Accident Analysis and Prevention 146 (2020) 105736

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

A copula-based approach for jointly modeling crash severity and number of vehicles involved in express bus crashes on expressways considering temporal stability of data

Reuben Tamakloe, Jungyeol Hong, Dongjoo Park*

Department of Transportation Engineering, The University of Seoul, 163 Seoulsiripdae-ro Dongdaemun-gu, Seoul, 02504, South Korea

ARTICLE INFO

Keywords:
Copula
Ordered probit model
Temporal stability
Crash severity
Expressway
Bus-involved crashes

ABSTRACT

The consequences of crashes, including injury, loss of lives, and damage to properties, are further worsened when buses plying expressways are involved in the crash. Previous studies have separately analyzed crash severity in terms of monetary cost, injuries and loss of lives, and the size of crashes in terms of the number of vehicles involved. However, as both outcome variables are correlated, it is imperative to perform a combined analysis using an appropriate econometric model to achieve a better model fit. This study contributes to the literature by jointly exploring the factors influencing the severity and size of express bus-involved crashes that occur on expressways and characterizes the dependence between both outcome variables by employing a more plausible copula regression framework. Likelihood ratio tests were also conducted to investigate the temporal stability of the factors that affect both crash severity and size. Based on the goodness-of-fit statistics, the Frank copula model proved superior to the independent ordered probit model. The estimate of the underlying dependence between the outcome variables provided a better comprehension of the correlation between them. Temporal instability was detected for the individual parameters in the models and is attributed to the changing driving behavior due to the heightened road safety campaigns. The results suggest that traffic exposure measures are significantly associated with a higher propensity of observing increased bus crash severity and size. Insights into the factors influencing the size and severity of express bus crashes are discussed, and appropriate engineering, enforcement, and education-related countermeasures are proposed.

1. Introduction

Traveling by express bus/coach is generally perceived as a safe form of long-distance transportation since the number of express bus-involved crashes is lower than that of other candidate modes (Goh et al., 2014). Relative to other modes of transport, they are also noted to play critical roles in the sustainable growth of economies by reducing congestion on roadways (Kim et al., 2012). Nevertheless, their operational and physical characteristics still make them susceptible to crashes on expressways, and a single bus-involved crash affects many people's lives.

Despite the continuous pragmatic efforts put in place by policymakers and transport engineers to make traveling by express buses much safer, the consequences of crashes involving these high occupancy vehicles remain significant since they convey many people at a given time (Chang and Mannering, 1999). To mitigate the significant externalities of traffic safety (especially in bus transport since buses convey many people) and to suggest adequate safety strategies and countermeasures, it is imperative to carefully examine express bus crash data to gain a comprehensive understanding of the factors that contribute to crash outcomes.

Few studies conducted in the area of bus safety primarily focused on the severity outcomes. They employed variables related to crash characteristics, drivers contributory roles, vehicle attributes, and environmental factors in their investigations (Feng et al., 2016; Park et al., 2019; Peng et al., 2019; Sam et al., 2018). Crash severity is an essential indicator mostly considered when modeling crash data. When crash severity is treated as an ordinal discrete variable comprising levels such as fatal, severe, minor, and property damage only (PDO), an ordered discrete regression model is employed to examine the influence of contributing factors on crash severity (Christoforou et al., 2010). On the other hand, when crash severity is compiled as a non-ordinal variable, multinomial probit, or logit frameworks are used (Bham et al., 2012). It is argued that both ordered and unordered probit or logit models are

* Corresponding author.
E-mail addresses: drtamakloe@uos.ac.kr (R. Tamakloe), jyhong9868@uos.ac.kr (J. Hong), djpark@uos.ac.kr (D. Park).

https://doi.org/10.1016/j.aap.2020.105736
Received 23 May 2020; Received in revised form 25 July 2020; Accepted 10 August 2020
Available online 02 September 2020
0001-4575/

fixed-parameter models and do not allow model parameters to vary across observations. Thus, assuming fixed parameters across all observations, as seen in the current bus crash research, could lead to erroneous results (Hong et al., 2019c; Mannering et al., 2016). To mitigate this, transport safety researchers have used advanced model formulations, incorporating the influence of unobserved heterogeneity to enhance traditional ordered outcome models for improved estimation results (Aidoo and Ackaah, 2019; Yasmin and Eluru, 2013).

Transport safety literature abounds with studies that highlight the variable for the number of vehicles involved in a crash (hereinafter referred to as crash size) as a critical determinant of severity (Broyles et al., 2003; Chang and Chien, 2013; Chang and Mannering, 1999; Feng et al., 2016; Kockelman and Kweon, 2002). Intuitively, crash size has significant ramifications on the injury severity, the level of damage, and the monetary cost of crashes. In actual sense, the number of vehicle occupants exposed to a crash increase as the crash size increases. By treating crash size as an outcome variable, an investigation of crashes on South Korean expressways revealed that crash size is likely to rise when crashes occur on mainlines, and when the Annual Average Daily Traffic (AADT) increases. Driver violations such as improper safe distance and negligence have been blamed for leading to an increase in the crash size of bus-involved crashes (Hong et al., 2019b).

As established, both bus crash severity and crash size are two critical indicators and have been usually modeled independently. Nevertheless, there are commonly observed and unobserved factors such as driver behavior, roadway geometry, or vehicle condition that may affect both outcome variables, highlighting potential interrelationships and correlation among them. In some jurisdictions such as South Korea, crash severity is determined as a function of variables such as the monetary cost of the crash, damage caused by the crash, and the lives lost due to the crash (Hong et al., 2020a, 2019b). Clearly, employing an independent ordered modeling framework to identify the impact of crash-risk factors on crash severity and crash size would result in flawed estimates and conclusions as there exists some dependency/correlation between them, which is unaccounted for. Crash size (number of vehicles involved in the crash) has the tendency of affecting the crash severity (damage), and vice versa. Although correlations due to observed factors can be dealt with by specifying them in some models as independent variables, the unobserved factors cannot be treated similarly since they may not be known. Failure to account for these interrelationships between such indicators that exhibit some level of dependency, and independently modeling them may lead to possible inference issues and erroneous coefficient estimates in prediction models (Martey and Attoh-Okine, 2019, 2018a; Wang et al., 2015). Simulation-based simultaneous equation frameworks (Ye et al., 2008) and structural equation modeling (Lee et al., 2008) may deal with the complex interrelationships between multi-outcome and explanatory variables when they are modeled simultaneously; however, ignoring potential jointness may result in inconsistent parameter estimates (Yasmin et al., 2014). To capture this issue, researchers have employed copula-based methods to obtain reliable model estimates. As shown in Table 1, the use of the copula-based model is increasing among transportation safety researchers who seek to account for the limitations specified above.

A review of the literature shows that the temporal stability of the factors influencing the injury severity and crash size of bus-involved crashes has not been addressed to date. The assumption that the effect of the determinants of crash outcomes do not change over time has been used by most transport safety researchers when modeling crash data. The results of these analyses are used in forecasting and making policies and safety countermeasures to mitigate the increasing trend of road crashes. A state-of-the-art review of the literature spanning fields such as psychology, neuroscience, economics, and cognitive science conducted by Mannering (2018) suggests that the assumption of temporal stability may not always hold. In particular, the study unearthed the potential of temporal instability in models of accident likelihood and the resulting injury severity to exist.

Ignoring to account for the potential temporal shifts in estimated parameters and assuming constant temporal stability in models has adverse effects (Mannering, 2018). It may affect the conclusions drawn from the model estimates and cause the creation of ineffective safety countermeasures. In crash analysis studies, reasons suspected to cause the changes in injury severity of crashes with time have been identified (Behnood and Mannering, 2019). The possibility that the effects of injury severity determinants change over time is due to either the fact that human behavior and driving styles evolve with time due to changes in vehicles, the advancement in technology (Hasler et al., 2014; Leone et al., 2017), or unobserved factors such as lighting and visibility. Although some studies have found a significant reason to support the fact that injury severities of crashes significantly varied from year to year, others did not (Behnood and Mannering, 2019, 2016, 2015).

Road-related incident fatality rates in South Korea had been among the highest compared to the average of countries belonging to the Organization for Economic Co-operation and Development (OECD). To mitigate this increasing trend, the Government of Korea (GoK) implemented several strategies. One of which was the development of the 7th National Transport Safety Master Plan (2012 − 2016) for the road sector. The master plan aimed at reducing the number of traffic fatalities to 3000 by 2016 by setting specific policy-related countermeasures targeted at the various road users. Again, GoK implemented a plan for the promotion of safety at roundabouts from 2013 to 2022. By the end of 2013, a mandatory vehicle registration scheme for all kinds of buses was introduced, and relevant safety standards were heightened. Due to the implementation of stringent safety measures and strategic interventions such as the enforcement of safety management system, enhancement of emergency response system, and improvement of road user behaviors through education, South Korea recorded a 62 % decrease in road-related traffic deaths in 2013 compared to the number of deaths in 1991(13,439). An assessment of efficiency gains in terms of crash reduction in 2014 showed an unprecedented improvement in traffic safety (Adler and Ahrend, 2017; OECD/ITF, 2016; Sul et al., 2014).

