a Traffic Accident

The marginal effect of an explanatory variable of interest on the probability of $Y = j$ is given by taking the first-order partial derivative of $\pi_j(\mathbf{x})$ with respect to the explanatory variable as follows:

$$\frac{\partial \pi_j(\mathbf{x})}{\partial x_p} = \frac{\partial}{\partial x_p} \frac{\exp(\alpha_j + \mathbf{x}\beta_j)}{\sum_{k=0}^2 \exp(\alpha_k + \mathbf{x}\beta_k)} = \pi_j(\mathbf{x}) \left[\beta_{jp} - \sum_{k=0}^2 \pi_k(\mathbf{x}) \beta_{kp} \right], j=0, 1, 2, p=1, 2, \dots, 16. \quad (4)$$

Here $\beta_{0p} = 0, p=1, 2, \dots, 16$ and hence $\pi_0(\mathbf{x})\beta_{0p} = 0$. For example, we can estimate the marginal effects of the number of daily trains on the probability of accident severity by plugging the MLEs $(\hat{\alpha}_j, \hat{\beta}_j), j=1, 2$ (Table 2), and $(\hat{\alpha}_0, \hat{\beta}_0) = (0, \mathbf{0})$ into Eq. (4), while holding the train speed, number of daily trucks, and crossing angle at their sample mean and the highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present. Figure 3 shows the estimated marginal effects of the number of daily trains on the probabilities of different accident severity levels for the true range of the number of daily trains of 18–240. In addition, to clearly see the estimated marginal effects of the number of daily trains on the probabilities of different severity levels of collisions, we consider a hypothetical range of the number of daily trains of 18–400 and show the plots in Fig. 4. For plots of estimated marginal effects of the number of daily trains on the probabilities of accident severities versus the number of daily trains in Figs. 3 and 4, the train speed, number of daily trucks, and crossing angle are held at their sample mean and the highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present. Figure 4 shows that the marginal effect of the number of daily trains on the probability of $Y=0$ (i.e., accident severity of 0) increases with the increase in the number of daily trains; then levels off to 0. The marginal effect of the number of daily trains on the probability of $Y=1$ (i.e., accident severity of 1) decreases slowly when the number of daily trains is increased; then levels off to 0. However, the marginal effect of the number of daily trains on the probability of $Y=2$ (i.e., traffic accident severity of 2 or 3) decreases quickly with the increase in the number of daily trains; then levels off to 0. In addition, the marginal effect of the number of daily trains on the probability of $Y=0$ is always negative, indicating that the probability of $Y=0$ decreases with the increase in the number of daily trains. However, the marginal effect of the number of trains on the probability of $Y=1$ and that of $Y=2$ is always positive, indicating that the probability of $Y=1$ and that of $Y=2$ increase as the number of daily trains increases.

5.3. Predicted Probability of the Severity of a Traffic Accident Due to Law Enforcement Camera

The predicted probabilities of $Y = j, j = 0, 1, 2$ at the presence of law enforcement camera in the accident severity data set are 0.3072, 0.0652, and 0.6276, respectively; the predicted

probabilities of $Y=j, j=0, 1, 2$ at the absence of law enforcement camera in the accident severity data set are 0.0768, 0.0656, and 0.8576, respectively, while holding the daily trains, train speed, number of daily trucks, and crossing angle at their sample mean, highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, and guidance sign 34, approaching crossing markings, and stop line markings at being present. The probability of $Y=0$ (i.e., accident severity of 0) increases with the presence of law enforcement camera. However, the probabilities of $Y=1$ (i.e., accident severity of 1) and $Y=2$ (i.e., accident severity of 2) decrease with the presence of law enforcement camera.

5.4 Marginal Effect of Law Enforcement Camera on the Probability of Severity of a Traffic Accident

The estimated marginal effects of law enforcement camera on the probability of accident severity are 0.2304, -0.0004 , and -0.2300 , respectively, while holding the daily trains, train speed, number of daily trucks, and crossing angle at their sample mean and the highway type at county road, highway width at being dichotomized to be 1 (i.e. > 3.5 meters), highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, and stop line markings at being present. Similarly, the marginal effect of law enforcement camera on the probability of $Y=0$ (i.e., accident severity of 0) increases with the presence of law enforcement camera. However, the marginal effect of law enforcement camera on the probabilities of $Y=1$ (i.e., accident severity of 1) and $Y=2$ (i.e., accident severity of 2) decrease with the presence of law enforcement camera.

