

Research Article

Exploring the Application of the Linear Poisson Autoregressive Model for Analyzing the Dynamic Impact of Traffic Laws on Fatal Traffic Accident Frequency

Yue Zhang,¹ Yajie Zou ,¹ Lingtao Wu,² Jinjun Tang ,³ and Malik Muneeb Abid^{1,4}

¹Key Laboratory of Road and Traffic Engineering of Ministry of Education, Tongji University, No. 4800 Cao'an Road, Shanghai 201804, China

²Texas A&M Transportation Institute, Texas A&M University System, 3135 TAMU, College Station 77843, TX, USA

³School of Traffic and Transportation Engineering, Key Laboratory of Smart Transport in Hunan Province, Central South University, Changsha 410075, China

⁴Department of Civil Engineering, College of Engineering and Technology, University of Sargodha, Sargodha, Pakistan

Correspondence should be addressed to Yajie Zou; yajiezou@hotmail.com and Jinjun Tang; jinjuntang@csu.edu.cn

Received 3 July 2020; Revised 10 September 2020; Accepted 20 September 2020; Published 9 October 2020

Academic Editor: Hailong Liu

Copyright © 2020 Yue Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Annual fatal traffic accident data often demonstrate time series characteristics. The existing traffic safety analysis approaches (e.g., negative binomial (NB) model) often cannot accommodate the dynamic impact of factors in fatal traffic accident data and may result in biased parameter estimation results. Thus, a linear Poisson autoregressive (PAR) model is proposed in this study. The objective of this study is to apply the PAR model to analyze the dynamic impact of traffic laws and climate on the frequency of fatal traffic accidents occurred in a large time span (from 1975 to 2016) in Illinois. Besides, the NB model, NB with a time trend, and autoregressive integrated moving average model with exogenous input variables (ARIMAX) are also developed to compare their performances. The important conclusions from the modelling results can be summarized as follows. (1) The PAR model is more appropriate for analyzing the dynamic impacts of traffic laws on annual fatal traffic accidents, especially the instantaneous impacts. (2) The law that allows motorcycles and bicycles to proceed on a red light following the rules applicable after a “reasonable period of time” leads to an increase in the frequency of annual fatal traffic accidents by 14.98% in the short term and 30.69% in the long term. The climate factors such as average temperature and precipitation concentration period have insignificant impacts on annual fatal traffic accidents in Illinois. Thus, the modelling results suggest that the PAR model is more appropriate for annual fatal traffic accident data and has an advantage in estimating the dynamic impact of traffic laws.

1. Introduction

A report of National Highway Traffic Safety Administration (NHTSA) reveals that 37,461 people were killed in 34,439 motor vehicle crashes, an average of 102 deaths per day in year 2016. To reduce the number of people killed in traffic accidents, it is important to analyze the influential factors affecting the frequency of fatal traffic accidents. Among many factors, the traffic law is considered an effective measure to reduce the severity of injuries and the number of fatalities as a means of macroeconomic regulation and control. Some existing studies have analyzed the impact of

certain traffic laws on the number of traffic accidents, such as the seat belt law [1], driving under the influence (DUI) law [2, 3], and alcohol control law [4]. However, the dynamic impacts of these traffic laws on traffic accidents have not been adequately studied.

So far, numerous traffic safety analysis models have been developed. Since the frequency of traffic accident is non-negative and integer, many studies assumed such events follow a Poisson distribution and modelled the frequency of traffic accidents using a Poisson regression model [5, 6]. However, the Poisson model cannot handle overdispersed or underdispersed data and may result in biased estimation. In

order to analyze the overdispersed data, many studies proposed different mixed Poisson models, such as the Poisson-gamma model (the negative binomial (NB) model) [7–13], Poisson-lognormal model [14–16], and Poisson-inverse gamma model [17]. For the data with many zeros (i.e., excess zero-count data), the zero-inflated models were applied, including the zero-inflated Poisson model [18, 19], zero-inflated negative binomial model [20–22], and their extension models (i.e., multiple random parameter zero-inflated negative binomial regression model [20] and zero expansion Poisson regression model with random parameter effect [23]). Although rare, crash data can sometimes be characterized by underdispersion. The Conway–Maxwell–Poisson model [24] and diagonal inflated bivariate Poisson regression model [25] were appropriate.

Several recent works about regression models cannot properly address the time series characteristics of the traffic accident count data. Noland et al. [26] proposed a time trend variable as an explanatory variable in the count regression model to consider the series correlation. However, this model may not clearly account for the effects of serial correlation. An alternative approach was modelling possible dynamics in the traffic accident count data with a lagged dependent variable in the Poisson or NB models. These models failed to represent adequately the dynamics in persistent time series because they implied that the growth rate of the process was the exponentiated coefficient on the lagged dependent variable. Such a process may potentially generate time series data rather than dynamic data [27]. These two kinds of models were dynamic models with a trend, but not necessarily a cyclical or dynamic component. Another approach to handle the time series was the autoregressive integrated moving average (ARIMA) model and its extensions including the seasonality autoregressive integrated moving average (SARIMA) [28] and nonlinear autoregressive exogenous (NARX) [29]. These time series models may not be applicable to discrete time series variables (e.g., traffic accident count data). To consider both time series and discrete characteristics of the response variables, an integer-valued autoregressive (INAR) Poisson model [30–34] was developed. However, the dynamic characteristics of the influential factors were not adequately described in the INAR model. Few approaches can adequately model the dynamics and distribution of annual fatal traffic accident data.

To address the above issues, this study proposed a linear Poisson autoregressive (PAR) model. The objective of this study is to apply the PAR model to analyze the dynamic impact of traffic laws on annual fatal traffic accident frequency from the macroscopic point of view using the data collected in Illinois from year 1975 to 2016. The contribution is to demonstrate the performance of the PAR model in the analysis of the dynamic influence of factors on traffic accident frequency and quantitatively analyze the impacts of traffic laws.