Most of the studies on crash analysis have focused on identifying the relationship between some explanatory variables on specific crash outcomes; that is, either crash severity or crash size. However, researchers have paid little attention to modeling both crash outcomes simultaneously to explore the complex interrelationships or dependence between them, which may result from common unobserved crash attributes in the data. Also, to the best of our knowledge, there is currently no study that explicitly analyses express bus-involved crash severity and crash size simultaneously to capture the dependence between them, and to explore the temporal stability of variables that affect the severity of bus crashes.

Against this backdrop, this current study seeks to:

- Investigate the factors affecting crash severity and crash size of express bus-involved crashes that occurred on expressways by employing a bivariate copula regression framework that jointly models both outcome variables and addresses endogeneity due to correlations between both outcome variables
- Explore the dependence between both outcome variables, while accommodating for potential unobserved heterogeneity
- Investigate the temporal stability of variables in the express bus-involved crash data.

The proposed model is estimated using a 5-year (2012 − 2016) crash data from South Korea.

In particular, the research carried out in this current paper contributes to the express bus safety literature, both methodologically and empirically. First, as far as we are concerned, this is the first study that seeks to model both crash size and crash severity jointly using a robust copula approach. The most important advantage of the selected

Table 1
Previous transport-related studies employing the copula-based approach.

Type of study	Reference	Study region/reference period	Objective	Outcome/dependent variables	Copula Dependence parameter
Incident management	Tirtha et al. (2020)	Greater Orlando Region/2013 – 2017	To jointly model factors that affect crash clearance times while accommodating for observed and unobserved temporal effects	Incident type; Incident duration	Negative
	Laman et al. (2018)	Central Florida/2015	To jointly model incident duration components while accommodating for inherent dependencies between them	Reporting time; Response time; Clearance time	Negative
Crash/injury severity	Wang et al. (2019)	Connecticut crash data repository/2016 – 2017	To simultaneously estimate four common intersection crash consequence metrics while accounting for potential correlations due to common unobserved factors.	Driver error; Crash type; Vehicle damage; Crash severity	Negative
	Wali et al. (2018)	Virginia/2013	To simultaneously investigate the injury severity sustained by drivers involved in head-on collisions with respect to fault status designation while accounting for the potential correlation between injury outcomes of drivers.	Crash severity of at-fault driver; Crash severity of not-at-fault driver	Positive
	Ayuso et al. (2016)	Spain/2005 – 2007	To jointly model the association between the set of factors that might influence both categories of injury	Temporal disability; Permanent injuries	Positive
	Wang et al. (2015)	Madison/2005 – 2009	To explore the interrelationships between injury severity and vehicle damage	Injury severity; Vehicle damage	Positive
	Yasmin et al. (2014)	Victoria crash database of Australia/2006 – 2010	To jointly model two dimensions of the injury severity process	Injury severity; Collision type	Positive
	Eluru et al. (2010)	USA/2007	Simultaneously model injury severity based on seat position of vehicle occupant	Driver; Front passenger; Rear passenger	Positive
	Rana et al. (2010)	USA/2007	To jointly model both outcome variables while addressing endogeneity in the models of crash severity	Collision type; Injury severity Injury severity of two drivers involved in two-vehicle crash	Positive
Crash frequency	Bhowmik et al. (2019)	Central Florida/2016	To jointly model crash counts across different crash types	Motorized intersection crashes; Motorized road segment crashes; motorized off-road crashes; non-motorized crashes	Positive
	Yasmin et al. (2018)	Florida Department of Transportation (FDOT) Crash Analysis Reporting System (CARS)/2014	To jointly model crash frequency of different road user groups	Car; Light truck; Van; Other motorized vehicles; non-motorists	Positive
	Nashad et al. (2016)	Florida/2010 – 2012	To jointly model traffic related crashes while accommodating for potential heterogeneity (across zones) in the dependency structure between the outcome variables	Pedestrian crash frequency; Bicyclist crash frequency	Positive
Railway incident	Martey and Attoh-Okine (2018)	USA/2015	To simultaneously model the train derailment severity components while addressing the dependence between them	Monetary damage; Number of derailed cars	Positive

framework is that it accounts for correlations that may exist due to endogeneity, which may occur due to omitted or unobserved variables, measurement errors, or from simultaneous causality (Hong et al., 2019a; Martey and Attoh-Okine, 2018a; Wang et al., 2015). The methodology employed allows for the joint modeling of both crash severity and crash size in such a way that covariates may influence both outcome variables. The results could help identify the impact of crash factors on the outcome variables and guide the formation of strategic countermeasures for improving express bus transport safety. Also, this study is the first attempt to account for temporal stability in express bus crash data. Overlooking this may lead to drawing wrong conclusions (Al-Bdairi et al., 2020). The results from this study would help identify the impact of crash factors on the response variables, and to determine how the policies implemented through the 7th National Transport Safety Master Plan (2012 – 2016) affected the stability of individual parameter estimates on injury severity probabilities across time-period combinations.

2. Methodology

2.1. Overall framework

As shown in Fig. 1, two outcome variables were considered in our empirical analysis. Both crash severity and crash size are usually modeled independently using separate model formulations. Since crash size has a significant impact on crash severity (Feng et al., 2016), this study posits that both variables are interrelated. Accounting for the underlying dependence between such variables by jointly modeling them can result in more appropriate parameter estimates (Martey and Attoh-Okine, 2018a; Wali et al., 2018).

To obtain the objectives of our study, we employed an advanced bivariate copula regression framework together with an independent ordered probit (ORP) model to explore the factors affecting crash severity and crash size of bus-involved crashes that occur on expressways in South Korea using a five-year crash dataset (2012 – 2016). It is worth noting that the unit of analysis with regard to the crash severity used in this study is the entire crash. In South Korea, crash severity is determined based on the cost of damage to property, injury to vehicle occupants, and loss of lives caused by all vehicles involved in a bus crash s Also, as previously discussed, regarding the 2012 – 1016 7th National Transport Safety Master Plan, major safety standards and mandatory bus registration schemes were introduced by the end of 2013. As such, the temporal stability investigation would help us identify whether or not the effect of crash factors on crash outcomes have significantly changed. Besides, it is generally worth exploring the temporal stability of the factors used in a study when the data covers an extended period.

The framework used for the investigation in this study is made up of four main steps. First, we prepared the dataset by cleaning and removing observations with omitted variables. Crashes involving buses on the expressway were extracted for analysis, and the dataset was arranged into three-time periods (A= 2012 – 2013, B= 2014 – 2016, and T= 2012–2016), making sure that each group had an appreciable number of observations to ensure the statistical performance of our models. In the next step, we conducted a preliminary analysis to determine an appropriate copula framework that suitably fits the dataset. Upon selecting the appropriate copula framework, we statistically tested for temporal stability using likelihood ratio tests in the third stage. Finally, for the three data groups, we estimated the dependency between both outcome variables and identified significant contributory factors of crash severity and crash size using the selected bivariate copula regression framework and compared its results with the independent model which assumed no dependence between the two outcome variables. At this stage, we also computed marginal effects to better illustrate the magnitude of the effect of the independent variables on the outcome variables. The results from the study provided valuable insights into the temporal instability of factors affecting the severity and size of bus-involved crashes in South Korea. Based on the useful information obtained from the running the models, we proposed effective road traffic-related policies and strategies that can be adopted by policy-makers and transport planners to reduce the severity and size of crashes. In Fig. 2 below, the framework of this study is presented, and more details about each step mentioned in this paragraph are thoroughly discussed in subsequent sections.