6. CONCLUDING REMARKS

The generalized logit model is an appropriate methodological approach for exploring the key factors that significantly affect accident severity at an RGC. The generalized logit model with stepwise variable selection gives us better insight into the key factors that statistically significantly contribute to the different degrees of severity of collisions. The number of daily trains, highway separation, number of daily trucks, obstacle detection device, and approaching crossing markings statistically significantly affect the levels of accident severity at an RGC. It was found in the empirical analysis that traffic exposure variables such as the number of daily trains and number of daily trucks had positive effects on the more severe accidents since more potential conflicts might occur. On the other hand, the traffic control and management tools, highway separation and obstacle detection device surprisingly did not have an effect on preventing RGCs from the more severe accidents. This interesting but surprising result seems to contradict our expectation that the probability of the more severe accidents might be reduced given that an RGC is warned by highway separation with traffic markings or equipped with an obstacle detection device. One common characteristic found in the RGCs with more severe accidents is that these RGCs are usually located in urban areas where traffic exposure is relatively high (including truck traffic), and various means of crossing protective devices are generally found. However, if an accident happens at an RGC in urban areas, due to more traffic exposure/conflict, more severe accidents are found in these busy RGCs usually equipped with traffic control devices.

We also conducted a marginal effect analysis to evaluate the effect of the number of daily trains and law enforcement camera on accident severity. The sensitivity analysis enhances our understanding of the number of daily trains and other potential factors that may cause a severe collision at an RGC. These estimations and application results can serve as a good reference while preparing safety-improvement programs and fulfill the requirements of applying for HSIP resources. The empirical results can be beneficial to government agencies to identify the most dangerous RGCs (blackspots) in preparing the corresponding countermeasures for safety improvement purposes.

In the past two decades, several studies have focused on investigating the key factors associated with accident frequency (likelihood) on highways and railroads. In the past decade, the issue of highway or railroad accident severity has also received increasing attention. However, to fairly evaluate the safety level of an RGC, it would be more appropriate to model crash counts by severity to better understand the key factors that may contribute to a potential accident risk with regard to collisions by different severity levels. An investigation on accident frequency by severity by using a bivariate count data model (e.g., Song et al., 2006; Ma and Kockelman, 2006; Park and Lord, 2007; Ma et al., 2008) is capable of realistically exploring the accident risk at RGCs. Furthermore, to incorporate accident severity and some of the key factors such as vehicle occupancy into a total risk model (e.g. Miranda-Moreno et al., 2009) is crucial for hotspot identification. These are crucial topics that are worthy of future research. Finally, an in-depth investigation on driver's behaviors at individual RGCs which is currently conducted by the authors might answer the contradictory model estimation results found in this research.

Acknowledgments

This research was partially supported by grants NSC 95-2415-H-006-014-MY3 (S.R. Hu and C.K. Lee). This publication was made possible by Grant Number UL1 RR024146 from the National Center for Research Resources (NCRR), a component of the National Institutes of Health (NIH), and NIH Roadmap for Medical Research (C.S. Li). We thank Vani Shanker, PhD, for her help in editing the manuscript and Matthew P. Smeltzer, MS, for reading this manuscript.