The rest of the paper is organized as follows. Section 2 introduces the specification, estimation, and interpretation of the PAR (p) model. Section 3 describes the dataset used in

this study and the source of our data. In Section 4, the results of statistical modelling are shown to understand the contribution of different factors to the annual fatal traffic accidents in Illinois and compare the performances of various models. A conclusion and future recommendations are provided in Section 5.

2. Methodology

2.1. The PAR Model. Before presenting the model, the linear autoregressive (AR) process is firstly introduced. The AR model describes the random variables at some time by using the linear combination of random variables at earlier time as equation (1). It is a common form of time series.

$$y_t = \sum_{i=1}^p \rho_i y_{t-i} + \lambda, \quad (1)$$

where y_t is the traffic accident count value at time t , y_{t-i} is the past traffic accident count value at the i moments before time t , $\rho_i \in [-1, 1]$ is autocorrelation coefficient, and λ is a random error term.

Because there are explanatory variables in the PAR model, it is necessary to redefine the variables in the AR process. The conditional data Y_{t-i} replace y_{t-i} in the AR (p) model, which is a vector that included all the observed values of the dependent and independent variables at time t .

$$Y_{t-1} = (y_0, y_1, \dots, y_{t-1}; X_0, X_1, \dots, X_{t-1}), \quad (2)$$

where X_{t-i} are the past explanatory variables (factors affecting traffic accidents) at the i moments before time t . Y_{t-1} can be regarded as all the prior information about the series of interest at time t . Assume that y_t is a realization from a Markov process with the conditional transition probability $\Pr(y_t | Y_{t-1})$ and $E[Y_0] = \mu$. Let the conditional expectation $E[y_t | Y_{t-1}] = m_t$ at time t have a finite mean. Then, y_t is a p th order linear autoregressive process as shown in the following equation:

$$E[y_t | Y_{t-1}] = \sum_{i=1}^p \rho_i Y_{t-i} + \lambda. \quad (3)$$

Then, we can obtain equations (4) and (5) by using iterated expectations [35] of equation (3):

$$E[E[y_t | Y_{t-1}]] = E\left[\sum_{i=1}^p \rho_i Y_{t-i} + \lambda\right], \quad (4)$$

$$E[Y_t] = \sum_{i=1}^p \rho_i E[Y_{t-i}] + \lambda, \quad (5)$$

where equation (5) is a geometric series for ρ_i ; then,

$$\lim_{t \rightarrow \infty} E[Y_t] = \frac{\lambda}{(1 - \sum_{i=1}^p \rho_i)} \equiv \mu. \quad (6)$$

Since $[Y_0] = \mu$, equation (3) can be written as

$$E[y_t | Y_{t-1}] = \sum_{i=1}^p \rho_i Y_{t-i} + \left(1 - \sum_{i=1}^p \rho_i\right) \mu. \quad (7)$$

This is a linear AR(p) process, where the distribution of y is not used. The only role of distribution of y_t is to define the possible value of ρ_i .

The PAR (p) model can be defined as follows. The assumptions of the model are that the observed traffic accident counts y_t ($t = 1, 2, \dots, T$) are generated from a Poisson distribution on the condition of m_t . Then, the measurement equation for the observation value is obtained in the following equation:

$$\Pr(y_t | m_t) = \frac{m_t^{y_t} e^{-m_t}}{y_t!}. \quad (8)$$

Assume that m_t is the conditional mean of the linear AR process of $E[y_t | Y_{t-1}]$, which defines the state variable for the model. According to the measurement equation (8), this state density with its mean m_t and variance σ_t is in the exponential family. The mean or state variable of the marginal Poisson distribution evolves according to a stationary AR (p) process with autocorrelation parameters ρ_i , and its transition equation is expressed as follows:

$$m_t = \sum_{i=1}^p \rho_i y_{t-i} + \left(1 - \sum_{i=1}^p \rho_i\right) \exp(X_t \delta), \quad (9)$$

where X_t are $T \times k$ -matrices of explanatory variables composed of various factors affecting traffic accidents, δ is a $k \times 1$ -vector of regression parameters, and k is the number of factors.

Finally, assume that the density of the state variable has a gamma-distributed conjugate prior with gamma's location parameter a_{t-1} and scale parameter b_{t-1} , so

$$\Pr(m_t | Y_{t-1}) = \Gamma(\sigma_{t-1} m_{t-1}, \sigma_{t-1}), \quad m_{t-1} > 0, \sigma_{t-1} > 0, \quad (10)$$

where $m_t = E[y_t | Y_{t-1}]$ and $\sigma_{t-1} = \text{Var}[y_t | Y_{t-1}]$. The prior is constructed using the observed traffic accidents data. The prior can help to find the conditional mean and variance at time t based on the previous $t - 1$ observations.

Since the prior is gamma, using an extended Kalman filter, the conditional distribution at time t given $t - 1$ is also gamma, that is, $m_{t|t-1} \sim \Gamma((m_{t|t-1} \sigma_{t|t-1}), \sigma_{t|t-1})$. Since the measurement equation is Poisson and the state equation is gamma, an estimate of the posterior at time t is equation (11). This is a negative binomial distribution.

$$\begin{aligned} \Pr(y_t | Y_{t-1}) &= \int_{\theta} \Pr(y_t | \theta_t) \Pr(\theta_t | Y_{t-1}) d\theta \\ &= \int_{\theta} \frac{\theta_t^{y_t} e^{-\theta_t}}{y_t!} \cdot \frac{e^{-\sigma_{t|t-1} \theta} \theta^{\sigma_{t|t-1} m_{t|t-1} - 1} \sigma_t^{\sigma_{t|t-1} m_{t|t-1}}}{\Gamma(\sigma_{t|t-1} m_{t|t-1})} \\ &= \frac{\Gamma(\sigma_{t|t-1} m_{t|t-1} + y_t)}{\Gamma(y_t + 1) \Gamma(\sigma_{t|t-1} m_{t|t-1})} (\sigma_{t|t-1})^{\sigma_{t|t-1} m_{t|t-1}} \\ &\quad \times (1 + \sigma_{t|t-1})^{-(\sigma_{t|t-1} m_{t|t-1} + y_t)}. \end{aligned} \quad (11)$$