2.2. Copulas

Generally, a copula function is a mathematical construct that links multivariate distribution functions to their univariate marginal distribution functions taking into account the dependency between the outcome variables in the model (Martey and Attoh-Okine, 2018a). It is used to generate stochastic dependence between random variables with predefined marginal distributions (Bhat and Eluru, 2009). A multivariate distribution copula function $C(u_1, u_2, \ldots, u_n)$ defined on the unit hypercube $[0,1]^n$ having n random variables as uniformly distributed marginals represents an n-dimensional copula (Nelson, 2006). Consider a copula function $C(u, v)$ with random variables for crash severity and crash size U and V, respectively. If C: $[0, 1^2] \rightarrow [0,1]$ and satisfies the following conditions listed below, it is regarded as a bivariate copula (Martey and Attoh-Okine, 2018b; Nelson, 2006):

$$C(u, 0) = C(0, v) = 0, \ \forall \ u, v \in [0,1] \tag{1}$$

$$C(u, 1) = u; C(1, v) = v, \forall \ u, v \in [0,1] \tag{2}$$

$$C(u_2, v_2) - C(u_1, v_2) - C(u_2, v_1) + C(u_1, v_1) \geq 0, \ \forall \ 0 \leq u_1 \leq u_2 \leq 1; 0 \leq v_1 \leq v_2 \leq 1 \tag{3}$$

Sklar's theorem is central to the theory of copulas and is the

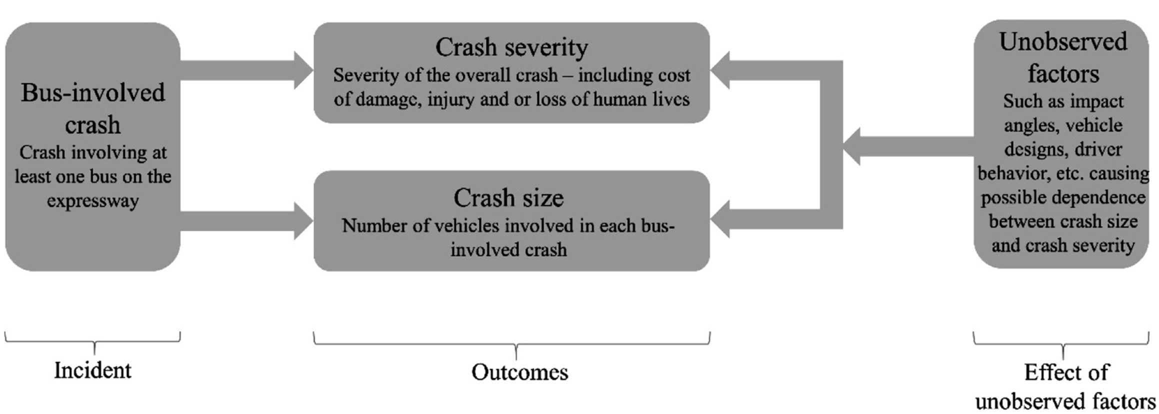


Fig. 1. Outcome variables used as dependent variables in the study.

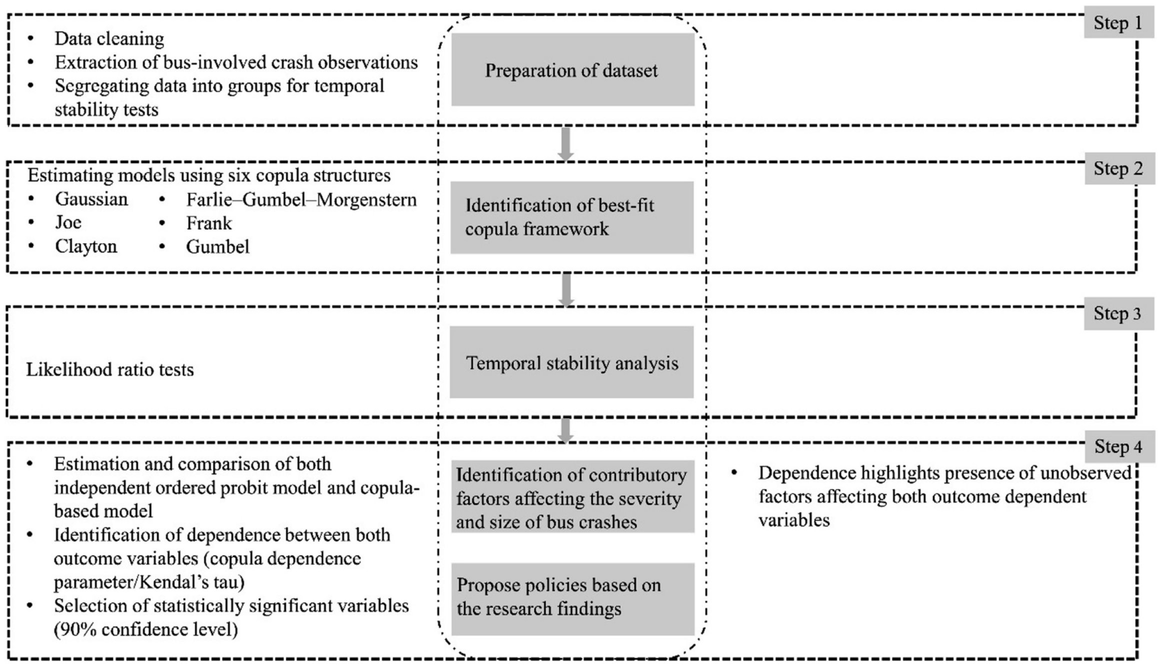


Fig. 2. Proposed study procedure.

foundation of most applications in statistics (Nelson, 2006). According to Sklar's theorem, for any n-dimensional distribution function with marginals $F_1, F_2, \ldots, F_n$, there exists an n-dimensional copula C: $[0,1]^n \to [0,1]$ such that $\forall\ (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$, $F(x_1, x_2, \ldots, x_n) = C\{F_1(x_1), F_2(x_2), \ldots, F_n(x_n)\}$ (Sklar, 1959). If $F_i(x)$ is continuous, then C is unique; otherwise, C is uniquely determined on Range of $F_1 \times$ Range of $F_2 \times \ldots \times$Range of F_n. Given the marginal $F_1, F_2, \ldots, F_n$, and their quasi-inverses, $F_1^{-1}, F_2^{-1}, \ldots, F_n^{-1}$, this theorem presents a valuable way of building copulas, such that;

$$C(x_1, x_2, \ldots x_n) = F(F_1^{-1}(x_1), F_2^{-1}(x_2), \ldots, F_n^{-1}(x_n)) \tag{4}$$

If U and V are two uniformly distributed random variables for crash severity and crash size, respectively, the 2-dimentional copula can be written as:

$$C_\theta(u, v) = Prob(U < u,\ V < v) \tag{5}$$

where θ is the copula parameter vector also referred to as the dependence parameter. This parameter characterizes the dependence between the random variables for crash severity and crash size.

There are five common types of copula structures. These are the popular 1-parameter elliptical Gaussian copula related to the multivariate normal distribution and the 1-parameter Archimedean copulas, which include Clayton, Gumbel, Frank, and Joe copulas (Martey and Attoh-Okine, 2018b; Nelson, 2006). Both Gaussian and Frank copulas allow for positive and negative dependence. The other copulas, namely Clayton, Gumbel, and Joe copulas, do not allow for negative dependence. Also, the dependence in their tails is asymmetric. Among the copula structures, the Clayton copula is the best choice to consider if the two random variables for crash severity and crash size are strongly correlated at low values but not so correlated at high values (Hernández-Alava and Pudney, 2016). A recently introduced copula approach, Farlie-Gumbel-Morgenstein (FGM), which adopts a full spatial specification approach and has been shown to have unique traits with regards to its substantial computational gains compared to the previously discussed copula structures was also used in this study (Bhat and Sener, 2009). Together, the Frank, FGM, and Gaussian copula structures ensure high dependencies that are symmetric around the mean of the distribution. The symmetric tail dependence structure makes them more appropriate for use in crash analysis. In particular, when there is a symmetric upper and lower dependence in the tails, the Frank copula performs better relative to the other frameworks (Ayuso et al., 2016).

Copulas are invariant under monotone transformations. As such, it is better to use scale-invariant measures of dependence for assessing the degree of dependence between random variables (measures that remain unchanged under strictly increasing transformations). The common scale-invariant measures used for evaluating the degree of dependence are Kendal's tau and Spearman's rho. The Kendal's tau is mostly preferred among researchers due to its numerous advantages summarized in Bedford and Cooke (2001). The groups of copulas mentioned above are summarized in Table 2 (Hernández-Alava and Pudney, 2016; Wali et al., 2018).

2.3. Independent crash severity model component

In this study, for the purpose of comparison, crash severity is modeled using the traditional ordered probit (ORP) model structure. Assume that $q\ (q = 1, 2, \ldots, Q)$ is the index for each express bus-involved crash observation. For each observation, let $j\ (j = 0,1, \ldots, J)$ be the index representing the crash severity level y_q. For the three-scale ordinal crash severity level, $j = 1$ denotes PDO, $j = 2$ denotes evident injury, and $j = 3$ denotes fatal/severe injury. The unobserved/latent crash severity propensity y_q^* using the ORP model formulation is specified as a function of independent variables as follows:

$$y_q^* = \alpha' x_q + \varepsilon_q$$
$$y_q = j \text{ if } \ \tau_{j-1} < y_q^* < \tau_j \tag{6}$$

where y_q^* is the latent crash severity propensity for crash observation q, x_q is a vector of exogenous variables which may be associated with crash severity outcomes. α' is the vector of parameters to be estimated, ε_q is the random disturbance term assumed to be normal, and τ_j is the threshold associated with the crash severity levels j satisfying the condition $-\infty < \tau_1 < \tau_2 < \cdots < \tau_{J-1} < +\infty$ (Wang et al., 2015). From the above, the probability of the jth crash severity level of the qth express bus-involved crash can be formulated as follows:

$$\text{Prob}(y_q = j) = \Phi(\tau_j - \alpha' x_q) - \Phi(\tau_{j-1} - \alpha' x_q) \tag{7}$$

Table 2
Mathematical formulations and properties of copulas.

