Hindawi
Journal of Advanced Transportation
Volume 2022, Article ID 2941327, 10 pages
https://doi.org/10.1155/2022/2941327

Research Article

Influence of Public Bus Driver's Driving Behaviors on Passenger Fall Incidents: An Analysis Using Digital Tachograph Data

Harim Jeong,[1] **Wonil Park**,[2] **Jaeyoung Lee**,[3,4] **Sungho Park**,[1] **and Ilsoo Yun**[1]

[1]*Department of Transportation System Engineering, Ajou University, Suwon, Republic of Korea*
[2]*Korea Bus Transportation Associations Federation, Seoul, Republic of Korea*
[3]*School of Traffic & Transportation Engineering, Central South University, Changsha, Hunan 410075, China*
[4]*Department of Civil, Environment and Construction Engineering, University of Central Florida, Orlando, FL 32816, USA*

Correspondence should be addressed to Sungho Park; fenix3339@ajou.ac.kr

Received 19 July 2022; Revised 25 November 2022; Accepted 30 November 2022; Published 22 December 2022

Academic Editor: Tiziana Campisi

Owing to the low occurrence of public transport accidents, most existing studies have focused on improving the traffic safety of passenger cars. However, traffic accidents related to public transport should also be investigated for the safety of public transport users, particularly the vulnerable ones. This study analyzed the behavioral factors affecting passenger fall incidents on buses to enhance the safety of public bus passengers. This study considered potential influential factors, such as acceleration, deceleration, braking, and steering maneuvers, calculated using data from digital tachographs installed on the buses. Negative binomial and Poisson lognormal regression models were built using the Bayesian method. The two models yielded similar results. In all cases, deceleration-related behavioral factors (abrupt deceleration and abrupt stop) significantly influenced passenger fall incidents. The findings from this study are significant for establishing effective strategies to reduce fall incidents by providing safety education to bus drivers.

1. Introduction

1.1. Background and Objectives. Public transport accidents are relatively rare compared with passenger car accidents, and the data collected on these accidents are typically insufficient. Thus, the associated risks tend to be evaluated as relatively low. Accordingly, only the positive effects of public transportation, such as traffic congestion alleviation and pollution reduction, are typically highlighted [1]. Many studies related to traffic accidents have focused on improving the traffic safety of passenger cars. However, traffic accidents related to public buses must also be considered. Many bus users belong to the vulnerable age group, such as children and the elderly. Accordingly, efforts have been made worldwide to reduce bus-related traffic accidents [2, 3].

Prior studies have focused on reducing the number of traffic accidents. Although the overall number of traffic accidents has decreased, the number of accidents involving buses has not decreased significantly. According to the results of a survey by the European Road Safety Observatory [4], the number of fatalities in bus accidents in the European Union member countries was 1,148 in 2007 and decreased to 594 in 2016. However, in 2007 and 2016, the number of fatalities in bus accidents accounted for 3% and 2% of all traffic accident fatalities, respectively. This indicates that the fatality rate in bus accidents did not decrease significantly. According to the Fatality and Injury Reporting System Tool statistics of the National Highway Traffic Safety Administration, in 2010, 11,544 bus accidents in the United States resulted in injuries, and 251 led to deaths. The corresponding values in 2019 were 13,678 and 232, indicating no significant change relative to the past.

Many studies have evaluated and analyzed crash risks in relation to bus accidents [2, 5–18]. However, most studies have focused on collision accidents. Studies that included noncollision traffic accidents did not distinguish between collision and noncollision accidents.

Standing passengers on public buses are affected by driver maneuvers, such as sudden acceleration and braking maneuvers, with the potential to cause a fall incident. On a bus, noncollision passenger fall incidents are more likely to occur owing to the bus driver maneuvers, such as sudden accelerations/decelerations or sudden turns. In particular, buses operating within a city have high a risk of passenger fall incidents because there are more standing passengers. Therefore, safety management in buses is more important than that in other vehicles. It is necessary to investigate the characteristics of events related to driver maneuvers (i.e., acceleration, braking, or turning). Moreover, driving styles might influence the risk of standing passengers losing balance. In this context, shortening the time required for standing passengers to be seated can be beneficial for reducing the risk of falling after boarding the bus, particularly for vulnerable road users.

This study analyzed the factors affecting the occurrence of passenger falls on buses by targeting the driver's driving behaviors, including acceleration, deceleration braking, and steering maneuvers. Negative binomial and Poisson lognormal models were developed using passenger fall incident data. The driving behaviors of drivers were calculated using data from digital tachographs (DTGs) attached to buses. Moreover, using the developed model, the major factors affecting the occurrence of passenger falls on buses with respect to the driver's driving behavior were determined, including acceleration, deceleration, braking, and steering maneuvers.

1.2. Research Scope and Procedure. This study involved city buses operating in Korea from 2016 to 2018. The passenger fall incident and dangerous driving behavior data calculated from the DTGs mounted on the buses were used for the analysis. The dangerous driving behavior of bus drivers was classified into eight categories by processing the DTG data (recorded in seconds). Detailed information, such as the calculation criteria for each dangerous driving behavior, is provided below. Using the collected data, Poisson lognormal and negative binomial models were developed using the Bayesian method. The effects of drivers' driving behaviors on the occurrence of passenger fall incidents were analyzed using the models.

2. Literature Review

2.1. Prior Studies. Existing research conducted on bus accidents primarily focused on predicting the number of traffic accidents, analyzing the factors affecting traffic accidents, improving safety, and measuring the impacts of safety measures. Many studies were conducted to predict and evaluate the effects of various factors on the number of traffic accidents [19–28]. However, owing to difficulties in data collection in most studies, human factors and significant factors, such as driver behavior and driving style, are rarely considered.