Replace the AR (1) process with m_t and obtain the PAR (1) model with a negative binomial predictive distribution. The one-step ahead conditional forecast function for the PAR (p) model is expressed as follows:

$$\begin{aligned} E[y_{t+1} | Y_t] &= m_{t+1|t} \\ &= \sum_{i=1}^p \rho_i m_{t|t-1} + \left(1 - \sum_{i=1}^p \rho_i\right) \mu. \end{aligned} \quad (12)$$

2.2. The Impact Multipliers. Because the PAR (p) model considers an explanatory variable matrix X_t and $\mu = \exp(X_t \delta)$, the interpretation differs from the Poisson and NB models significantly [36]. There is a concept of impact multiplier as in a Gaussian linear autoregressive model, which is the effect of a change in explanatory variable X_t . The instantaneous impact multiplier can be obtained by calculating the first derivative of the mean function for this change. The calculation process is shown in the following equation:

$$\begin{aligned} \frac{\partial m_t}{\partial X_t} &= \frac{\partial (\sum_{i=1}^p \rho_i Y_{t-i} + (1 - \sum_{i=1}^p \rho_i) \exp(X_t \delta))}{\partial X_t} \\ &= \left(1 - \sum_{i=1}^p \rho_i\right) \exp(X_t \delta) \cdot \delta, \end{aligned} \quad (13)$$

where δ is the coefficient of the explanatory variable X_t . This is the instantaneous effect of a shock in factors affecting traffic accidents X_t on the mean of traffic accidents m_t . The long-run impact multiplier which means the total effect of a shock to X_t can be calculated by equation (14), as in the Gaussian time series analysis. The long-run multiplier can be compared with the parameter estimation of other count regression models, which measures the impact of a shock on the conditional mean number of events.

$$\begin{aligned} \frac{\partial m_t / \partial X_t}{(1 - \sum_{i=1}^p \rho_i)} &= \frac{(1 - \sum_{i=1}^p \rho_i) \exp(X_t \delta) \cdot \delta}{(1 - \sum_{i=1}^p \rho_i)} \\ &= \exp(X_t \delta) \cdot \delta. \end{aligned} \quad (14)$$

The long-run impact multiplier and the instantaneous impact multiplier correspond to the concepts of average impact and marginal impact in economics, respectively. In economics, the average impact corresponds to the whole time, while the marginal impact corresponds to the "present" in time. The long-run impact is compared with the whole of the past, which should consider the impact of the past. However, the instantaneous impact multiplier or the marginal impact (the marginal effect is obtained by the partial derivation of X_t as shown in equation (13)) focuses on the impact on the future without considering the past.

In the Poisson and NB models, the long-run and instantaneous impact multipliers are the same and they can also be calculated by equation (14). The reason for this difference is

that the PAR (p) model accounts for the influence of explanatory variables on traffic accidents and the dynamic responses to the changes in explanatory variables over time.

3. Data Description

The annual fatal traffic accident frequency data for Illinois from year 1975 to 2016 were obtained from NHTSA's Fatality Analysis Reporting System (FARS). The annual fatal traffic accident was taken as the dependent variable to avoid underreporting due to various definitions of traffic accidents. Some traffic laws were considered to evaluate the instantaneous and long-run impact on traffic accidents from a quantitative perspective. According to WHO [37], five road safety risk factors (i.e., speeding, drunk-driving, the use of helmets, seatbelts, and child restraint systems) played an important role in traffic injuries and deaths. And Senna et al. [38] concluded that driving under the influence of alcohol is always a dominant problem. Due to the far-reaching influence of law on traffic accidents, the research data of law are basically based on the year [1, 39]. Thus, the traffic laws in Table 1 were selected for analysis of annual fatal traffic accidents. The variables related to law were set as binary variables here. For example, the initial effective date of safety belt law in Illinois was January 1, 1988, and then this traffic law variable equalled 0 for the first 13 periods and 1 thereafter (every traffic law used in this study and their effective date are shown in Table 1). If the date of implementation of the law was in the latter half of the year, the law was considered to work from the second year.

Besides, various factors identified as related to traffic safety are composed of our dataset to analyze how the explanatory variables affect the annual fatal traffic accident. In order to be consistent with legal factors, we also select some macroscopic indicators ranging from 1975 to 2016 with an interval of year in Illinois. The dataset is assembled from a variety of sources including the U.S. Energy Information Administration, Federal Highway Administration, and National Institute on Alcohol Abuse and Alcoholism. The dataset covers economic, social, driver, climate, and law factors. Summary statistics of the dataset are shown in Table 1. Note that the climate factors mentioned are different from the microscopic weather factor, which indicates the wet and dry conditions over a large space-time range. In the dataset, gross domestic product (GDP), total vehicle miles of travel (VMT), rural VMT as proportion of total VMT, per capita beer consumption, gasoline price, and safety belt law are used to analyze the influence of various factors on traffic accidents [3]. Geedipally et al. [3] demonstrated that DUI laws, beer consumption, the proportion of rural VMT, and shocks in the economy had a significant effect in traffic fatalities. Note that all the economic indicators are converted to 2016 dollars using the consumer price index (CPI) calculator at the Bureau of Labor Statistics.

4. Modelling Results

Four models (i.e., the ARIMAX, PAR, NB, and NB with a time trend model) are developed using the Illinois data, where the ARIMAX, NB, and NB with a time trend model

are considered as benchmark models. All the models constructed in this paper are implemented with R. The main purpose is to identify the influential traffic laws affecting the fatal traffic accidents. Figure 1 shows the trend of the fatal traffic frequency accident and VMT over time. As can be seen from Figure 1, the frequency of fatal traffic accident decreases significantly with time and shows sequence correlation, while VMT increases with time.