References

- Austin RD, Carson JL. An alternative accident prediction model for highway-rail interfaces. *Accident Analysis and Prevention* 2002;34(1):31–42. [PubMed: 11789573]
- Ayyub, BM. *Risk Analysis in Engineering and Economics*. Chapman & Hall/CRC; New York: 2003.
- Bauer, KM.; Harwood, DW. *Statistical Models of At-grade Intersection Accidents – Addendum, FHWA-RD-99-094*, the Turner-Fairbank Highway Research Center, FHWA. U.S. Department of Transportation; Washington, D.C: 2000.
- Carson J, Mannering F. The effect of ice warning signs on ice-accident frequencies and severities. *Accident Analysis and Prevention* 2001;33(1):99–109. [PubMed: 11189127]
- Chang LY, Mannering F. Analysis of injury severity and vehicle occupancy in truck-and non-truck-involved accidents. *Accident Analysis and Prevention* 1999;31(5):579–592. [PubMed: 10440555]
- Chang, SH.; Lee, CK.; Hu, SR.; Tao, CC. Technical report. Transportation and Communications Management Unit, Ministry of Transportation and Communications; Taiwan: 2006. A Website and Information Platform for Railroad Grade Crossing.
- de Guevara FL, Washington SP, Oh J. Forecasting crashes at the planning level: Simultaneous negative binomial crash model applied in Tucson, Arizona. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897 2004:191–199.
- El-Basyouny K, Sayed T. Comparison of two-binomial regression techniques in developing accident prediction models. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1950 2006:9–16.
- Gitelman V, Hakkert AS. The evaluation of road-rail crossing safety with limited accident statistics. *Accident Analysis and Prevention* 1997;29(2):171–179. [PubMed: 9088356]
- Hadayeghi A, Shalaby AS, Persaud BN. Macrolevel accident prediction models for evaluating safety of urban transportation systems. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1840 2003:87–95.
- Harwood, DW.; Council, FM.; Hauer, E.; Hughes, WE.; Vogt, A. *Prediction of the Expected Safety Performance of Rural Two-lane Highways*. FHWA, U.S. Department of Transportation; 2000. FHWA-RD-99-207
- Heydecker BG, Wu J. Identification of sites for accident remedial work by Bayesian statistical methods: An example of uncertain inference. *Advances in Engineering Software* 2001;32:859–869.

- Khorashadi A, Niemeier D, Shankar V, Mannering F. Differences in rural and urban driver-injury severities in accidents involving large-trucks: An exploratory analysis. *Accident Analysis and Prevention* 2005;37(5):910–921. [PubMed: 15935320]
- Kockelman KM, Kweon YJ. Driver injury severity: an application of ordered probit models. *Accident Analysis and Prevention* 2002;34(3):313–321. [PubMed: 11939360]
- Kweon YJ, Kockelman KM. Overall injury risk to different drivers: combining exposure, frequency, and severity models. *Accident Analysis and Prevention* 2003;35(4):441–450. [PubMed: 12729808]
- 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* 2002;34(2):149–161. [PubMed: 11829285]
- Liao, TF. Interpreting probability models: Logit, Probit, and Other Generalized Linear Models, Series: Quantitative Applications in the Social Sciences. SAGE Publications; Thousand Oaks, CA: 1994.
- Lord D, Washington SP, Ivan JN. Poisson, Poisson-gamma and zero-inflated regression models of motor vehicle crashes: balancing statistical fit and theory. *Accident Analysis and Prevention* 2005;37(1): 35–46. [PubMed: 15607273]
- Lovegrove GR, Sayed T. Using macrolevel collision prediction models in road safety planning applications. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1950 2006:73–82.
- Ma, J.; Kockelman, KM. *Transportation Research Record* 1950. Transportation Research Board; Washington, D.C: 2006. Bayesian multivariate Poisson regression for models of injury count, by severity; p. 24–34.
- Ma J, Kockelman KM, Damien P. A multivariate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods. *Accident Analysis and Prevention* 2008;40:964–975. [PubMed: 18460364]
- McCullagh P. Regression models for ordinal data (with discussion). *Journal of the Royal Statistical Society, Series B* 1980;42:109–142.
- Miaou SP. The relationship between truck accident and geometric design of road sections: Poisson versus negative binomial regressions. *Accident Analysis and Prevention* 1994;26(4):471–482. [PubMed: 7916855]
- Miaou SP, Hu PS, Wright T, Rathi AK, Davis SC. Relationship between truck accidents and highway geometric design: A Poisson regression approach. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1376 1992:10–18.
- Miaou SP, Lord D. Modeling traffic crash-flow relationships for intersections: Dispersion parameter, functional form, and Bayes versus empirical Bayes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1840 2003:31–40.
- Miaou SP, Lum H. Modeling vehicle accidents and highway geometric design relationships. *Accident Analysis and Prevention* 1993;25(6):689–709. [PubMed: 8297437]
- Mickey RM, Greenland S. The impact of confounder selection criteria of effect estimation. *American Journal of Epidemiology* 1989;129:125–137. [PubMed: 2910056]
- Milton JC, Shankar V, Mannering F. Highway accident severities and the mixed logit model: An exploratory empirical analysis. *Accident Analysis and Prevention* 2008;40(1):260–266. [PubMed: 18215557]
- Ministry of Transportation and Communications, MOTC. A Study on the Planning of Railway Level Crossing Safety Improvement in Taiwan. MOTC; Taiwan: 1998.
- Miranda-Moreno LF, Fu L, Saccomanno FF, Labbe A. Alternative risk models for ranking locations for safety improvement. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1908 2005:1–8.
- Miranda-Moreno, LF.; Fu, L.; Ukkusuri, S.; Lord, D. How to incorporate accident severity and vehicle occupancy into the hotspot identification process?. CD-ROM of the 88th Annual Meeting of the Transportation Research Board, Paper No. 09–2824; Washington, D.C. 2009.
- O'Donnell CJ, Connor DH. Predicting the severity of motor vehicle accident injuries using models of ordered multiple choice. *Accident Analysis and Prevention* 1996;28(6):739–753. [PubMed: 9006642]