This study investigated research related to the effects of human factors on drivers of commercial vehicle traffic accidents. Related studies can be classified into two categories. The first category comprises studies analyzing the relationships between traffic accidents and drivers' driving behaviors using information visualized with data, such as that from DTGs. The second category includes studies analyzing the relationships between traffic accidents and factors that can be investigated through surveys, such as the drivers' age, driving history, and the number of law violations.

Blower et al. [7] analyzed the effects of bus drivers' driving behaviors on safety. A logistic regression model was developed based on driving records, law violation records, accident records for drivers of buses involved in fatal accidents, and bus operation types for three years, surveyed through a questionnaire. The analysis revealed that only the bus operation type had a statistically significant influence on the occurrence of bus fatalities.

Lee et al. [14] analyzed the factors affecting the occurrence of accidents according to the types of business vehicles, such as buses, trucks, and taxis. Their study compared drivers driving different types of vehicles in similar regions and environments. A mixed logit model was developed using personal information data from 627,594 commercial vehicle drivers in South Korea. The drivers' personal information included the number of traffic accidents, gender, driving years, job turnover, and the number of law violations. The analysis revealed that extensive driving experience reduced the occurrence of accidents. By contrast, the numbers of law violations, job turnover, and traffic accidents increased the occurrence of traffic accidents.

Zhou and Zhang [18] analyzed the potentially dangerous driving behaviors of drivers. They used DTG data from 4,373 trips over an 11-month period collected from 70 truck drivers. Principal component analysis and density-based spatial clustering were performed to classify driver types based on DTG data. The drivers were classified into five clusters based on the results of the clustering analysis (very safe, slightly safe, slightly dangerous, dangerous, and very dangerous) according to the primary components related to speeding, fatigue, and acceleration (jerk).

Hassan et al. [11] analyzed the effects of human (driver) factors on speed-related crash accidents. A logistic regression model was developed based on self-questionnaire data from 442 drivers. The analysis revealed that the drivers' gender, age, and nationality significantly affected the number of accidents.

Silvano and Ohlin [2] investigated the degree of injury in bus accidents according to the driver's manipulation (acceleration or braking) and passenger conditions (boarding, moving, and alighting). The data covered three and a half years (2015–2018). All passengers stood inside the bus at the time of the accident and were treated in hospitals after the incident. In general, driver manipulation and passenger condition are significant factors influencing fall accidents. In the accelerated maneuvers, older passengers (≥65 years) were most vulnerable to falls immediately after boarding. Falls during braking maneuvers were most common during travel and mostly involved the age group of 25–64 years.

Palacio et al. [29] attempted to derive a strategy to prevent the elderly from falling over a bus. They performed an experiment with a passenger standing on a bus and holding a handle. Acceleration during actual bus driving was measured using a laptop computer equipped with a portable accelerometer and was applied to the simulation. The bus driver was unaware of the data collection. According to the simulation analysis applying the acceleration/deceleration profile of an actual bus, the maximum deceleration was 0.32 g. This exceeded 0.15 g, the threshold for a passenger holding the handle to maintain balance. Thus, the maximum probability of knee and head injuries on passengers who lost balance, colliding with a bus seat, wall, and so on, was 53% and 35%, respectively.

Zhou et al. [30] attempted to identify the factors increasing the severity of passenger injuries in noncollision accidents occurring on public buses. A dataset of 17,383 passengers injured while riding a bus in Hong Kong within 10 years (2007–2017) was analyzed. Based on the collected dataset, they built a random parameter logistic regression model to estimate the likelihood of fatal and severe injuries due to various factors. The analysis revealed that bus speed significantly affected the severity of passenger accidents. Their study classified the drivers' behaviors into comprehensive characteristics, such as "driving too fast," "losing control of the vehicle," and "failing to ensure passenger safety." "Driving too fast" significantly affected the occurrence of fatal and severe injuries.

Barabino et al. [5] proposed a framework for evaluating the severity of bus routes by considering both accident frequency and severity in relation to public bus accidents. This framework includes risk factors in terms of frequency, severity, and exposure to bus accidents among various safety-related factors (e.g., road topography, pavement, vehicle mileage, driving propensity, and accident history). In addition, they developed a model that includes these factors. The risk of accidents on bus routes was calculated using this model, and a safety performance ranking was provided.

These results indicate that acceleration and braking maneuvers may need to be studied separately. The driving style can affect a standing passenger's risk of losing balance. Several studies also analyzed fall incidents on buses [31, 32].

Although prior studies dealt with collision and noncollision accidents occurring on public buses, few studies focused on the effect of bus movement on passenger comfort.

Eboli et al. [33] mentioned the importance of acceleration as a kinematic parameter significantly affecting the comfort of bus passengers. This study analyzed the degree of comfort passengers feel in the bus when the bus travels along the route based on acceleration. For the analysis, the level of comfort was investigated with a questionnaire for passengers on the bus, and acceleration data were obtained using a global positioning system (GPS), and a 3-axis accelerometer mounted on a smartphone.

Barabino et al. [34] cited on-board comfort as a key factor influencing the quality of public bus services and suggested a framework that can evaluate it in real-time. Two types of data were collected in this framework. The first was the question-and-answer data that the passengers subjectively evaluated according to the 10 grades of the bus driver's driving style. The second is longitudinal, lateral, and vertical acceleration data collected using a GPS device mounted on a smartphone and a 3-axis accelerometer. The authors developed a regression model using these data. The independent variable of the regression model is the root mean square weighted acceleration, quantifying whole-body vibrations inducing comfort and motion sickness. The dependent variable is the driving style evaluated on a scale of 1 to 10.