For the PAR(p) model mentioned in the methodology, the order p of the PAR (p) model is determined firstly (Table 2). Based on the PAR (1) model, stepwise regression is used to select all significant variables as the combination of explanatory variables of each model. Thus, the final models include only a subset of the original explanatory variables, which is shown in Tables 2 and 3. The Akaike information criterion (AIC) which is an estimator of the relative quality of statistical models for a given set of data provided another means for order selection. The smaller its value is, the better the fitting effect of the model is. When $p > 4$, the model complexity increases but AIC does not decrease significantly. As discussed by Eluru et al. [40], different from the AIC, the Bayesian information criterion (BIC) imposes a larger penalty on model overfitting with excess parameters. As can be seen from Table 2, the BIC value for the PAR (4) and PAR (5) models differ slightly because the PAR (5) model has more parameters than the PAR (4) model. Besides, the estimated insignificant parameters are shown in bold in Table 2. There are insignificant parameters from $p > 2$, which will affect the analysis of the impact multiplier of explanatory variables. Based on the modelling results in Table 2, the PAR (2) model is chosen as the final model.

In addition to the PAR (2) model, the ARIMAX, NB, and NB with a time trend models are also compared as alternative models. Because of the time series in the data, the ARIMAX model considering the explanatory variables is selected as one of the alternative models. Based on AIC values, the final model is determined as ARIMAX (1, 1, 0). Besides, the NB model, which is most commonly used in traffic accident frequency analysis, is considered as one of the alternative models. Furthermore, the NB with a time trend model which can consider the time series and discrete characteristic of the traffic accident frequency by using a simple solution is also compared. The parameter values of these models are estimated using the maximum likelihood estimation method.

The results of these models are shown in Table 3. According to the AIC and BIC values in Table 3, the PAR (2) model fits this dataset best, followed by the NB with a time trend model and ARIMAX (1, 1, 0) model. The NB model provides the least fitting performance. The models with a time trend structure seem to fit best since the traffic accidents appear to be serially correlated. Taking AIC as example, the fitting performance of the PAR model increases by 12% compared to the NB model, 6% compared to the ARIMAX model, and 5% compared to the NB with a time trend model. However, the coefficients of explanatory variables estimated by the ARIMAX model are not significant; in other words, the ARIMAX model is not able to explain how these factors affect annual fatal traffic accidents. For the

TABLE 1: Summary statistics of annual fatal traffic accidents and factors in Illinois.

Variables	Min	Max	Mean	St. error
Fatal traffic accidents	832	1877	1316	299.31
Total VMT (million)	60943	109135	89696	16914.83
GDP (in \$100,000)	405739	807458	593554	134797
Rural VMT proportion (%)	0.24	0.34	0.30	0.03
Gasoline price (\$ per gallon)	1.54	3.71	2.41	0.64
Per capita beer consumption (gallons)	1.12	1.45	1.30	0.10
Average temperature (°F)	49.38	55.78	52.16	1.52
Precipitation concentration period	28.54	51.18	39.68	5.68
Safety belt law	January 1988		Before: 0; after: 1 (year 1988: 1)	
DUI toughened penalties	January 1982		Before: 0; after: 1 (year 1982: 1)	
Drivers under age 18 may not use a wireless phone while driving	July 2005		Before: 0; after: 1 (year 2005: 0)	
Persons under age 18 who obtain a graduated driver’s license may not drive during the first 6 months of the license, or until the person reaches age 18, with more than one person under age 20 in the vehicle	January 2004		Before: 0; after: 1 (year 2004: 1)	
Disallow parents/guardians to knowingly permit the consumption of alcohol by underage invitees at their residence	August 2007		Before: 0; after: 1 (year 2007: 0)	
Motorcycles and bicycles may proceed on a red light following the rules applicable after a “reasonable period of time”	January 2012		Before: 0; after: 1 (year 2012: 1)	
Graduated driver’s license provisions strengthened	January 2008		Before: 0; after: 1 (year 2008: 1)	
Limit the use of mobile phones while operating a motor vehicle	January 2010		Before: 0; after: 1 (year 2010: 1)	

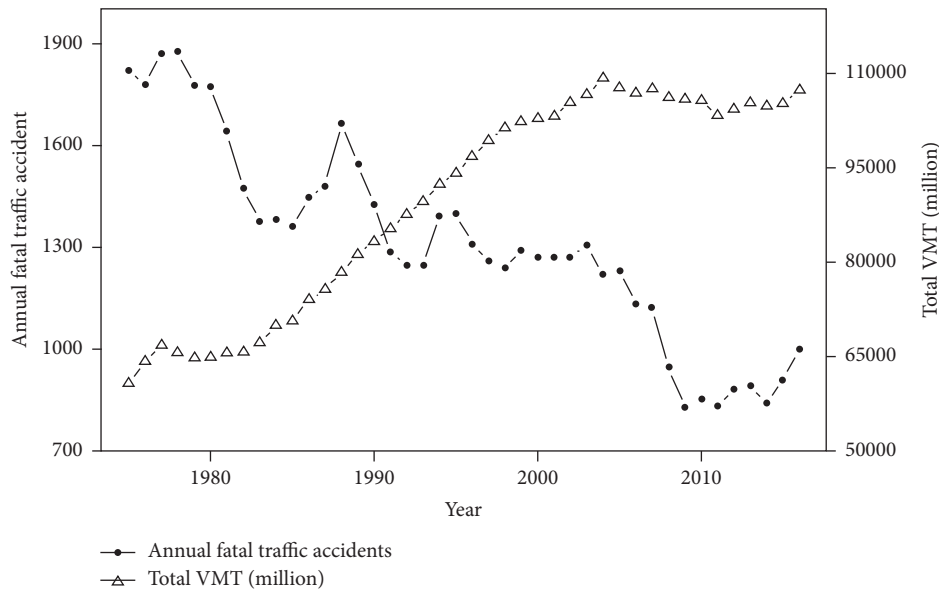


FIGURE 1: Annual fatal traffic accidents and VMT in Illinois.