- Ogden, BD. 2. Federal Highway Administration, U.S. Department of Transportation; Washington, D.C: 2007. Railroad-highway Grade Crossing Handbook, rev.
- Oh J, Washington SP, Choi K. Development of accident prediction models for rural highway intersections. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897 2004:18–27.
- Park, ES.; Lord, D. *Transportation Research Record* 2019, Transportation Research Board. Washington, D.C: 2007. Multivariate Poisson-lognormal models for jointly modeling crash frequency by severity; p. 1-6.
- Park PYJ, Saccomanno FF. Reducing treatment selection bias for estimating treatment effects using propensity score method. *Journal of Transportation Engineering, ASCE*, February 2007;2007:112–118.
- Persaud B, Lyon C, Nguyen T. Empirical Bayes procedure for ranking sites for safety investigation by potential for safety improvement. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1665 1999:7–12.
- Qin X, Ivan JN, Ravishanker N. Selecting exposure measures in crash rate prediction for two-lane highway segments. *Accident Analysis and Prevention* 2004;36(3):183–191. [PubMed: 14642873]
- Saccomanno FF, Fu L, Miranda-Moreno LF. Risk-based model for identifying high-rail grade crossing blackspots. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1862 2004:127–135.
- Saccomanno FF, Park PYJ, Fu L. Estimating countermeasure effects for reducing collisions at highway-railway grade crossings. *Accident Analysis and Prevention* 2007;39:406–416. [PubMed: 17087912]
- Shankar VN, Chayanon S, Sittikariya S, Shyu MB, Juvva NK, Milton JC. Marginal impacts of design, traffic, weather, and related interactions on roadside crashes. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1897 2004:156–163.
- Shankar V, Mannering F. An exploratory multinomial logit analysis of single-vehicle motorcycle accident severity. *Journal of Safety Research* 1996;27(3):183–194.
- Shankar V, Mannering F, Barfield W. Statistical analysis of accident severity on rural freeways. *Accident Analysis and Prevention* 1996;28(3):391–401. [PubMed: 8799444]
- Shankar V, Milton J, Mannering F. Modeling accident frequencies as zero-altered probability processes: An empirical inquiry. *Accident Analysis and Prevention* 1997;29(6):829–837. [PubMed: 9370019]
- Song JJ, Ghosh M, Miaou S, Mallick B. Bayesian multivariate spatial models for roadway traffic crash mapping. *Journal of Multivariate Analysis* 2006;97:246–273.
- Stokes, ME.; Davis, CS.; Koch, GG. *Categorical Data Analysis Using the SAS System*. 2. SAS Institute, Inc; Cary, NC: 2000.
- Vogt, A. *Crash Models for Rural Intersections: Four-lane by Two-lane Stop-Controlled and Two-lane by Two-lane Signalized*. FHWA, U.S. Department of Transportation; Washington, D.C: 1999. FHWA-RD-99–128
- Vogt, A.; Bared, J. *Accident Prediction Models for Two-lane Rural Roads: Segments and Intersections*. FHWA, U.S. Department of Transportation; Washington, D.C.: 1998. FHWA-RD-98–133
- Yamamoto T, Shankar V. Bivariate ordered-response probit model of driver's and passenger's injury severities in collisions with fixed objects. *Accident Analysis and Prevention* 2004;36(5):869–876. [PubMed: 15203364]
- Yau KKW. Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong. *Accident Analysis and Prevention* 2004;36(3):333–340. [PubMed: 15003577]
- Zhang C, Ivan JN. Effects of geometric characteristics on head-on crash incidence on two roads in Connecticut. *Transportation Research Record: Journal of the Transportation Research Board*, No. 1908 2005:159–164.