Copula Function $C(v, v)$	Formulation	Dependence parameter range	Kendal's tau $\tau(\theta)$	Kendal's tau range
JO	$1-[(1-u)^{\theta}+(1-v)^{\theta}-(1-u)^{\theta}(1-v)^{\theta}]^{\frac{1}{\theta}}$	$\theta \in (1, \infty)$	$1+\frac{4}{\theta^2}\int_0^1 t ln(t)(1-t)^{\frac{2(1-\theta)}{\theta}}dt$	$0 \le \tau < 1$
GU	$exp\left\{-[(-\ln u)^{\theta}+(-\ln v)^{\theta}]^{\frac{1}{\theta}}\right\}$	$\theta \in [1, \infty)$	$\frac{\theta-1}{\theta}$	$0 \le \tau < 1$
FR	$-\frac{1}{\theta}ln\left[1+\frac{(e^{-\theta u}-1)(e^{-\theta v}-1)}{e^{-\theta}-1}\right]$	$\theta \in (-\infty, \infty)\setminus\{0\}$	$1-\frac{4}{\theta}\left(1-\frac{1}{\theta}\int_0^{\theta}\frac{t}{e^t-1}dt\right)$	$-1 < \tau < 1$
CL	$(u^{-\theta}+v^{-\theta}-1)^{-\frac{1}{\theta}}$	$\theta \in [-1, \infty)\setminus\{0\}$	$\frac{\theta}{\theta+2}$	$0 \le \tau < 1$
FGM	$uv[1+\theta(1-u)(1-v)]$	$\theta \in [-1,1]$	$\frac{2\theta}{9}$	$-\frac{2}{9} \le \tau \le \frac{2}{9}$
GA	$\Phi(\Phi^{-1}(u), \Phi^{-1}(v); \theta)$	$\theta \in [-1,1]$	$\frac{2}{\pi}\sin^{-1}\theta$	$-1 \le \tau \le 1$

GA = Frank Copula; FR = Frank Copula; CL = Clayton Copula; FGM = Farlie-Gumbel-Morgenstein; GU = Gumbel Copula; JO = Joe Copula.

where $\Phi(\bullet)$ is the standard normal distribution function.

2.4. Independent crash size model component

The crash size model was also expressed via an ORP formulation where the latent propensity of crash size (u_q^*) for the q th crash observation is

$$u_q^* = \beta' z_q + \lambda_q$$
$$u_q = k \text{ if } \psi_{k-1} < u_q^* < \psi_k \tag{8}$$

where z_q is a vector of exogenous variables which may be associated with crash size outcomes. β' is the vector of parameter estimates, λ_q is the random disturbance term assumed to be normal, and ψ_k is the threshold associated with the crash size levels k ($k = 1$ represents express bus-involved crashes with one vehicle involved, $k = 2$ denotes express bus-involved crashes with crash size of two, and $k = 3$ indicates express bus-involved crashes with crash size three or more) satisfying the condition $-\infty < \psi_1 < \psi_2 < \cdots < \psi_{k-1} < +\infty$. Similarly, the probability the probability of the kth crash size in the q th express bus-involved crash observation can be expressed as:

$$\text{Prob}(u_q = k) = \Lambda(\psi_k - \beta' z_q) - \Lambda(\psi_{k-1} - \beta' z_q) \tag{9}$$

where $\Lambda(\bullet)$ is the standard normal distribution function.

2.5. Joint modeling of crash severity and crash size: application of the copula-based approach

The joint probability that an express bus-involved crash leads to crash severity level j and crash size k for crash observation q can be formulated as (Laman et al., 2018):

$$\text{Prob}(y_q = j,\ u_q = k) = \text{Prob}\left[\frac{[(\tau_{j-1} - \alpha' x_q < \varepsilon_q < \tau_j - \alpha' x_q)]}{[(\psi_{k-1} - \beta' z_q < \lambda_q < \psi_k - \beta' z_q)]}\right]$$
$$= \text{Prob}[\varepsilon_q < (\tau_j - \alpha' x_q), \lambda_q < (\psi_k - \beta' z_q)]$$
$$- \text{Prob}[\varepsilon_q < (\tau_j - \alpha' x_q), \lambda_q < (\psi_{k-1} - \beta' z_q)]$$
$$- \text{Prob}[\varepsilon_q < (\tau_{j-1} - \alpha' x_q), \lambda_q < (\psi_k - \beta' z_q)]$$
$$+ \text{Prob}[\varepsilon_q < (\tau_{j-1} - \alpha' x_q), \lambda_q < (\psi_{k-1} - \beta' z_q)] \tag{10}$$

Through the joint probability function above, the correlation between the residuals of crash severity and crash size, ε_q and λ_q are accommodated for using the copula formulation:

$$\text{Prob}(y_{qj} = j,\ u_{qk} = k) = C_\theta(\delta_{qj}, \delta_{qk}) - C_\theta(\delta_{qj}, \delta_{qk-1}) - C_\theta(\delta_{qj-1}, \delta_{qk}) + C_\theta(\delta_{qj-1}, \delta_{qk-1}) \tag{11}$$

$$\delta_{qj} = F_\varepsilon(\tau_j - \alpha' x_q),\ \delta_{qj-1} = F_\varepsilon(\tau_{j-1} - \alpha' x_q) \tag{12}$$

$$\delta_{qk} = F_\lambda(\psi_k - \beta' z_q),\ \delta_{qk-1} = F_\lambda(\psi_{k-1} - \beta' z_q) \tag{13}$$

where C_θ is a specific copula representation with dependency parameter θ, δ is a function of thresholds (τ_j and ψ_k) and parameter estimates (α and β). F_ε and F_λ represent the cumulative distribution functions of crash severity equation with residual term ε and crash size equation with residual term λ, respectively. The likelihood function can be computed as a summation of individual likelihoods for each express bus-involved crash observation (Wali et al., 2018):

$$\text{L}= \prod_{q=1}^{Q}\left\{\prod_{j=1}^{J}\prod_{k=1}^{K}(\text{Prob}(y_q = j,\ u_q = k))^{\Omega_{qkj}}\right\} \tag{14}$$

where
$$\Omega_{qkj} = \begin{cases} 1\ \textit{if qth crash results in the jth crash severity level with the kth crash size} \\ 0\ \textit{otherwise} \end{cases}$$

Detailed discussion of the copula methodology is described by Bhat and Eluru (2009), and Martey and Attoh-Okine (2018a).

3. Empirical setting

The focus of this study is to investigate the factors influencing crash severity and crash size of bus-involved crashes in South Korea using a bivariate copula-based methodology while addressing endogeneity due to correlations between both outcome variables. The analysis considered crash data that spans the period from 2010 to 2016. The study's crash data were grouped into three-time periods to explore the temporal stability of the coefficients (A= 2012 − 2013, B= 2014 − 2016, and T= 2012–2016).

Since many smaller cities are not connected to the South Korean rail network, the only way to travel to such places is by car or the express buses. In South Korea, traveling by express buses or express buses is the most popular way to move from one region to another due to its timeliness, reasonable price offerings, and accessibility owing to the extensive country-wide expressway network and small country size. Express-buses may make a stop only at rest areas during the journey. Per the South Korean regulations, express buses should have an engine output of more than 20 horsepower's per ton of vehicle total weight. At present, the average speed of the express buses is around 105 km/h. They have seating capacities ranging from 21 to 45, depending on the level of comfort or luxury (Ministry of Land, Infrastructure and Transport, 2020).

Expressway crash-related data was sourced from the Korean Expressway Corporation (KEC) database. In South Korea, trained crash investigators from the KEC are dispatched to the crash scene to collect crash-related data, which is compiled and managed for crash-risk assessment and safety management. Each reported crash is given a unique identification number to avoid duplicates (Hong et al., 2020a, 2019a). The data includes information for crash observations involving at least one bus traveling on one of the country's 38 expressway routes. The total length of the expressway in South Korea is approximately 4746 km, and the posted speed limits range from 100 km/h to 110 km/h. For the empirical analysis, the raw crash data were first preprocessed, and express bus-specific crash observations were extracted. At this stage, express bus-involved crash observations with missing information were omitted. For the five-year period (2012–2016), a total of 2997 express bus-involved crashes remained after the data preprocessing and were used for the analysis. For each crash observation in the preprocessed data, there existed 157 important variables describing the conditions and factors contributing to the crash. Since some of them have very few observations, they were combined to form meaningful variables. It is worth noting here that, in this study, a bus-involved crash is any roadway crash involving a bus and any other vehicle type, a bus and roadway facilities, or a bus and pedestrians.