NB with a time trend model, most of the explanatory variables are significant and the values of parameter estimates are similar to those of the NB and PAR (2) models. However, the belt law and beer consumption variables are insignificant. This phenomenon shows that this model may not help us to explain the impact of various legal factors on the frequency of traffic accidents completely. Since the PAR (2) model and NB model have all statistically significant variables, their fitting performance and parameter estimates are compared in the following paragraphs.

Except the ARIMAX model, other models belong to the regression model. During modelling the fatal traffic accidents, total VMT is considered as an offset term because there is a linear relationship between total VMT and fatal traffic accidents [3]. From a qualitative point of view, both the coefficients estimated by the NB and PAR models show that beer consumption has the greatest impact among these factors on the frequency of annual fatal traffic accidents (Table 3). However, the AIC and BIC values of the PAR (2) model are much smaller than those of the NB model. Note

TABLE 2: Modelling results for different orders of PAR models.

Model	PAR (1)	PAR (2)	PAR (3)	PAR (4)	PAR (5)
Intercept	-6.216 (0.375)	-6.821 (0.328)	-6.811 (0.322)	-7.028 (0.341)	-6.787 (NA)
Gasoline price	-0.152 (0.038)	-0.074 (0.028)	-0.074 (0.029)	-0.101 (0.037)	-0.124 (0.048)
Beer consumption	2.12 (0.244)	2.37 (0.204)	2.363 (0.197)	2.476 (0.209)	2.491 (0.225)
Belt law	-0.196 (0.046)	-0.157 (0.037)	-0.159 (0.038)	-0.118 (0.042)	-0.093 (0.049)
DUI toughened penalties	-0.379 (0.05)	-0.279 (0.043)	-0.278 (0.05)	-0.206 (0.076)	-0.451 (NA)
Red running	0.393 (0.083)	0.301 (0.059)	0.296 (0.065)	0.383 (0.089)	0.434 (0.113)
Alcohol law	-0.237 (0.067)	-0.219 (0.049)	-0.216 (0.052)	-0.247 (0.062)	-0.276 (0.075)
ρ_1	0.628 (0.057)	0.706 (0.093)	0.704 (0.095)	0.754 (0.095)	0.718 (0.098)
ρ_2	—	-0.194 (0.079)	-0.179 (0.119)	-0.218 (0.12)	-0.157 (0.127)
ρ_3	—	—	-0.015 (0.086)	-0.265 (0.118)	-0.262 (0.12)
ρ_4	—	—	—	0.287 (0.082)	0.144 (0.12)
ρ_5	—	—	—	—	0.147 (0.09)
N^1	42	42	42	42	42
Log-likelihood	-234.125	-222.116	-217.543	-200.139	-197.746
AIC	482.249	460.232	453.069	420.277	417.492
BIC	494.414	474.133	468.725	437.655	436.606

¹ N is the number of samples.

TABLE 3: Fatal traffic accident analysis (standard errors in parentheses).

Model	ARIMAX (1, 1, 0)	PAR (2)	NB	NB with a time trend
Intercept	—	-6.821 (0.328)	-5.835 (0.347)	7.118 (0.29)
Gasoline price	3.836 (32.897)	-0.074 (0.028)	-0.077 (0.028)	-0.05 (0.02)
Beer consumption	575.665 (514.852)	2.37 (0.204)	1.805 (0.242)	0.393 (0.211)
Belt law	191.826 (66.103)	-0.157 (0.037)	-0.225 (0.048)	0.029 (0.039)
DUI toughened penalties	-98.515 (77.537)	-0.279 (0.043)	-0.445 (0.048)	-0.227 (0.042)
Red running	58.631 (63.765)	0.301 (0.059)	0.162 (0.059)	0.096 (0.039)
Alcohol law	-144.166 (67.999)	-0.219 (0.049)	-0.221 (0.06)	-0.228 (0.041)
Linear trend	—	—	—	-0.007 (0.003)
ρ_1	—	0.706 (0.093)	—	—
ρ_2	—	-0.194 (0.079)	—	—
α^1	—	—	0.006	0.002
N^2	42	42	42	42
Log-likelihood	-237.85	-222.116	-254.633	-232.543
AIC	491.71	460.232	525.27	483.087
BIC	505.61	474.133	539.168	498.726

¹The dispersion parameter in the NB model. ² N is the number of samples.

that the PAR model can capture dynamics in fatal traffic accident data and provide better fitting performance. In addition to the goodness-of-fit statistics, we further compare the modelling results of the PAR (2) model and the NB model.

For dynamic models, the results cannot be fully obtained by displaying coefficients in Table 3, which represented the average effect of explanatory variables [41]. Besides, the generalized linear model includes a link function, which makes it difficult to explain the original coefficients of the model independently [41]. In other words, the estimated coefficients in Table 3 cannot directly quantify the impact of factors per unit change on fatal traffic accident frequency. Thus, in order to compare the effects of explanatory variables on fatal traffic accidents in different models, the long-run and instantaneous multipliers are calculated by using equations (13) and (14). Since this study focuses on the impact of traffic laws on annual fatal traffic accidents, only the impact multipliers of law are presented in the first

column of Tables 4 and 5. Because of the dynamic structure of the PAR (2) model, the value of long-run impact multiplier is larger than that of instantaneous impact multiplier ($0 < \sum_{i=1}^2 \rho_i < 1$). The three laws (e.g., belt law, DUI toughened penalties, and alcohol law) lead to a decrease in fatal traffic accident, among which the DUI toughened penalties law has the greatest influence. After the implementation of the DUI toughened penalties, the frequency of fatal traffic accident decreases by 91 in the short run and 186 in the long run. However, implementation of the law that allows motorcycles and bicycles to proceed on a red light following the rules applicable after a “reasonable period of time” (red running) leads to an increase in the frequency of fatal traffic accidents. Before this study, Pai and Jou [42] have revealed the high association between bicyclist red-running and accidents in Taiwan. The effect of red running is a total increase of 201 fatal traffic accidents in the long run and 72 fatal traffic accidents in the short run, which indicates that this law is not conducive to traffic safety. This law may be

TABLE 4: Fatal traffic accident analysis (long-run effect).