Fig. 1.

Plot of predicted probabilities of different accident severity levels versus the number of daily trains (in the true range of 18–240 trains). The train speed, number of daily trucks, and crossing angle were held at their sample mean and the highway type was held at county road, highway width at being dichotomized to be 1, highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present.

Fig. 2.

Plot of predicted probabilities of different accident severity levels versus the number of daily trains (in the hypothetical range of 18–400 trains). The train speed, number of daily trucks, and crossing angle were held at their sample mean and the highway type was held at county road, highway width at being dichotomized to be 1, highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present.

Fig. 3. Marginal effects on the probability of accident severity versus the number of daily trains (in the true range of 18–240 trains). The train speed, number of daily trucks, and crossing angle were held at their sample mean and the highway type was held at county road, highway width at being dichotomized to be 1, highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present.

Fig. 4. Marginal effects on the probability of accident severity versus the number of daily trains (in the hypothetical range of 18–400 trains). The train speed, number of daily trucks, and crossing angle were held at their sample mean and the highway type was held at county road, highway width at being dichotomized to be 1, highway separation at marking separation, obstacle detection device at being active, guidance sign 34, approaching crossing markings, stop line markings, and law enforcement camera at being present.

Table 1Observed accident severity frequency and frequency of $Y = y$

Accident severity	Observed frequency (%)	$Y = y$	Observed frequency (%)
0	472 (79.93%)	0	472 (79.93%)
1	84 (14.19%)	1	84 (14.19%)
2	33 (5.57%)	2	36 (6.08%)
3	3 (0.51%)		

Table 2

Results from the generalized logit model

Category	Effect	J^a	MLE (SE) ^b	Wald χ^2 ^c	DF	p-value ^d
Railway feature	No. of daily trains	2	0.012 (0.006)	13.932	2	0.0009
		1	0.012 (0.004)			
	Train speed	2	0.028 (0.015)	4.995	2	0.0823
		1	-0.008(0.009)			
Highway feature	Highway type					
	County road	2	0.413 (1.225)	7.337	8	0.5008
		1	1.030 (0.790)			
	Provincial highway	2	1.473(1.589)			
		1	1.567(1.093)			
	Rural road	2	0.869(1.108)			
		1	0.528(0.743)			
	Urban street	2	0.723(1.220)			
		1	1.267(0.813)			
	Highway width					
Crossing feature	> 3.5 meters	2	-0.651(0.434)	5.151	2	0.0761
		1	-0.607(0.325)			
	Highway separation					
	Marking separation	2	1.858(0.489)	19.076	4	0.0008
		1	0.676(0.337)			
	Divided highways	2	1.934(0.959)			
		1	-0.266(0.754)			
	Daily trucks/100	2	0.013(0.046)	8.975	2	0.0112
		1	0.082(0.029)			
	Crossing angle	2	0.003(0.014)	0.876	2	0.6453
		1	-0.008(0.009)			
	Obstacle detection device					
	Yes	2	2.761(0.774)	12.730	2	0.0017