Bridgelall [35] proposed a composite roughness index (CRI) that can evaluate the ride quality using the longitudinal, lateral, and vertical forces generated when a moving object travels along a path and the rotational speed of each axis. The CRI was applied to five routes for comparative analysis. Data on the six forces were collected using accelerometers and gyroscopes mounted on a smartphone.

Most previous studies that analyzed bus movement in relation to passenger comfort used longitudinal, lateral, and vertical acceleration information collected through an accelerometer. This study used the data recorded on the DTG installed in the vehicle, whereas previous studies used a sensor mounted on a machine, such as a smartphone or a laptop. However, similar to previous studies, this study used data on the vehicle's position, speed, and acceleration. Previous studies used data collected based on the vehicle's perspective, whereas this study used data collected based on the driver's perspective. In addition, previous studies evaluated comfort, rather than accidents, by integrating this acceleration information. Moreover, movement by detailed direction, such as sudden turns and stops, was not analyzed sufficiently.

Most previous studies used data on a driver's personal information (age, gender, driving experience, and so on) in considering human factors. By contrast, studies based on data directly reflecting drivers' driving behavior are limited. In addition, the data used in this study were of noncollision accidents, in which passengers in the vehicle were overturned owing to the bus movement. Therefore, it is inappropriate to directly compare the results with those of studies conducted on collision accidents. This study compared the drivers' behavior and characteristics with the vehicle's movement, such as speeding, sudden acceleration, and abrupt deceleration and turns, based on the DTG data. The data are expressed in terms of the bus movement and used to predict the occurrence of incidents.

3. Data Description

3.1. Public Bus Accidents in Korea. Most companies operating public buses in Korea are affiliated with the National Bus Mutual-Aid Association to manage liability for damage caused by accidents. Therefore, the data of the National Bus Mutual-Aid Association includes detailed information on most bus traffic accidents, including those that were not reported to the police.

3.2. Digital Tachograph (DTG)-Based Dangerous Driving Behaviors. In Korea, the installation of DTGs has been mandatory for commercial vehicles, such as buses, trucks, and taxis, since 2011. DTG-based dangerous driving behavior data are generated based on vehicle speed, acceleration, revolutions per minute, brake signals, and GPS location information recorded in the DTG. The Korea Transportation Safety Authority has defined eight dangerous driving behaviors based on the collected DTG data, as shown in Table 1. Dangerous driving behaviors have different standards according to the type of vehicle (trucks, buses, and taxis).

For reference, the DTG data installed in commercial vehicles, such as buses, trucks, and taxis, are periodically submitted to Empowering Tomorrow's Automotive Software [ETAS (https://etas.kotsa.or.kr/index.jsp)], the Digital Tachograph Analysis System of the Korea Transportation Safety Authority. The submitted DTG data were preprocessed, processed, stored, and utilized for the management of commercial vehicles.

In the case of DTG data, erroneous information may be recorded owing to various factors, such as poor radio reception and device errors; therefore, it is necessary to filter out data that are determined to be outliers or errors. The Korea Transportation Safety Authority removes these errors through preprocessing.

3.3. Data Summary. The DTG-based dangerous driving behavior data for three years can be aggregated into vehicle units. However, fall incident data on buses are aggregated into bus operating company units. Therefore, the DTG-based dangerous driving behavior data were aggregated according to the operating company. This study analyzed 414 bus operating companies. According to the data on dangerous driving behaviors of the companies studied, sudden acceleration and abrupt deceleration occurred at a higher frequency on average than the other dangerous behaviors, as shown in Table 2.

4. Methodology

4.1. Introduction. A method was established to analyze the dangerous driving behaviors of bus drivers influencing passenger fall incidents on buses. The analysis methodology included the following steps: first, the variables, models for the analysis, and model estimation method were selected. The variable selection involved selecting an appropriate variable for constructing a regression model while considering the multicollinearity problem among the collected data. Second, models suitable for the analysis were selected by considering the characteristics of the passenger fall incident data. Third, the major behaviors affecting in-vehicle passenger fall incidents were derived from the eight dangerous driving behaviors of drivers using the developed model. Finally, the implications of the two models are analyzed.

4.2. Variable and Model Selection. The variables for building the model must be selected before developing regression models using DTG-based dangerous driving behavior data. The significance of the corresponding variable cannot be confirmed in the context of multicollinearity, in which a highly linear relationship exists between the independent variables. In the Bayesian method, this multicollinearity problem can be solved using an informative prior distribution [36]. However, this study assessed the existence of multicollinearity through the variance inflation factor (VIF) because there was no sufficient information on the prior distribution; variables considered to have multicollinearity were removed.

Subsequently, an appropriate model was selected to analyze whether the driver's driving behavior, expressed using DTG data, was related to the number of in-vehicle passenger fall incidents. A traffic accident is an independent random event. Therefore, a Poisson distribution is a suitable statistical analysis tool for explaining the randomness of traffic accidents. However, it is inappropriate to apply a Poisson distribution in actual traffic accidents. This is because a Poisson distribution assumes equal mean and variance; the variance in actual accidents is greater than the mean. A negative binomial regression model was used in various studies to overcome this overdispersion problem, including those forming the basis of the U.S. Highway Safety Manual [20, 37–41]. Therefore, this study used a negative binomial regression model for the analysis and the Bayesian method for estimating the model parameters. In addition, a Poisson lognormal model was constructed and compared with a negative binomial regression model [19, 42–46].

4.3. Estimation of Poisson Lognormal Regression Using Bayesian Method. As previously discussed, data that occur randomly and exhibit overdispersion, such as actual traffic accidents, can be expressed using a mixed distribution, with the number of traffic accidents as a dependent variable following a Poisson distribution and showing a varying mean [47]. Thus, negative binomial and Poisson lognormal regression models can be used because these models use a mixed distribution.