The data comprise comprehensive information on variables such as crash severity level and crash size of each crash observations collected at the crash level, alongside the following:

- Temporal characteristics: time of the crash (12 a.m. to 5:59 am, 6 a.m. to 11:59 am, 12 pm to 5:59 pm, and 6 pm to 11:59 pm), day of the crash (weekday – Monday to Friday, and weekend – Saturday to Sunday), the season of the crash (spring – March to May, summer – June to August, fall – September to November, and winter – December to February).
- Crash location and roadway characteristics: roadway segment of crash (mainline, ramp, tollgate/tunnel – combined due to the small number of observations), pavement condition (normal, slippery, presence of potholes/fallen objects), horizontal alignment (straight without curves, curved segments – with radius > 1000 m or < = 1000 m), vertical alignment (grade, zero grade), median barrier types (fixed concrete median barrier (height - 81/127 cm), lawn median/median guardrails/movable median barrier, no median barrier), shoulder barrier types (shoulder guardrail, concrete/rock/pipe shoulder barrier, no shoulder barrier).
- Vehicle characteristics: the size of bus - based on the number of seats (small-sized - 21-seater, medium-sized – 28-seater, large-sized – 45-seater).
- Traffic characteristics: (logarithm of AADT, the proportion of trucks, the proportion of private cars, and the proportion of buses).
- Crash characteristics: crash type (vehicle to vehicle crash, vehicle to facility crash, vehicle to pedestrian crash).
- Weather characteristics (fine, inclement weather – comprising rainy, cloudy, snow/windy/foggy).
- Drivers characteristics: gender (male, female), age (<30 years, $30-49$ years, >49 years), drivers condition (normal, impaired by alcohol, drugs or tiredness), drivers faults (negligence, over speeding, wrongful passing/reversing/headway).

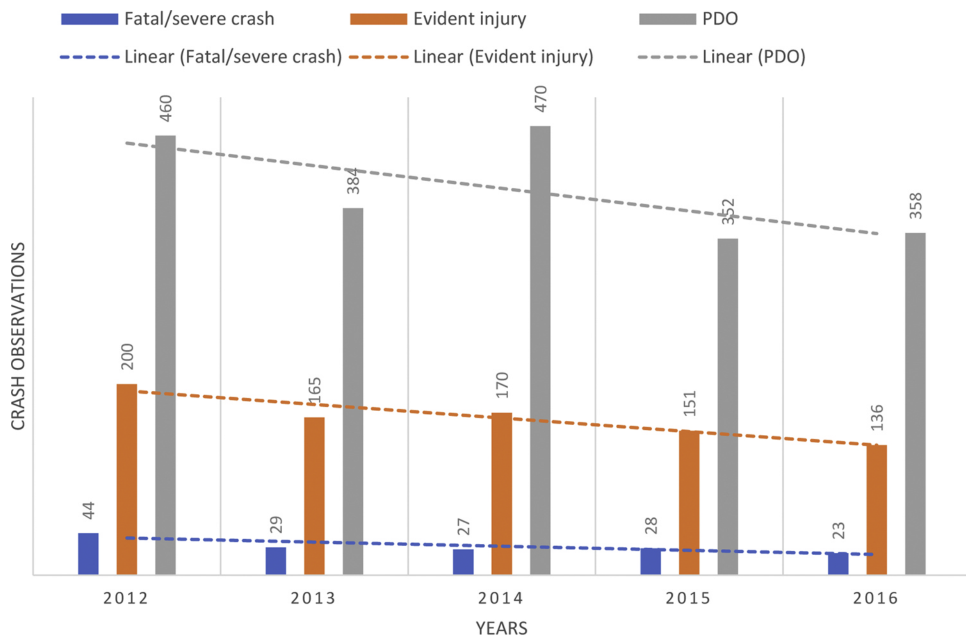
Considering the key focus of this study, both crash severity and crash size were considered as outcome variables. Due to the method of determination of the crash severity in South Korea by the authorities solely in charge of managing incidents on the Korean expressways, KEC, the unit of analysis of crash severity in this study is the entire crash (damage and loss of lives/injury caused by all vehicles in one crash). As discussed, the severity indicator is at the crash-level. However, since buses have unique features, particularly with regard to their body masses and number of occupants, they end up causing much harm when they are involved in crashes. Therefore, it is imperative to study bus-involved crashes at the crash-level and to identify valuable insights that can be employed in determining appropriate interventions that can be used to reduce them.

In South Korea, the KEC records crash/injury severity using a four-point ordinal scale which quantifies the severity of all entities that are involved in a single crash: 1) A – fatal injury (Number of deaths > 3, injured persons > 20 or damage cost > 1 billion won), 2) B – severe injury (1 < Number of deaths ≤ 3, 5 < injured persons ≤ 20 or 2.5 million won < damage cost ≤ 1 billion won), 3) C – evident injury (1 < injured persons ≤ 5 or 300 thousand won < damage cost ≤ 2.5 million won), and 4) D – property damage only, PDO (Damage cost ≤ 300 thousand won) (Hong et al., 2020a, 2019b). It is essential to note that crash severity, as recorded by the KEC, represents the total severity of all entities involved in one crash. As shown, the severity of each crash observation is based on the cost of damage, injury, or loss of human lives. Due to the low frequency of fatalities (approx. 0.42 % of the total crash observations), the variables for fatal and severe crash severity levels were amalgamated into one category (fatal or severe injury). It is important to note here that the severity outcome recorded for each crash observation in the dataset refers to the severity as a result of the crash involving either a bus and other vehicles, or a bus and roadway facilities.

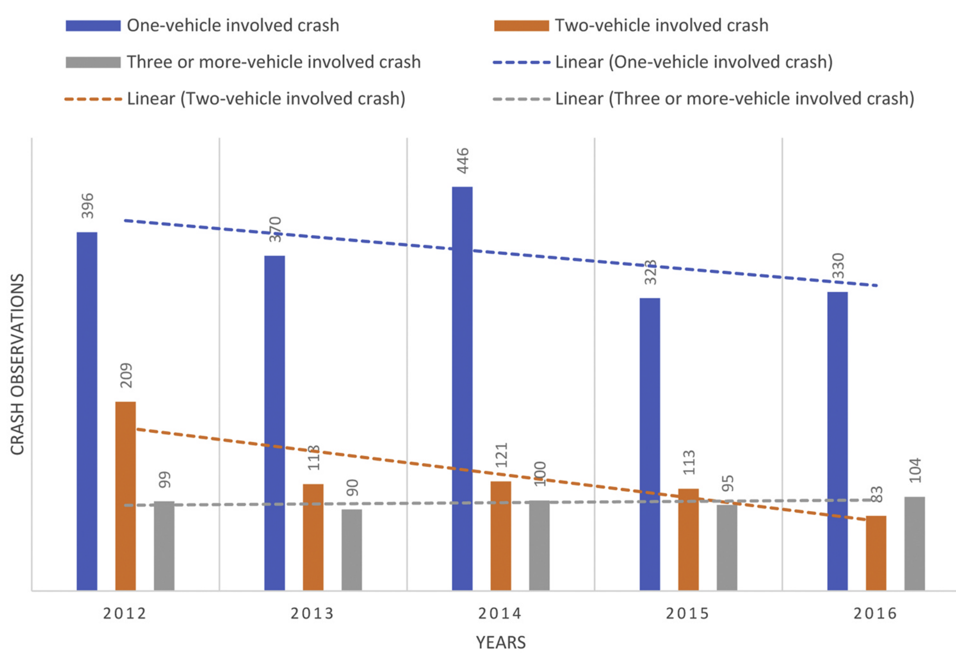
Concerning the crash size, the study identified that the number of vehicles in express bus-involved crashes ranged from 1 to 17. The crash size was classified into three discrete groups, ranging from one vehicle-involved crashes (counts – 1865 (62.23 %)), two vehicles involved in a crash (counts – 644 (21.49 %)), and three or more vehicles involved in a single crash (counts – 488 (16.28 %)). The last group was created to comprise all crashes involving three or more vehicles due to the lower proportion of crash observations with more than three vehicles involved. In this study, crash size connotes the number of vehicles involved in a crash. Hence, a crash size of level 1 (one vehicle-involved crash) represents a crash involving an express-bus and roadway facilities or a pedestrian. Crash size of level 2 (two vehicles-involved crashes) represents a crash involving an express-bus with another express-bus or any other vehicle type. Crash size of level 3 signifies a crash involving more than two vehicles, of which an express-bus is a part. In all, two/three or more vehicle-involved crashes connote multiple vehicle-involved crashes. Noting that each crash-size group shows unique severity profiles, the main reason for selecting the three groups is to find a distinction between one, two, and three or more vehicle crashes.

A figure showing the distribution of crash counts by crash severity level and crash size outcomes is presented (see Fig. 3). In general, there was no consistent pattern; however, overall, a slight and gradual reduction in crash severity and crash size in most years from 2012 to 2016 (with a sharp increase in 2014). Our discussions would be based on the trend line. From Fig. 3(a), crash observations with PDO and evident severities decreased for most of the years except for 2014 and 2016. Observations with fatal/severe crashes decreased slightly. In Fig. 3(b), while express bus-involved crash observations for both one-vehicle and two-vehicle involved crashes were decreasing (again, except that of 2014 in some cases), that of three or more vehicle-involved crashes increased significantly. A prior intuition that suggested the correlation between both outcome variables was tested. The product-moment correlation coefficient was estimated as $\rho = 0.3657$. This confirms some dependence between both outcome variables. This observation adequately motivates the need for employing the copula-based methodology to account for the dependence appropriately.