Category	Effect	j^a	MLE (SE) b	Wald χ^2 c	DF	p-value d
Traffic controls	Guidance sign 34	1	0.834(1.107)			
	Yes	2	0.605(0.693)	3.628	2	0.163
		1	-0.778(0.492)			
	Approaching crossing markings					
	Yes	2	0.588(0.456)	15.998	2	0.0003
		1	-1.089(0.307)			
	Stop line markings					
	Yes	2	0.623(0.459)	4.826	2	0.0895
		1	0.651(0.337)			
	Law enforcement camera					
	Yes	2	-1.699(0.844)	5.538	2	0.0627
		1	-1.392(0.689)			

a_j : an index for the j th logit: $\text{logit}[\pi_j(x)] = \log(\pi_j(x)/\pi_0(x)) = a_j + \sum_{j=1}^2 x_j \beta_j$, $j=1, 2$.
The MLE (SE) of the intercept parameter of the 1st logistic regression model was -0.574 (1.896); p -value of Wald-type chi-squared test for intercept parameter equal to zero was 0.7619.
The MLE (SE) of the intercept parameter of the 2nd logistic regression model was -7.768 (2.951); p -value of Wald-type chi-squared test for intercept parameter equal to zero was 0.0085.

b MLE: maximum likelihood estimate; SE: standard error.

c Wald-type chi-squared statistic of testing for parameters equal to zero.

d p -value of Wald-type chi-squared test for parameters equal to zero.

Table 3

Odds ratio (OR) estimates

Category	Effect	J^a	OR (95% CI)
Railway feature	No. of daily trains	2	1.012 (1, 1.023)
		1	1.012 (1.005, 1.02)
	Train speed	2	1.028 (0.999, 1.058)
		1	0.992 (0.974, 1.01)
Highway feature	Highway type		
	County road vs. Farm road	2	1.511 (0.137, 16.663)
		1	2.801 (0.595, 13.176)
	Provincial highway vs. Farm road	2	4.363 (0.194, 98.172)
		1	4.79 (0.563, 40.785)
	Rural road vs. Farm road	2	2.384 (0.272, 20.926)
		1	1.695 (0.395, 7.279)
	Urban street vs. Farm road	2	2.061 (0.188, 22.534)
		1	3.551 (0.721, 17.479)
	Highway width		
	> 3.5 meters vs. ≤ 3.5 meters	2	0.522 (0.223, 1.222)
		1	0.545 (0.288, 1.031)
	Highway separation		
	Marking Separation vs. No separation	2	6.411 (2.46, 16.707)
		1	1.966 (1.015, 3.808)
	Divided highways separation vs. No separation	2	6.918 (1.055, 45.352)
		1	0.767 (0.175, 3.36)
Daily trucks/100	2	1.013 (0.925, 1.109)	
	1	1.085 (1.026, 1.148)	
Crossing feature	Crossing angle	2	1.003 (0.976, 1.03)
		1	0.992 (0.974, 1.01)
	Obstacle detection device		
	Yes vs. No	2	15.818 (3.471, 72.094)
1		2.303 (0.263, 20.156)	
Traffic controls	Guidance sign ³⁴		
	Yes vs. No	2	1.83 (0.471, 7.113)
		1	0.459 (0.175, 1.204)
	Approaching crossing markings		
	Yes vs. No	2	1.8 (0.736, 4.401)
		1	0.337 (0.184, 0.615)
	Stop line markings		
	Yes vs. No	2	1.864 (0.758, 4.584)
1		1.918 (0.991, 3.712)	

Category	Effect	j^a	OR (95% CI)
	Law enforcement camera		
	Yes vs. No	2	0.183 (0.035, 0.956)
		1	0.249 (0.064, 0.959)

j^a : an index for the j th logit: $\text{logit}[\pi_j(\mathbf{x})] = \log(\pi_j(\mathbf{x})/\pi_0(\mathbf{x})) = \alpha_j + \mathbf{x}\beta_j$, $j=1, 2$.