Maximum likelihood estimation (MLE) was used to estimate the parameters of the negative binomial regression model [48, 49]. However, the likelihood function for the dependent variable does not appear in a closed form in the Poisson lognormal regression model; thus, it is difficult to apply the MLE method [50, 51]. In addition, studies have used the Bayesian and Markov chain Monte Carlo (MCMC) methods to numerically obtain the expected value of the posterior distribution [19, 42–46, 52–55]. Moreover, when the sample size is not large and has a highly skewed distribution, such as with traffic accident data, the Bayesian method can provide a more reliable analysis than the MLE method [56–58]. Based on these characteristics, the parameters of the Poisson lognormal and negative binomial models were estimated using the Bayesian method. The theoretical framework of this method is expressed by the following equation:

Table 1: Dangerous driving behavior calculation criteria.

Types	Calculation criteria
Speeding	Exceeding the road speed limit by 20 km/h
Sudden acceleration	Accelerating more than 6.0 km/h per second at speeds of 6.0 km/h or more
Sudden start	Accelerating more than 8.0 km/h per second at speeds below 5.0 km/h
Abrupt deceleration	Decelerating more than 9.0 km/h per second at speeds of 6.0 km/h or more
Abrupt stop	When the speed is reduced to 5.0 km/h or less by decelerating more than 9.0 km/h per second
Abrupt turn left or turn right	Sharp turn to left/right (accumulated rotation angle is in the range of 60–120°) within 4 s at a speed of 25 km/h or more
Abrupt U-turn	Drive left/right (range 160–180°) in 8 s with a speed of more than 20 km/h

Table 2: Statistics of the DTG data.

Types	Average	Standard deviation	Min. value	Max. value
No. of fall accidents per bus	0.7	0.5	0.0	3.7
Speeding per 100 km·bus	7.5	6.0	0.2	32.2
Sudden acceleration per 100 km·bus	24.9	21.9	0.1	88.6
Sudden start per 100 km·bus	7.5	8.0	0.0	33.9
Abrupt deceleration per 100 km·bus	16.5	12.5	1.3	74.1
Abrupt stop per 100 km·bus	4.9	5.3	0.0	35.6
Abrupt turn left per 100 km·bus	5.5	6.5	0.2	63.2
Abrupt turn right per 100 km·bus	2.7	2.1	0.2	19.3
Abrupt U-turn per 100 km·bus	1.7	3.1	0.0	23.4

$$\pi(\theta|y) = \frac{L(y|\theta)\pi(\theta)}{\int L(y|\theta)\pi(\theta)\mathrm{d}\theta}, \quad (1)$$

where y denotes the observed data, θ is a vector of the numbers constituting the likelihood function, $L(y|\theta)$ is the likelihood function, $\pi(\theta)$ is the prior distribution of θ, $\int L(y|\theta)\pi(\theta)\mathrm{d}\theta$ is the marginal distribution of the observed data, and $\pi(\theta|y)$ is the posterior distribution.

In equation (1), the posterior distribution is obtained by combining the observed data with the prior distribution information. The Poisson lognormal model is defined in the following equation:

$$Y_i|\theta_i \sim \text{Poisson}(\theta_i), \quad (2)$$

where Y_i denotes the passenger fall accidents of bus operating company i.

In the Poisson lognormal model, assumptions, such as those in equations (3) and (4), are assumed to address overdispersion resulting from unobserved or unmeasured heterogeneity.

$$\theta_i = \mu_i \exp(\varepsilon_i), \quad (3)$$

$$\log(\mu_i) = \beta_0 + \beta X + \log(\text{offset}), \quad (4)$$

where X denotes the vector of dangerous driving behavior (independent variable) and β is the vector of regression coefficients for each independent variable.

The error term representing the random effect can be expressed as follows: the variance parameter is specified to follow the gamma prior distribution and is equal to the following equation:

$$\exp(\varepsilon_i) \sim \text{lognormal}\left(0, \sigma_\varepsilon^2\right). \quad (5)$$

The initial values of the parameters can be made robust using appropriate informative priors, such as information and experience from related research (Wang et al., 2015). However, there was no sufficient informative prior to this study; therefore, a noninformative distribution was used along with a normal distribution with a mean and variance of 0 and 100,000, respectively.

A statistical analysis programming language was used to build the model using the Bayesian method. The Bayesian method uses a sampling algorithm such as the MCMC to estimate the posterior distribution of each parameter. In this study, the no-U-turn sampler (NUTS) algorithm provided by RStan was used. The NUTS was developed based on the MCMC sampling algorithm. It converges faster than existing sampling methods (Salvatier et al., 2016). In addition, the Gelman—Rubin statistic was used to determine whether the parameters estimated through sampling had converged. The Gelman—Rubin statistic compares the variance of the entire chain sampled for each parameter with that of each chain. When this value was less than 1.1, the corresponding parameter was considered to have converged [59].

5. Analysis and Result

5.1. Multicollinearity Analysis for Variable Selection. Appropriate variables for building the regression model were selected according to the methodology described in Section 4. Multicollinearity may exist for a variable with a VIF value equal to or greater than 5; when it is greater than or equal to 10, there is significant multicollinearity [60]. In this study, considering that the data to be analyzed were the driving behaviors of the drivers, the correlation between each independent variable could be higher than that with other data. Therefore, the more conservative VIF value of 5

Table 3: VIF calculation results after excluding abrupt U-turns and sudden acceleration variables.

Types	Speeding	Sudden acceleration	Sudden start	Abrupt deceleration	Abrupt stop	Abrupt turn left	Abrupt turn right	Abrupt U-turn
VIF	1.14	—	1.59	1.60	1.11	2.16	2.21	—

Table 4: Bayesian negative binomial regression modeling result.