A contingency table for the joint frequency distribution of both outcome variables (crash severity and crash size) for the three-time periods ($2012-2013$, $2014-2016$, and $2012-2016$) studied in this paper is shown in Table 3. Clearly, fatal/severe crashes are frequent among two-vehicle, and three or more vehicles involved crash observations, relative to single-vehicle crashes. In essence, the number of fatal/severe express bus-involved crashes increased as the number of vehicles involved in the crash (crash size) increased. An opposite trend is observed with regards to PDO crashes. Also, but for 2016 alone, the number of vehicles involved in the crash decreased as the crash size increased. This suggests that the majority of vehicles were involved in vehicle-facility/object crashes on expressways. Preliminary analysis from Fig. 3 shows that the frequency of observations with two-vehicle involved crashes and fatal/severe crashes was highest in 2012 and lowest in 2016. Also, observations for single-vehicle crashes and those



(a) Crash severity levels



(b) Crash sizes

Fig. 3. Crash severity and crash size distribution.

resulting in PDO were both highest in 2014 and lowest in 2015. Table 4 provides a summary of the descriptive statistics of the variables of interest considered for the final model estimation.

4. Model comparison and selection

This study aims to explore the factors contributing to the crash outcomes and the dependence between crash severity and crash size of express bus-involved crash outcomes given some covariates. For the empirical analysis, the study explored six different copula structures, namely the Gaussian, Frank, FGM, Clayton, Gumbel, and Joe copulas, together with an independent ORP model that assumed no dependence between the two outcome variables studied in this paper. The independent ORP model serves as a starting point for the estimation of the joint copula models. A comparison exercise was conducted using the Bayesian information criterion (BIC), and Akaike information criterion (AIC) estimates to judge the models' performance. The BIC and AIC for a given empirical model are computed using the formulae below:

$$BIC = Kln(N) - 2LL \tag{15}$$

$$AIC = 2K - 2LL \tag{16}$$

Table 3
Contingency table for the joint distribution of crash severity and crash sizes for the different time periods (2012-2013, 2014-2016, and 2012-2016).

Year	Crash severity	Crash size						Total	
		Number of vehicles involved = 1		Number of vehicles involved = 2		Number of vehicles involved ≥ 3			
		Count	% of total	Count	% of total	Count	% of total	Count	% of total
2012 − 2013 (1282 crash observations)	**Fatal/severe**	20	0.667	26	0.868	27	0.901	73	2.436
	Evident	141	4.705	133	4.438	91	3.036	365	12.179
	PDO	605	20.187	168	5.606	71	2.369	844	28.161
Total		*766*	*25.559*	*327*	*10.911*	*189*	*6.306*	1282	*42.776*
2014 − 2016 (1715 crash observations)	**Fatal/severe**	19	0.634	30	1.001	29	0.968	78	2.603
	Evident	185	6.173	106	3.537	166	5.539	457	15.249
	PDO	895	29.863	181	6.039	104	3.470	1180	39.373
Total		1099	*36.670*	*317*	*10.577*	*299*	*9.977*	1715	*57.224*
2012 − 2016 (2997 crash observations)	**Fatal/severe**	39	1.301	56	1.869	56	1.869	151	5.038
	Evident	326	10.878	239	7.975	257	8.575	822	27.427
	PDO	1500	50.050	349	11.645	175	5.839	2024	67.534
Total		1865	*62.229*	*644*	*21.488*	*488*	*16.283*	2997	*100.000*

where LL is the log-likelihood at convergence, K is the number of estimated parameters, and N is the number of express bus involved-crash observations. The performance of the models is displayed for the three-time periods analyzed in this study are shown in Table 5.

The goodness-of-fit measures of all the models differ slightly. All the copula structures had lower LL, BIC and AIC estimates compared to the independent model, which shows that there exists some correlation caused by unobserved factors between crash severity and crash size. Accounting for these correlations improves model performance. The model structure with the lowest LL, AIC, and BIC was selected as the most suitable model.

Overall, based on the goodness-of-fit measures (Table 5), we observed that the Frank copula structure proved superior compared to the other copula structures and the independent bivariate model for the 2012 − 2013 and 2012 − 2016, and 2014 − 2016 datasets. This is because they had the lowest BIC estimates compared to the others. This confirms that the copula framework is more appropriate to account for the complex dependence between express bus-involved crash size and crash severities for the three-time periods compared to the independent model (A= 2012 − 2013, B= 2014 − 2016, and T= 2012 − 2016). The dependency parameter for the time periods $\hat{\theta}_A = 1.689$; $\hat{\theta}_B = 2.373$; $\hat{\theta}_T = 2.762$ were identified to be highly significant, and they correspond to Kendal's tau's of approximately 0.183, 0.250, and 0.286 ($-1 < \tau < 1$), respectively. The AIC and BIC values, together with the positive Kendal's tau values for the three-time periods suggests a strong positive correlation or dependency between crash severity and crash size of express bus involved crashes in South Korea.

5. Test for overall temporal stability of coefficients

The crash data was divided into different time periods in order to analyze the stability of coefficients over time. (A= 2012 − 2013, B= 2014 − 2016, and T= 2012–2016). In particular, this test could reveal the effect of the safety interventions that targeted buses after 2013. The total crash data was also modeled to compare the trends variable effects more comprehensively. Compared to crashes involving other vehicle types, the number of bus-involved crashes is low. The number of crash observations for the years 2012, 2013, 2014, 2015, and 2016 are 704, 578, 667, 531, and 517, respectively. As such, critical care was taken to ensure that a significant number of observations are maintained for each time period to ensure better forecast values. To ensure reduced bias, we think that dividing the data into groups A and B is appropriate to achieve our objective.

We employed a series of likelihood ratio tests to statistically explore the coefficients' stability over different time periods. The null hypothesis for this test is that, for the three-time periods, the parameters are equal. In equation form, the test statistic used for this test is shown below (Al-Bdairi et al., 2020; Islam and Mannering, 2020; Washington et al., 2003):

$$X^2 = -2[LL(\beta_T) - LL(\beta_A) - LL(\beta_B)] \tag{17}$$

where $LL(\beta_T)$ is the log-likelihood at convergence of a model estimated with data from both time periods under consideration ($A + B$). $LL(\beta_A)$ is the log-likelihood at convergence of the model using data from time period A. Similarly, $LL(\beta_B)$ is the log-likelihood at convergence of the model using data from time period B. The resulting X^2 obtained by using Eq. (18) is χ^2 distributed with degrees of freedom (df) equal to the difference between the sum of number of estimated parameters in time periods A and B, and the number of estimated parameters in the model with data from time period T (Washington et al., 2003). It should be noted that, only parameters that are statistically significant in the models should be counted (Islam and Mannering, 2020), and the same variables should enter all the models estimated (Al-Bdairi et al., 2020). The test was conducted using Eq. (18) by estimating both the independent ORP model and the Frank copula model previously identified to be superior relative to the other frameworks. In comparison, the log likelihood values obtained from using the independent ORP model ($T = -3{,}997.26$, $A = -1{,}770.44$, $B = -2{,}178.26$) were higher than those obtained from the copula-based model ($T = -3{,}893.21$, $A = -1{,}705.34$, $B = -2{,}097.49$). Their corresponding X^2 statistics gives us a 99 % confidence to fail to accept the null hypothesis that the parameters in the three-time periods are the same ($X^2_{ORP} = 99.10$, $\mathrm{df}_{ORP} = 15$; $X^2_{Copula} = 180.76$, $\mathrm{df}_{Copula} = 29$).

6. Discussion of estimation results

Tables 6–8 show the coefficient estimation results of both the independent ORP and Frank copula models for the three-year periods (2012 − 2013, 2014 − 2016, and 2012 − 2016). Only variables that were statistically significant at 90 % confidence level were selected and presented. From the tables, it is clear that the impact of the variables on the outcome variables follows the same trend under the independent ORP and Frank copula models; however, their magnitudes differ. This result is because the copula framework accounts for the complex interrelationships between both outcome variables used in the study. In each case (crash severity model/crash size model), the propensity to increase the level of crash severity or crash size is represented by a positive coefficient, and a negative coefficient shows a propensity to decrease the level of crash severity or crash size. Since the parameter estimates from both models alone are not adequate to describe the magnitude of the effect of regressors on the probability of crash severity and crash size outcomes, we estimated the marginal effects for all

Table 4
Descriptive statistics of explanatory variables for express bus-involved crashes.