Variable	PAR (2)		NB	
	Long-run multiplier	Percentage change in mean (%)	Long-run multiplier	Percentage change in mean (%)
Belt law	-105.137	-15.00	-155.447	-20.15
DUI toughened penalties	-186.371	-25.65	-309.795	-35.92
Red running	201.172	30.69	112.659	17.59
Alcohol law	-146.495	-20.63	-153.839	-19.83

TABLE 5: Fatal traffic accident analysis (instantaneous effect).

Variable	PAR (2)		NB	
	Instantaneous multiplier	Percentage change in mean (%)	Instantaneous multiplier	Percentage change in mean (%)
Belt law	-51.322	-7.32	-155.447	-20.15
DUI toughened penalties	-90.975	-12.52	-309.795	-35.92
Red running	98.2	14.98	112.659	17.59
Alcohol law	-71.51	-10.07	-153.839	-19.83

designed to increase the traffic efficiency of nonmotorized vehicles and reduce travel time, but it is not desirable to improve traffic efficiency at the expense of traffic safety.

Finally, the impact multipliers of the NB and PAR (2) models are compared (Tables 4 and 5). All signs of parameters estimated by the NB and PAR models are the same. It can be observed that the frequency of fatal traffic accidents decreases with the increase of gasoline prices, the implementation of the belt law, and the enforcement of the DUI penalty law. With beer consumption rising and red running allowed, the frequency of fatal traffic accident increases. However, the values of impact multipliers estimated by the NB and PAR (2) models differ significantly. Taking the implementation of red running law as an example, which is the only law variable leading to an increase in the frequency of traffic accidents in the dataset, the instantaneous impact of the red running law is about 98 for the PAR (2) model, which means that the implementation of red running law has increased the number of accidents by 98 at this point. The long-run multiplier of red running law is 201, which means the frequency of accidents increased by 201 in the long run. The NB model estimates the impact of enforcing the red running law by a multiplier of 113. These percentage changes are shown in Tables 4 and 5. For the PAR (2) model, the total percentage change due to this intervention in the number of fatal traffic accidents is an increase of 30.69% while the instantaneous percentage change is 14.98%. For the NB model, the total change in the number of fatal traffic accidents is 17.59%. The instantaneous effect of the PAR (2) model is smaller than the instantaneous effect of the NB model, and long-run effects of the PAR (2) model are larger than the long-run effect of the NB model. The reason for this phenomenon is that the estimated coefficients of the PAR (2) model include dynamic characteristics. The long-run multiplier takes into account the impact of the previous stage on the present, while the instantaneous effect only takes into account the current impact. Since δ describes an average effect, the multipliers calculated from $\exp(X_t \delta) \cdot \delta$ can only describe the average effect. Thus, the instantaneous effect of

the NB model actually reflects the average effect, which leads to overestimation of the short-run effect. The NB model cannot properly consider the dynamics of fatal traffic accident data, which leads to overestimation of the instantaneous impact of explanatory variables on fatal traffic accidents.

For the remaining three variables, both the long-run and instantaneous impact multipliers estimated by the PAR model are smaller than those estimated by the NB model. Taking the belt law as an example, the instantaneous and long-run percentage changes due to the intervention estimated by the PAR model are -7.32% and -15%, respectively, and the estimated multiplier of the NB model is -20.15%. This phenomenon indicates again that the NB model overestimates the impact of explanatory variables, especially for the instantaneous impacts. The dynamic nature of the PAR (p) model makes it more suitable for estimating the dynamic impact of traffic laws on annual fatal traffic accidents. The instantaneous impact of a safety intervention strategy can inform the transportation management agencies to design more appropriate traffic laws, while the NB model cannot provide such information.

5. Conclusions

Annual fatal traffic accidents are count data with time series characteristics. The existing traffic accident analysis models cannot fully model their dynamic characteristics and analyze the dynamic influence of explanatory variables on annual fatal traffic accidents. Among many explanatory variables of traffic accident analysis, the dynamic effect of the enforcement of traffic laws has not been widely concerned. In this study, a linear Poisson autoregressive model is proposed to analyze the long-run and instantaneous impact of traffic laws on annual fatal traffic accidents. Then, the modelling results of PAR (p), ARIMAX, NB, and NB with a time trend models are compared. Several major conclusions are summarized as follows:

- (1) The PAR model can outperform the ARIMAX, NB, and NB with a time trend models in terms of fitting performance and estimation of dynamic effects. The PAR (p) model is more suitable for analyzing the dynamic impact of traffic laws on annual fatal traffic accidents. Compared with the ARIMAX model, the PAR (p) model can consider discrete characteristics in the accident data and analyze the influence of factors. Compared with the NB with a time trend model, the PAR (p) model can accurately analyze the influence of more explanatory variables on the frequency of fatal traffic accidents. Compared with the NB model, the PAR (p) model can capture the time series in annual fatal traffic accident frequency and calculate the dynamic effect of traffic laws and other explanatory variables. The omission of the dynamics from the NB model leads to biased parameter estimates, especially the inability to estimate the instantaneous multipliers of factors. However, instantaneous multipliers can indicate the immediate effects of traffic law interventions for traffic safety management agencies and help to make new laws.
- (2) Some climate and traffic law factors are considered to quantitatively evaluate their impact on annual fatal traffic accidents in Illinois. The average temperature and precipitation concentration period have insignificant impacts. The law of DUI toughened penalties results in a decrease of annual fatal traffic accidents by 12.52% in the short run and 25.65% in the long run, which has the greatest inhibitory effect on fatal traffic accidents among the analyzed laws. However, the law allowing red running leads to an increase of annual fatal traffic accidents by 14.98% in the short term and 30.69% in the long term. Therefore, controlling the DUI behaviors and modifying the red running law may significantly reduce the frequency of annual fatal traffic accidents, which provide guidance for future traffic law development.