Variable	Mean	Standard deviation	2.5%	97.5%	R-hat
Constant	−0.919	0.072	−1.059	−0.781	1.000
Speeding (x_1)	−0.002	0.005	−0.012	0.008	1.000
Sudden start (x_3)	−0.005	0.004	−0.013	0.004	1.000
Abrupt deceleration (x_4)	0.031	0.003	0.025	0.037	1.000
Abrupt stop (x_5)	0.020	0.006	0.009	0.032	1.000
Abrupt turn left (x_6)	0.007	0.007	−0.005	0.021	1.000
Abrupt turn right (x_7)	−0.034	0.020	−0.074	0.007	1.000
Dispersion parameter (α)	0.283	0.273	0.244	0.331	1.000

Table 5: Bayesian Poisson lognormal regression modeling result.

Variable	Mean	Standard deviation	2.5%	97.5%	R-hat
Constant	−1.054	0.071	−1.194	−0.918	1.000
Speeding (x_1)	−0.003	0.005	−0.013	0.007	1.000
Sudden start (x_3)	−0.001	0.004	−0.010	0.007	1.000
Abrupt deceleration (x_4)	0.029	0.004	0.024	0.035	1.000
Abrupt stop (x_5)	0.019	0.006	0.008	0.031	1.000
Abrupt turn left (x_6)	0.009	0.007	−0.004	0.022	1.010
Abrupt turn right (x_7)	−0.035	0.020	−0.074	0.006	1.000
Dispersion parameter (σ_e)	0.739	0.015	0.710	0.771	1.000

was used as the criterion to determine multicollinearity (James et al., 2013).

The VIF values for the abrupt left turn and U-turn were identified as 18.63 and 15.86, respectively, suggesting the existence of multicollinearity. Considering that the road sections requiring U-turns on bus routes were typically fewer than those requiring a left turn, the VIF was recalculated while excluding the abrupt U-turn. After removing the abrupt U-turn variable, the VIF of the abrupt left turn variable decreased to 2.18; the multicollinearity between the abrupt left turn and U-turn disappeared. However, the VIF for sudden acceleration and start remained higher than 5.

In the sudden start variable, according to the calculation criteria for dangerous driving behaviors in Table 1, the operation of acceleration from a state close to a stop was counted based on the DTG raw data. These conditions were determined to significantly affect passenger fall accidents more than sudden acceleration while already accelerating during driving. Therefore, VIF was calculated after removing the sudden acceleration variable, as shown in Table 3. The calculated VIF values for all variables were lower than 5, indicating no multicollinearity problem. Six variables were selected to build the model: speeding, sudden start, abrupt deceleration, abrupt stop, and abrupt left and right turns.

5.2. Modeling Result. Negative binomial regression and Poisson lognormal models were constructed using the six selected variables. As aforementioned, the parameters of the two models were estimated using the Bayesian method. Tables 4 and 5 show the estimation results from the negative binomial regression and Poisson lognormal regression models, respectively. Two chains were constructed to sample the posterior distribution. In total, 16,000 posterior samples were obtained by performing 10,000 samplings for each chain, including 2,000 warm-up times, which were subsequently excluded. In the negative binomial regression model, the 95% confidence intervals of the regression coefficients for the speeding, sudden start, and abrupt left and right turn variables (i.e., those other than abrupt deceleration and abrupt stop) contained zeros. Thus, the speeding, sudden start, and abrupt left and right turn variables did not have statistically significant effects on passenger fall accidents on a bus. The Poisson lognormal regression model also showed that only two variables (abrupt deceleration and abrupt stop) significantly influenced passenger fall accidents, similar to the results of the negative binomial regression model. In addition, the values of the Gelman–Rubin (R-hat) statistic for both model parameters were less than 1.1; thus, the parameters converged.

Table 6: Bayesian regression model comparison: negative binomial vs. Poisson lognormal.

	Negative binomial	Poisson lognormal
WAIC	3744.1	2974.9

The negative binomial and Poisson lognormal regression models estimated using Bayesian methods showed similar results. However, the widely applicable information criterion (WAIC) of the Poisson lognormal regression model was smaller, as shown in Table 6. Thus, the fit of the Poisson lognormal regression model was better than that of the negative binomial model. Deceleration-related driving behaviors, such as abrupt deceleration and abrupt stop, were found to be the major factors influencing the occurrence of passenger fall incidents on buses. In addition, acceleration-related driving behaviors, such as speeding and sudden starts, or rotation-related driving behaviors, such as abrupt left and right turns, did not have statistically significant effects. This is because the passengers had time to anticipate and respond to these maneuvers. By contrast, abrupt deceleration and abrupt stop are primarily performed by drivers in response to an unexpected event, such as a sudden stop of the vehicle in front of it or the appearance of an obstacle.

6. Conclusion and Future Research

6.1. Conclusion. This study analyzed the drivers' behaviors affecting passenger fall incidents on city buses with many standing passengers. DTG-based dangerous driving behavior data reflecting drivers' driving styles, known to significantly influence the occurrence of traffic accidents, were used. Two regression models were developed using DTG-based dangerous driving behavior and passenger fall incident data, with the number of passenger fall incidents and the number of dangerous driving behaviors as the dependent and independent variables, respectively. The negative binomial and Poisson lognormal regression models were built using the Bayesian method.

Both models yielded similar results. In all cases, the deceleration-related variables, such as abrupt deceleration and abrupt stop, significantly influenced the number of passenger fall incidents on buses. Moreover, acceleration-related variables, such as speeding and sudden start, and rotation-related variables, such as abrupt left and right turns, did not have statistically significant effects.

The results of this study showed that acceleration significantly affected the number of passenger fall incidents, similar to the results of previous studies [2, 29, 30]. Abrupt deceleration and abrupt stopping, with large changes in acceleration, significantly influenced the occurrence of passenger fall accidents.