Variable	2012 – 2013		2014 – 2016		2012 – 2016	
	Percentage	Count	Percentage	Count	Percentage	Count
Temporal characteristics						
Season						
Spring (1 if crash occurred from March to May; 0 otherwise)	25 %	324	23 %	397	24 %	721
Summer (1 if crash occurred from June to August; 0 otherwise)	31 %	400	25 %	436	28%	836
Fall (1 if crash occurred from September to November; 0 otherwise)	22 %	276	26 %	453	24 %	729
Winter (1 if crash occurred from December to February; 0 otherwise)	22 %	282	25 %	429	24 %	711
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Day of crash						
Weekday (1 if crash occurred on weekdays; 0 otherwise)	71 %	912	71 %	1226	71 %	2138
Weekend (1 if crash occurred on weekends; 0 otherwise)	29 %	370	29 %	489	29 %	859
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Time of crash						
Midnight to early morning (1 if crash occurred from 12am to 5:59am; 0 otherwise)	12 %	148	12 %	201	12 %	349
Morning peak to late morning (1 if crash occurred from 6am to 11:59am; 0 otherwise)	33 %	420	30 %	507	31 %	927
Noon to late afternoon (1 if crash occurred from 12 pm to 5:59pm; 0 otherwise)	34 %	441	36%	621	35 %	1062
Evening peak/ to late night (1 if crash occurred from 6 pm to 11:59pm; 0 otherwise)	21 %	273	23 %	386	22 %	659
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Drivers characteristics						
Gender						
Male (1 if driver is male; 0 otherwise)	91 %	1163	91 %	1556	91 %	2719
Female (1 if driver is female; 0 otherwise) (1 if driver is male; 0 otherwise)	9%	119	9%	159	9%	278
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Age						
Young (1 if driver is young < 30; 0 otherwise)	26 %	337	37%	631	32%	968
Middle aged (1 if driver is middle-aged 30 – 49; 0 otherwise)	40 %	514	29 %	505	34 %	1019
Old (1 if driver is old (≥50); 0 otherwise)	34 %	431	34 %	579	34 %	1010
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Drivers fault						
Negligence (1 if driver was negligent; 0 otherwise)	25 %	323	26 %	443	26 %	766
Over speeding (1 if driver was over speeding; 0 otherwise)	23 %	297	22 %	377	22 %	674
Wrongful passing, reversing, and headway (1 if driver was indulging in wrongful passing, reversing, and headway; 0 otherwise)	52 %	662	52 %	895	52 %	1557
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Drivers condition						
Normal (1 if driver was normal at the time of crash; 0 otherwise)	91 %	1169	93%	1594	92%	2763
Impaired (1 if driver was tired or under the influence of alcohol or drugs; 0 otherwise)	9%	113	7%	121	8%	234
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Weather characteristics						
Weather during crash						
Fine weather (1 if weather was normal at the time of the crash; 0 otherwise)	58 %	739	61%	1043	59%	1782
Inclement weather (1 if weather was cloudy, foggy, rainy, snowy; 0 otherwise)	42 %	543	39%	672	41%	1215
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Crash characteristics						
Crash type						
Vehicle-facility crash (1 if bus crushed into roadway facilities; 0 otherwise)	55 %	702	56%	959	55 %	1661
Vehicle-vehicle crash (1 if bus crushed with another vehicle; 0 otherwise)	24 %	314	22 %	384	23 %	698
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	21 %	266	22 %	372	21 %	638
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Crash location and roadway characteristics						
Pavement condition						
Slippery (1 if pavement was slippery; 0 otherwise)	12 %	153	14 %	246	13%	399
Normal pavement (1 if pavement was normal; 0 otherwise)	82 %	1047	80%	1374	81%	2421
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	6%	82	6%	95	6%	177
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Roadway segment of crash						
Mainline (1 if crash occurred on mainline section; 0 otherwise)	76 %	970	70%	1208	73%	2178
Ramp (1 if crash occurred at ramp section; 0 otherwise)	11 %	135	11 %	196	11 %	331
Tollgate/tunnel (1 if crash occurred at tollgate/tunnel section; 0 otherwise)	14 %	177	18%	311	16%	488
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Horizontal curve						
Straight (1 if crash occurred on curve-less roadway section; 0 otherwise)	78 %	1004	79%	1362	79%	2366
Curved (1 if crash occurred on roadway with horizontal curves; 0 otherwise)	22 %	278	21 %	353	21 %	631
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*

(*continued on next page*)

Table 4 (*continued*)

Variable	2012 – 2013		2014 – 2016		2012 – 2016	
	Percentage	Count	Percentage	Count	Percentage	Count
Vertical alignment						
Grade (1 if crash occurred on roadway with vertical curves; 0 otherwise)	34 %	442	36%	615	35 %	1057
Zero grade (1 if crash occurred on roadway had no grade; 0 otherwise) (1 if crash occurred on roadway with vertical curves; 0 otherwise)	66 %	840	64%	1100	65%	1940
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Median barrier type						
Fixed concrete median barrier (1 if crash occurred on roadway section with fixed concrete barrier of height 81/127 cm; 0 otherwise)	62 %	789	58 %	992	59%	1781
Other median barrier types (1 if crash occurred on roadway section with median lawns, movable guardrail; 0 otherwise)	17 %	213	18%	310	17 %	523
No median barrier (1 if crash occurred on roadway section with no median barrier; 0 otherwise)	22 %	280	24 %	413	23 %	693
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Shoulder barrier type						
Guardrail shoulder barrier type (1 if crash occurred on roadway section with guardrail shoulder barrier type barrier type; 0 otherwise)	42 %	540	45%	766	44%	1306
Other shoulder barrier types (1 if crash occurred on roadway section with concrete/rock/pipe shoulder barrier type; 0 otherwise)	23 %	296	20%	350	22 %	646
No shoulder barrier (1 if crash occurred on roadway section with no shoulder barrier; 0 otherwise)	35 %	446	35 %	599	35 %	1045
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Vehicle characteristics						
Size of bus						
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	50 %	636	50 %	860	50 %	1496
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	30 %	380	28%	472	28%	852
Large-sized (1 if bus was small-sized: 45-seater coach; 0 otherwise)	21 %	266	22 %	383	22 %	649
Total	100 %	*1282*	100 %	*1715*	100 %	*2997*
Traffic characteristics	**Avg./St. Dev./Min/Max**		**Avg./St. Dev./Min/Max**		**Avg./St. Dev./Min/Max**	
Logarithm of AADT	10.673/0.789/7.045/12.426		10.718/0.780/6.936/12.350		10.699/0.784/6.936/12.426	
Proportion of private cars	0.673/0.081/0.385/0.883		0.693/0.069/0.384/0.895		0.684/0.075/0.384/0.895	
Proportion of trucks	0.042/0.018/0.003/0.114		0.036/0.018/0.005/0.113		0.038/0.018/0.003/0.114	
Proportion of buses	0.285/0.082/0.083/0.581		0.271/0.071/0.070/0.588		0.277/0.076/0.070/0.588	

Table 5
Model performance measures.

Model	2012 – 2013				2014 – 2016				2012 – 2016			
	LL	K	AIC	BIC	LL	K	AIC	BIC	LL	K	AIC	BIC
GA	−1917.47	30	3894.94	4049.62	−2476.64	30	5013.28	5176.69	−4394.73	30	8859.45	9069.64
FR	**−1917.23**	30	**3894.46**	**4049.14**	**−2473.55**	30	**5007.11**	**5170.52**	**−4394.31**	30	**8858.62**	**9068.81**
CL	−1918.77	30	3897.54	4052.23	−2474.93	30	5009.85	5173.26	−4398.56	30	8867.12	9077.31
FGM	−1920.37	30	3900.74	4055.43	−2476.41	30	5012.81	5176.23	−4419.16	30	8898.32	9078.48
GU	−1919.23	30	3898.46	4053.15	−2485.37	30	5030.73	5194.15	−4402.75	30	8875.50	9085.68
JO	−1920.91	30	3901.82	4056.50	−2490.89	30	5041.78	5205.20	−4409.10	30	8888.19	9098.38
IND	−1930.20	29	3918.41	4067.94	−2513.25	29	5084.50	5242.46	−4443.92	29	8955.83	9160.02

GA = Frank Copula; FR = Frank Copula; CL = Clayton Copula; FGM = Farlie-Gumbel-Morgenstein; GU = Gumbel Copula; JO = Joe Copula; IND = Independent; K = number of parameters.

regressors in the three-year groups with regards to both outcome variables as shown in Tables 9–11. Again, it is worth noting that we considered only statistically significant variables. Given a dummy or continuous regressor, the marginal effect measures either show a discrete or instantaneous rate of change, respectively. It essentially shows how probabilities change as the dummy regressor changes from 0 to 1 or the continuous regressor changes by 1 unit.

Once more, crash severity, as defined by the KEC, factors in the amount of damage caused to properties and humans, whereas crash size defines the number of vehicles involved in a single crash. These outcome variables have been explained in detail in Section 3. It is worth noting that not all the variables studied in this research were used in the model, as some were used as baselines in the model estimation. We begin this section by discussing the temporal stability of the model parameters, the copula parameters, and the various parameters in the model estimation results together with the marginal effects. The explanation of the results will focus on those from the copula models.