The PAR (p) model can be widely applied to analyze the time series count data. Besides the traffic laws mentioned in this paper, the applicable explanatory variables exhibiting a sudden change can be extended to the factors such as the emergence of policies and regulations, and the dynamic impact of these kinds of variables can be well explained by the PAR (p) model. For future research, the PAR (p) model can be applied to investigate traffic accident data collected from other sites. Furthermore, with the development of data acquisition technology, multisource datasets [43–46] can be used to analyze traffic accidents.

Data Availability

The experimental data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was sponsored jointly by the National Key Research and Development Program of China (grant no. 2018YFE0102800), the Shanghai Science and Technology Committee (grant no. 19210745700), and the Fundamental Research Funds for the Central Universities (grant no. 22120200035).

References

- [1] J. Lv, D. Lord, Y. Zhang, and Z. Chen, "Investigating Peltzman effects in adopting mandatory seat belt laws in the US: evidence from non-occupant fatalities," *Transport Policy*, vol. 44, pp. 58–64, 2015.
- [2] V. R. Campos, R. d. S. E. Silva, S. Duailibi, J. F. D. Santos, R. Laranjeira, and I. Pinsky, "The effect of the new traffic law on drinking and driving in São Paulo, Brazil," *Accident Analysis & Prevention*, vol. 50, pp. 622–627, 2013.
- [3] S. R. Geedipally, "In-depth investigation of factors that contributed to the decline in fatalities from 2008 to 2012 in the United States," in *Proceedings of the Transportation Research Board 97th Annual Meeting*, Washington, DC, USA, 2018.
- [4] K. Chang, C.-C. Wu, and Y.-H. Ying, "The effectiveness of alcohol control policies on alcohol-related traffic fatalities in the United States," *Accident Analysis & Prevention*, vol. 45, pp. 406–415, 2012.
- [5] M. Hosseinpour, S. Sahebi, Z. H. Zamzuri, A. S. Yahaya, and N. Ismail, "Predicting crash frequency for multi-vehicle collision types using multivariate Poisson-lognormal spatial model: a comparative analysis," *Accident Analysis & Prevention*, vol. 118, pp. 277–288, 2018.
- [6] X. Ye, "A semi-nonparametric Poisson regression model for analyzing motor vehicle crash data," *PLoS One*, vol. 13, 2018.
- [7] C. Dong, "An improved deep learning model for traffic crash prediction," *Journal of Advanced Transportation*, vol. 2018, p. 13, 2018.
- [8] C. Ma, W. Hao, A. Wang, and H. Zhao, "Developing a coordinated signal control system for urban ring road under the vehicle-infrastructure connected environment," *Ieee Access*, vol. 6, pp. 52471–52478, 2018.
- [9] W. Ma and Z. Yuan, "Analysis and comparison of traffic accident regression prediction model," in *Proceedings 3rd International Conference on Electromechanical Control Technology and Transportation (ICECTT 2020)*, New York, NY, USA, 2018.
- [10] M. R. R. Shaon, X. Qin, M. Shirazi, D. Lord, and S. R. Geedipally, "Developing a random parameters negative binomial-lindley model to analyze highly over-dispersed crash count data," *Analytic Methods in Accident Research*, vol. 18, pp. 33–44, 2018.
- [11] K. Wang and X. Ye, *Development of Alternative Stochastic Frontier Models for Estimating Time-Space Prism Vertices*, Transportation, New York, NY, USA, 2020.
- [12] Y. Zou, "Empirical Bayes estimates of finite mixture of negative binomial regression models and its application to highway safety," *Journal of Applied Statistics*, vol. 14, pp. 1–18, 2017.