This study quantified bus movement according to the driver's actual driving behavior using DTG data. In addition, this study analyzed the movements affecting the number of passenger fall incidents among the quantified bus movements. The model constructed in this study showed that deceleration-related driving behaviors significantly affect the occurrence of accidents; however, the actual accidents may be because of unobserved factors not reflected in this study.

In this study, it was not possible to directly match the bus driver and their DTG data according to the national policy to protect the driver's personal information. Therefore, aggregated data for each transportation company was used for model development. However, for more precise analysis of the cause of the actual passenger fall incident, time series data of individual buses before and after the accident is required. The time series data includes serial changes in road geometry, passenger boarding positions, number of passengers, and acceleration/deceleration of bus. Through these data, it is expected that factors affecting actual accidents that were not observed in this study could be identified. Therefore, in future research, analysis using time series data of individual buses is considered necessary.

The safety facilities for bus passengers, driver qualifications, vehicle safety standards, and laws and systems differ between countries. In addition, these data may be difficult to use if there are no legal regulations regarding the installation of DTG-based devices. However, the methodology presented in this study can be used to develop standards for evaluating drivers' safe-driving abilities from the passengers' perspective. In addition, when a public institution that manages public transportation safety wants to manage the safety level of transportation companies based on the frequency of dangerous driving behavior and the frequency of actual passenger fall incidents in the bus, the institution can utilize the results derived from this study. However, among the factors that actually cause the passenger fall incident, there may be factors that were not considered in this study. Therefore, it is judged that the direction of safety management proposed in this study is appropriate to give incentives to safe companies and drivers, rather than to penalize dangerous companies and drivers. These standards can also be applied to autonomously driving buses and shuttles currently under development.

6.2. Future Research. This study analyzed the driving behaviors affecting passenger fall accidents on buses. Based on the results of this study, future research is required to improve the safety of passengers on buses. The potential directions for future research are summarized below.

First, in this study, several drivers drove the same bus alternately. In addition, one-to-one matching between the vehicle and driver was not possible owing to legal restrictions. A more precise analysis is possible if the driving behaviors of individual drivers and accidents can be combined using additional data.

Second, this study used dangerous driving behavior data by processing the DTG data. These data quantify the driver's manipulation into a single behavior, such as sudden or abrupt deceleration. However, in an actual traffic accident, such sudden driving maneuvers may occur in a continuous and complex situation. In future research, it is necessary to analyze repetitive or complex dangerous driving behaviors highly related to traffic accidents.

Third, specific circumstances were not considered in this study, such as the locations where dangerous driving behavior occurred and the number of passengers on board. In a location where dangerous driving behavior occurs, it is possible to further consider geometrical factors, such as the slope and curvature of the relevant road. In addition, if dangerous driving behaviors occur multiple times without passengers on board, it would be inappropriate to analyze the relationships between the number of passengers and these accidents. The above data can be supplemented with additional data, such as those from closed-circuit televisions in vehicles and transportation cards.

Fourth, the model constructed in this study showed that deceleration-related driving behaviors had a significant effect on accident occurrence, but the actual accident may be due to unobserved factors that were not reflected in this study. Therefore, in future research, it is necessary to examine the relationship between vehicle dynamic movement and passenger fall incidents using more precise data collected from autonomous buses.

Data Availability

The data used in this study cannot be made available due to the Korea Transportation Safety Authority (TS)' policy.

Conflicts of Interest

The authors declare that they have no conflicts of interest or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by Korea Institute of Police Technology (KIPoT) grant funded by the Korea Government (KNPA) (no. 092021C29S01000, Development of Traffic Congestion Management System for Urban Network).

References

[1] T. Brenac and N. Clabaux, "The indirect involvement of buses in traffic accident processes," *Safety Science*, vol. 43, no. 10, pp. 835–843, 2005.

[2] A. P. Silvano and M. Ohlin, "Non-collision incidents on buses due to acceleration and braking manoeuvres leading to falling events among standing passengers," *Journal of Transport & Health*, vol. 14, Article ID 100560, 2019.

[3] S. Cafiso, A. Di Graziano, and G. Pappalardo, "Road safety issues for bus transport management," *Accident Analysis & Prevention*, vol. 60, pp. 324–333, 2013.

[4] European Road Safety Observatory, "Traffic safety basic facts 2018 – heavy goods vehicles and buses," 2018, https://road-safety.transport.ec.europa.eu/system/files/2021-07/bfs2018_hgvs.pdf.

[5] B. Barabino, M. Bonera, G. Maternini, A. Olivo, and F. Porcu, "Bus crash risk evaluation: an adjusted framework and its application in a real network," *Accident Analysis & Prevention*, vol. 159, pp. 106258–106316, 2021.

[6] U. Barua and R. Tay, "Severity of urban transit bus crashes in Bangladesh," *Journal of Advanced Transportation*, vol. 44, no. 1, pp. 34–41, 2010.

[7] D. Blower, P. E. Green, and A. Matteson, *Bus Operator Types and Driver Factors in Fatal Bus Crashes: Results from the Buses Involved in Fatal Accidents Survey*, Transportation Research Institute, Ann Arbor, MI, USA, 2008.

[8] C. Cheung, A. S. Shalaby, B. N. Persaud, and A. Hadayeghi, "Models for safety analysis of road surface transit," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2063, no. 1, pp. 168–175, 2008.

[9] D. Chimba, T. Sando, and V. Kwigizile, "Effect of bus size and operation to crash occurrences," *Accident Analysis & Prevention*, vol. 42, no. 6, pp. 2063–2067, 2010.

[10] K. C. K. Goh, G. Currie, M. Sarvi, and D. Logan, "Bus accident analysis of routes with/without bus priority," *Accident Analysis & Prevention*, vol. 65, pp. 18–27, 2014.

[11] H. M. Hassan, M. Shawky, M. Kishta, A. M. Garib, and H. A. Al-Harthei, "Investigation of drivers' behavior towards speeds using crash data and self-reported questionnaire," *Accident Analysis & Prevention*, vol. 98, pp. 348–358, 2017.

[12] J. Huting, J. Reid, U. Nwoke, E. Bacarella, and K. E. Ky, "Identifying factors that increase bus accident risk by using random forests and trip-level data," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2539, no. 1, pp. 149–158, 2016.

[13] S. Kaplan and C. G. Prato, "Risk factors associated with bus accident severity in the United States: a generalized ordered logit model," *Journal of Safety Research*, vol. 43, no. 3, pp. 171–180, 2012.

[14] J. Lee, J. Yeo, I. Yun, and S. Kang, "Factors affecting crash involvement of commercial vehicle drivers: evaluation of commercial vehicle drivers' characteristics in South Korea," *Journal of Advanced Transportation*, vol. 2020, Article ID 5868379, 8 pages, 2020.

[15] C. G. Prato and S. Kaplan, "Bus accident severity and passenger injury: evidence from Denmark," *European Transport Research Review*, vol. 6, no. 1, pp. 17–30, 2014.

[16] E. F. Sam, S. Daniels, K. Brijs, T. Brijs, and G. Wets, "Modelling public bus/minibus transport accident severity in Ghana," *Accident Analysis & Prevention*, vol. 119, pp. 114–121, 2018.

[17] O. Uçar and H. Tatlıdil, "Factors influencing the severity of damage in bus crashes in Turkey during 2002: a application of the ordered probit model," *Hacettepe J. Math. Statistics*, vol. 36, no. 1, 2007.

[18] T. Zhou and J. Zhang, "Analysis of commercial truck driver's potentially dangerous driving behaviors based on 11-month digital tachograph data and multilevel modeling approach," *Accident Analysis & Prevention*, vol. 132, Article ID 105256, 2019.

[19] M. A. Abdel-Aty and A. E. Radwan, "Modeling traffic accident occurrence and involvement," *Accident Analysis & Prevention*, vol. 32, no. 5, pp. 633–642, 2000.

[20] A. Das and M. A. Abdel-Aty, "A combined frequency–severity approach for the analysis of rear-end crashes on urban arterials," *Safety Science*, vol. 49, no. 8-9, pp. 1156–1163, 2011.

[21] V. V. Dixit, A. Pande, M. Abdel-Aty, A. Das, and E. Radwan, "Quality of traffic flow on urban arterial streets and its relationship with safety," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1610–1616, 2011.

[22] A. Farid, M. Abdel-Aty, J. Lee, N. Eluru, and J.-H. Wang, "Exploring the transferability of safety performance functions," *Accident Analysis & Prevention*, vol. 94, pp. 143–152, 2016.

[23] E. Hauer, "On the estimation of the expected number of accidents," *Accident Analysis & Prevention*, vol. 18, no. 1, pp. 1–12, 1986.

[24] E. Hauer, *Observational Before-After Studies in Road Safety*, Pergamon, Oxford, UK, 1997.

[25] E. Hauer, "Overdispersion in modelling accidents on road sections and in Empirical Bayes estimation," *Accident Analysis & Prevention*, vol. 33, no. 6, pp. 799–808, 2001.

[26] K. A. Kaaf and M. Abdel-Aty, "Transferability and calibration of Highway Safety Manual performance functions and development of new models for urban four-lane divided roads in Riyadh, Saudi Arabia," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2515, no. 1, pp. 70–77, 2015.

[27] A. Khattak, L. Zhang, J. L. Hochstein, and S. C. Tee, "Crash analysis of expressway intersections in Nebraska," in *Proceedings of the Transportation Research Board's 85th Annual Meeting*, Washington, DC, USA, 2006.

[28] J. Ma, K. M. Kockelman, and P. Damien, "A multi variate Poisson-lognormal regression model for prediction of crash counts by severity, using Bayesian methods," *Accident Analysis & Prevention*, vol. 40, no. 3, pp. 964–975, 2008.

[29] A. Palacio, G. Tamburro, D. O'Neill, and C. K. Simms, "Non-collision injuries in urban buses-Strategies for prevention," *Accident Analysis & Prevention*, vol. 41, no. 1, pp. 1–9, 2009.

[30] H. Zhou, C. Yuan, N. Dong, S. C. Wong, and P. Xu, "Severity of passenger injuries on public buses: a comparative analysis of collision injuries and non-collision injuries," *Journal of Safety Research*, vol. 74, pp. 55–69, 2020.

[31] S. Krašna, A. Keller, A. Linder et al., "Human response to longitudinal perturbations of standing passengers on public transport during regular operation," *Frontiers in Bioengineering and Biotechnology*, vol. 9, Article ID 680883, 2021.

[32] J. C. Xu, A. P. Silvano, A. Keller et al., "Identifying and characterizing types of balance recovery strategies among females and males to prevent injuries in free-standing public transport passengers," *Frontiers in Bioengineering and Biotechnology*, vol. 9, Article ID 670498, 2021.

[33] L. Eboli, G. Mazzulla, and G. Pungillo, "Measuring bus comfort levels by using acceleration instantaneous values," *Transportation Research Procedia*, vol. 18, pp. 27–34, 2016.

[34] B. Barabino, M. Coni, A. Olivo, G. Pungillo, and N. Rassu, "Standing passenger comfort: a new scale for evaluating the real-time driving style of bus transit services," *IEEE Transactions on Intelligent Transportation Systems*, vol. 20, no. 12, pp. 4665–4678, 2019.