- [13] Y. Zou, L. Wu, and D. Lord, "Modeling over-dispersed crash data with a long tail: examining the accuracy of the dispersion parameter in Negative Binomial models," *Analytic Methods in Accident Research*, vol. 5-6, pp. 1-16, 2015.
- [14] M. Shirazi and D. Lord, "Characteristics based heuristics to select a logical distribution between the Poisson-gamma and the Poisson-lognormal for crash data modelling," *Transportmetrica A: Transport Science*, vol. 14, pp. 1-22, 2019.
- [15] K. Wang, S. Zhao, and E. Jackson, "Multivariate Poisson lognormal modeling of weather-related crashes on freeways," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2672, no. 38, pp. 184-198, 2018.
- [16] M. Zhao, "Multivariate Poisson-lognormal model for analysis of crashes on urban signalized intersections approach," *Journal of Transportation Safety & Security*, vol. 10, no. 3, pp. 251-265, 2018.
- [17] S. H. Khazraee, V. Johnson, and D. Lord, "Bayesian Poisson hierarchical models for crash data analysis: investigating the impact of model choice on site-specific predictions," *Accident Analysis & Prevention*, vol. 117, pp. 181-195, 2018.
- [18] V. N. Shankar, G. F. Ulfarsson, R. M. Pendyala, and M. B. Nebergall, "Modeling crashes involving pedestrians and motorized traffic," *Safety Science*, vol. 41, no. 7, pp. 627-640, 2003.
- [19] S. Unhapipat, M. Tiensuwan, and N. Pal, "Bayesian predictive inference for zero-inflated Poisson (ZIP) distribution with applications," *American Journal of Mathematical and Management Sciences*, vol. 37, no. 1, pp. 66-79, 2018.
- [20] C. Liu, M. Zhao, W. Li, and A. Sharma, "Multivariate random parameters zero-inflated negative binomial regression for analyzing urban midblock crashes," *Analytic Methods in Accident Research*, vol. 17, pp. 32-46, 2018.
- [21] M. A. Raihan, P. Alluri, W. Wu, and A. Gan, "Estimation of bicycle crash modification factors (CMFs) on urban facilities using zero inflated negative binomial models," *Accident Analysis & Prevention*, vol. 123, pp. 303-313, 2019.
- [22] H. Wang, Y. Yang, and Y. Ma, "Research on the Yangtze River accident casualties using zero-inflated negative binomial regression technique," in *Proceedings of the 2016 IEEE International Conference on Intelligent Transportation Engineering (ICITE)*, Berlin, Germany, 2016.
- [23] M. Haque, H. C. Chin, and H. Huang, "Modeling random effect and excess zeros in road traffic accident prediction," in *Proceedings of the 19th KKCNN Symposium on Civil Engineering*, Berlin, Germany, 2006.
- [24] D. Lord, S. R. Geedipally, and S. D. Guikema, "Extension of the application of conway-maxwell-Poisson models: analyzing traffic crash data exhibiting underdispersion," *Risk Analysis*, vol. 30, no. 8, pp. 1268-1276, 2010.
- [25] Y. Lao, Y.-J. Wu, J. Corey, and Y. Wang, "Modeling animal-vehicle collisions using diagonal inflated bivariate Poisson regression," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 220-227, 2011.
- [26] R. Noland, M. Quddus, and W. Ochieng, "The effect of the congestion charge on traffic casualties in London: an intervention analysis," in *Proceedings of the Transportation Research Board (TRB) Annual Meeting*, Washington, DC, USA, January 2006.
- [27] P. T. Brandt, J. T. Williams, B. O. Fordham, and B. Pollins, "Dynamic modeling for persistent event-count time series," *American Journal of Political Science*, vol. 44, no. 4, pp. 823-843, 2000.
- [28] S. Yousefzadeh-Chabok, "A time series model for assessing the trend and forecasting the road traffic accident mortality," *Archives of Trauma Research*, vol. 5, no. 3, 2016.
- [29] A. Sukhai, A. P. Jones, B. S. Love, and R. Haynes, "Temporal variations in road traffic fatalities in South Africa," *Accident Analysis & Prevention*, vol. 43, no. 1, pp. 421-428, 2011.
- [30] T. Brijs, D. Karlis, and G. Wets, "Studying the effect of weather conditions on daily crash counts using a discrete time-series model," *Accident Analysis & Prevention*, vol. 40, no. 3, pp. 1180-1190, 2008.
- [31] Q. Duan, X. Ye, J. Li, and K. Wang, "Empirical modeling analysis of potential commute demand for carsharing in shanghai, China," *Sustainability*, vol. 12, no. 2, p. 620, 2020.
- [32] C. Ma, R. He, and W. Zhang, "Path optimization of taxi carpooling," *PLoS One*, vol. 13, no. 8, 2018.
- [33] M. Quddus, "Time-eries regression models for analysing transport safety data," in *Safe Mobility: Challenges, Methodology and Solutions*, L. Dominique and W. Simon, Eds., pp. 279-296, Emerald Publishing Limited, Berlin, Germany, 2018.
- [34] Y. Sunecher, N. M. Khan, and V. Jowaheer, "Estimation methods for a flexible INAR (1) COM-Poisson time series model," *Journal of Applied Mathematics, Statistics and Informatics*, vol. 14, no. 1, pp. 57-82, 2018.
- [35] G. K. Grunwald, R. J. Hyndman, and L. M. Tedesco, "A unified view of linear AR (1) Models," 1995.
- [36] P. T. Brandt and J. T. Williams, "A linear Poisson autoregressive model: the Poisson AR(p) model," *Political Analysis*, vol. 9, no. 2, pp. 164-184, 2001.
- [37] W. H. O. D. O. Violence, *Global Status Report on Road Safety: Time for Action*, World Health Organization, Geneva, Switzerland, 2009.
- [38] M.-C. Senna, M. Augsburg, B. Aebi et al., "First nationwide study on driving under the influence of drugs in Switzerland," *Forensic Science International*, vol. 198, no. 1-3, pp. 11-16, 2010.
- [39] D. Mark Anderson, B. Hansen, and D. I. Rees, "Medical marijuana laws, traffic fatalities, and alcohol consumption," *The Journal of Law and Economics*, vol. 56, no. 2, pp. 333-369, 2013.
- [40] N. Eluru, M. Bagheri, L. F. Miranda-Moreno, and L. Fu, "A latent class modeling approach for identifying vehicle driver injury severity factors at highway-railway crossings," *Accident Analysis & Prevention*, vol. 47, pp. 119-127, 2012.
- [41] B. J. Fogarty and J. E. Monogan Iii., "Modeling time-series count data: the unique challenges facing political communication studies," *Social Science Research*, vol. 45, pp. 73-88, 2014.
- [42] C.-W. Pai and R.-C. Jou, "Cyclists' red-light running behaviours: an examination of risk-taking, opportunistic, and law-obeying behaviours," *Accident; Analysis and Prevention*, vol. 62C, pp. 191-198, 2013.
- [43] X. Yang, Y. Zou, J. Tang, J. Liang, and M. Ijaz, "Evaluation of short-term freeway speed prediction based on periodic analysis using statistical models and machine learning models," *Journal of Advanced Transportation*, vol. 2020, pp. 1-16, 2020.
- [44] X. Zhong, Y. Zou, Z. Dong, S. Yuan, and M. Ijaz, "Finite mixture survival model for examining the variability of urban arterial travel time for buses, passenger cars and taxis," *IET Intelligent Transport Systems*, vol. 45, 2020 In Press.
- [45] X. Chen, X. Xu, Y. Yang, H. Wu, J. Tang, and J. Zhao, "Augmented ship tracking under occlusion conditions from maritime surveillance videos," *IEEE Access*, vol. 8, pp. 42884-42897, 2020.
- [46] X. Chen, S. Wang, C. Shi, H. Wu, J. Zhao, and J. Fu, "Robust ship tracking via multi-view learning and sparse representation," *Journal of Navigation*, vol. 72, no. 1, pp. 176-192, 2018.