[35] R. Bridgelall, "Characterizing ride quality with a composite roughness index," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 9, pp. 15288–15297, 2022.

[36] M. H. Pesaran and R. P. Smith, "A Bayesian analysis of linear regression models with highly collinear regressors," *Econometrics and Statistics*, vol. 11, pp. 1–21, 2019.

[37] S. Daniels, T. Brijs, E. Nuys, and G. Wets, "Externality of risk and crash severity at roundabouts," *Accident Analysis and Prevention*, vol. 42, pp. 1966–1973, 2010.

[38] S. Cafiso, A. Di Graziano, G. Di silvestro, G. La Cava, and B. Persaud, "Development of comprehensive accident models for two-lane rural highways using exposure, geometry, consistency and context variables," *Accident Analysis & Prevention*, vol. 42, pp. 1072–1079, 2010.

[39] C. Siddiqui, M. Abdel-Aty, and H. Huang, "Aggregate nonparametric safety analysis of traffic zones," *Accident Analysis & Prevention*, vol. 45, pp. 317–325, 2012.

[40] M. Abdel-Aty, J. Lee, C. Siddiqui, and K. Choi, "Geographical unit based analysis in the context of transportation safety planning," *Transportation Research Part A*, vol. 49, pp. 62–75, 2013.

[41] J. Lee, M. Abdel-Aty, K. Choi, and C. Siddiqui, "Analysis of residence characteristics of drivers, pedestrians and bicyclists involed in traffic crashes," *Transportation Research Board 92nd Annual Meeting, Transportation Research Board of the National Academics, Washington, D.C., USA*, 2013.

[42] M. M. Haque, H. C. Chin, and H. Huang, "Applying Bayesian hierarchical models to examine motorcycle crashes at signalized intersections," *Accident Analysis & Prevention*, vol. 42, no. 1, pp. 203–212, 2010.

[43] H. Huang, H. C. Chin, and M. M. Haque, "Empirical evaluation of alternative approaches in identifying crash hotspots: naive ranking, empirical Bayes and full Bayes," *Transportation Research Record*, vol. 2103, no. 1, pp. 32–41, 2009.

[44] R. Kulmala, *Safety at rural three- and four-arm junctions: development and application of accident prediction models*, Technical Research Center at Finland, VTT Publications, Espoo, 1995.

[45] S. P. Miaou, "The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regressions," *Accident Analysis & Prevention*, vol. 26, no. 4, pp. 471–482, 1994.

[46] V. N. Shankar, F. L. Mannering, and W. Barfield, "Effect of roadway geometrics and environmental factors on rural freeway accident frequencies," *Accident Analysis & Prevention*, vol. 27, no. 3, pp. 371–389, 1995.

[47] S. H. Khazraee, V. Johnson, and D. Lord, "Bayesian Poisson hierarchical models for crash data analysis: i," *Accident Analysis & Prevention*, vol. 117, pp. 181–195, 2018.

[48] J. A. Bonneson and P. T. Mccoy, "Effect of median treatment on urban arterial safety an accident prediction model," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1581, no. 1, pp. 27–36, 1997.

[49] E. Hauer, J. C. N. Ng, and J. Lovell, "Estimation of safety at signalized intersections," *Transportation Research Record*, vol. 1185, pp. 48–61, 1988.

[50] C. Gourieroux and A. Monfort, "Simulation based inference in models with heterogeneity," *Annales d'Economie et Statistique*, pp. 69–107, 1990.

[51] J. Hinde, "Compound Poisson regression models, GLIM 82," in *Proceedings of the International Conference on Generalised Linear Models*, vol. 14, pp. 109–121, Berlin, Germany, January 1982.

[52] J. Aguero-Valverde, "Full Bayes Poisson gamma, Poisson lognormal, and zero inflated random effects models: comparing the precision of crash frequency estimates," *Accident Analysis & Prevention*, vol. 50, pp. 289–297, 2013.

[53] S. P. Miaou and D. Lord, "Modeling traffic crash-flow relationships for intersections: dispersion parameter, functional form, and bayes versus empirical bayes methods," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1840, no. 1, pp. 31–40, 2003.

[54] S. P. Miaou and J. J. Song, "Bayesian ranking of sites for engineering safety improvements: decision parameter, treatability concept, statistical criterion and spatial dependence," *Accident Analysis & Prevention*, vol. 37, no. 4, pp. 699–720, 2005.

[55] P. J. Schluter, J. J. Deely, and A. J. Nicholson, "Ranking and selecting motor vehicle accident sites by using a hierarchical Bayesian model," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 46, no. 3, pp. 293–316, 1997.

[56] J. F. Angers and A. Biswas, "A Bayesian analysis of zero-inflated generalized Poisson model," *Computational Statistics & Data Analysis*, vol. 42, no. 1-2, pp. 37–46, 2003.

[57] S. K. Ghosh, P. Mukhopadhyay, and J. C. Lu, "Bayesian analysis of zero-inflated regression models," *Journal of*

Statistical Planning and Inference, vol. 136, no. 4, pp. 1360–1375, 2006.

[58] J. Rodrigues, “Bayesian analysis of zero-inflated distributions,” *Communications in Statistics - Theory and Methods*, vol. 32, no. 2, pp. 281–289, 2003.

[59] A. Gelman, J. B. Carlin, H. S. Stern, and D. B. Rubin, *Bayesian Data Analysis*, Chapman & Hall/CRC, Boca Raton, FL, USA, 2004.

[60] S. Menard, *Applied Logistic Regression Analysis: Sage University Series on Quantitative Applications in the Social Sciences*, Sage, Thousand Oaks, CA, USA, 1995.