6.1. *Temporal stability of model parameters*

With regards to the temporal stability tests, the results from both models show that the variables in the crash data are not stable across different time periods studied, hence, estimating different models for the different time periods is warranted. In the literature, similar results were obtained by researchers. Several researchers also concluded that different models should be estimated for different time periods (Behnood and Mannering, 2016; Islam and Mannering, 2020). The temporal instability of the model's parameters could be as a result of several factors. It is difficult to identify the primary source of temporal instability. Some studies have suggested that temporal instability could be as a result of either the long-term changes in factors that affect the outcome variables studied in the models or economic recession (Al-

Table 6
Model estimation results for the period 2012-2013.

Variable	Independent model				Frank copula-based joint model			
	Crash severity		Crash size		Crash severity		Crash size	
	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score
Temporal characteristics								
Season (baseline: winter)								
Summer (1 if crash occurred from June to August; 0 otherwise)	−0.186	−2.33	−0.240	−3.09	−0.191	−2.08	−0.206	−2.60
Time of crash (baseline: midnight to early morning −12 am to 5:59am)								
Morning peak to late morning (1 if crash occurred from 6am to 11:59am; 0 otherwise)	0.265	3.52	–	–	0.259	2.93	–	–
Drivers characteristics								
Age (baseline: old driver (≥50))								
Young (1 if driver is young < 30; 0 otherwise)	−0.157	−1.85	0.546	6.79	−0.180	−1.82	0.489	6.06
Drivers fault (baseline: wrongful passing, reversing, and headway)								
Negligence (1 if driver was negligent; 0 otherwise)	0.292	3.41	0.576	6.84	0.303	2.96	0.578	5.26
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	−0.235	−2.30	–	–	−0.293	−2.50
Drivers condition (baseline: impaired)								
Normal (1 if driver was normal at the time of crash; 0 otherwise)	−0.599	−5.06	–	–	−0.637	−5.11	–	–
Weather characteristics								
Weather during crash (baseline: fine weather)								
Inclement weather (1 if weather was cloudy, foggy, rainy, snowy; 0 otherwise)	–	–	−0.207	−2.74	–	–	−0.181	−2.38
Crash characteristics								
Crash type (baseline: vehicle-vehicle crash)								
Vehicle-facility crash (1 if bus crushed into roadway facilities; 0 otherwise)	−0.387	−3.54	–	–	−0.357	−2.93	–	–
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	−0.541	−4.59	−0.564	−5.89	−0.571	−4.23	−0.657	−6.27
Crash location and roadway characteristics								
Pavement condition (baseline: normal pavement)								
Slippery (1 if pavement was slippery; 0 otherwise)	−0.284	−2.40	–	–	−0.380	−2.59	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	−0.532	−3.14	–	–	−0.649	−2.90	–	–
Roadway segment of crash (baseline: tollgate/tunnel)								
Mainline (1 if crash occurred on mainline section; 0 otherwise)	0.409	4.38	0.651	5.42	0.500	4.03	0.830	5.50
Median barrier type (baseline: no median barrier)								
Fixed concrete median barrier (1 if crash occurred on roadway section with fixed concrete barrier of height 81/127 cm; 0 otherwise)	–	–	0.226	2.30	–	–	0.237	2.21
Shoulder barrier type (baseline: guardrail shoulder barrier type)								
Other shoulder barrier types (1 if crash occurred on roadway section with concrete/rock/pipe shoulder barrier type; 0 otherwise)	–	–	0.203	2.54	–	–	0.179	2.31
Vehicle characteristics								
Size of bus (baseline: large-sized)								
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	−0.416	−4.36	−0.672	−7.34	−0.488	−4.65	−0.738	−6.59
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	−0.268	−2.65	−0.287	−2.95	−0.299	−2.82	−0.310	−3.32
Traffic characteristics								
Proportion of trucks	6.938	3.47	6.438	3.25	8.175	3.72	6.554	3.02
Thresholds, log-likelihood at convergence and copula parameter								
Threshold 1	−0.105	−0.54	1.243	6.39	−0.174	−0.77	1.449	5.53
Threshold 2	0.781	5.41	1.765	11.78	0.777	5.06	1.884	9.33
Log-likelihood at convergence	−1917.469				−1903.164			
Dependence	Copula dependence parameter (θ)				1.689	5.07		
	Kendal's tau $\tau(\theta)$				0.183			

Bdairi et al., 2020; Behnood and Mannering, 2016). In particular, the instability of a model's parameters is usually attributed to changes in driver behavior and driving patterns (Wang et al., 2019).

In this study, the significant differences in parameters given the different time periods can be attributed to the effect of the 7th National Transport Safety Master Plan and the promotion of safe roundabouts championed by the Government of Korea in both 2012 and 2013, respectively. In line with past literature, improved educational programs, enforcement of driving rules, and improved roadway designs leads to changes in risky driving behavior and the reduction in driver error over time (Mannering, 2018). As a result of the Government intervention, the number and roadway fatalities slightly reduced in 2016 compared to that of 2012, as previously shown in Fig. 1.

6.2. Copula parameters

Since temporal instability in the models was identified, it was valid to estimate models for the three-time periods. As mentioned earlier, the Frank copula-based model provided the best fit, and three different values of copula parameters were obtained for the time periods studied. To deeply investigate the dependence, we computed their corresponding Kendal tau estimates since they are invariant under strictly increasing transformations (a property that Pearson's correlation lacks (Melchiori, 2003)). From the estimated parameters, insights about the potential correlations between both outcome variables can be obtained.

In essence, the results provided evidence to support our initial hypothesis and highlight the presence of some level of dependency among the outcome variables. Since the Frank copula dependency structure is symmetric, positive values observed in the models denote a positive

Table 7
Model estimation results for the period 2014-2016.

Variable	Independent model				Frank copula-based joint model			
	Crash severity		Crash size		Crash severity		Crash size	
	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score
Temporal characteristics								
Season (baseline: winter)								
Spring (1 if crash occurred from March to May; 0 otherwise)	0.262	3.68	–	–	0.267	3.78	–	–
Drivers characteristics								
Age (baseline: old driver (≥50))								
Young (1 if driver is young < 30; 0 otherwise)	−0.246	−3.56	0.301	4.30	−0.244	−3.53	0.294	0.29
Drivers fault (baseline: wrongful passing, reversing, and headway)								
Negligence (1 if driver was negligent; 0 otherwise)	–	–	0.195	2.43	–	–	0.206	0.21
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	−0.218	−2.19	–	–	−0.216	−0.22
Normal (1 if driver was normal at the time of crash; 0 otherwise)	−0.389	−3.32	0.318	2.26	−0.389	−3.31	0.310	0.31
Crash characteristics								
Crash type (baseline: vehicle-vehicle crash)								
Vehicle-facility crash (1 if bus crushed into roadway facilities; 0 otherwise)	−0.858	−10.93	−2.035	−23.49	−0.855	−10.90	−2.031	−2.03
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	−0.880	−8.77	−1.827	−17.37	−0.870	−8.67	−1.823	−1.82
Crash location and roadway characteristics								
Pavement condition (baseline: normal pavement)								
Slippery (1 if pavement was slippery; 0 otherwise)	−0.203	−2.12	–	–	−0.208	−2.16	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	−0.554	−3.04	0.387	2.70	−0.542	−3.00	0.387	0.39
Roadway segment of crash (baseline: tollgate/tunnel)								
Mainline (1 if crash occurred on mainline section; 0 otherwise)	0.288	3.53	0.506	5.26	0.299	3.66	0.520	0.52
Ramp (1 if crash occurred at ramp section; 0 otherwise)	–	–	−0.298	−1.80	–	–	−0.276	−0.28
Shoulder barrier type (baseline: guardrail shoulder barrier type)								
No shoulder barrier (1 if crash occurred on roadway section with no shoulder barrier; 0 otherwise)	−0.141	−1.93	–	–	−0.143	−1.96	–	–
Vehicle characteristics								
Size of bus (baseline: large-sized)								
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	−0.238	−2.98	−0.184	−2.71	−0.242	−3.06	−0.187	−0.19
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	−0.177	−2.04	–	–	−0.163	−1.89	–	–
Traffic characteristics								
Logarithm of AADT	0.197	4.58	0.189	4.17	0.198	4.59	0.185	0.19
Proportion of private cars	−1.120	−2.41	–	–	−1.102	−2.38	–	–
Thresholds, log-likelihood at convergence and copula parameter								
Threshold 1	0.751	1.44	2.149	4.10	0.783	1.50	2.186	4.18
Threshold 2	1.612	3.18	2.589	5.09	1.574	3.11	2.551	5.02
Log-likelihood at convergence	−2155.599				−2153.602			
Dependence	Copula dependence parameter (θ)				2.373	8.17		
	Kendal's tau $\tau(\theta)$				0.250			

correlation among the unobserved factors unknown to analysts but influence both the severity and size of express bus-involved crashes. As such, considering a model that adequately accounts for unobserved factors relating to driver behavior, roadway geometry, or vehicle condition in crash data, as used in this current study, is justified. Identifying this dependency and using the proposed copula methodology enabled us to improve upon the performance of the model estimation results.

6.3. Temporal characteristics

The impact of temporal characteristics shows significant variations based on the variable for the time of crash and season of crash. Results from the models reveal that, compared to crashes that occur at midnight until dawn, those that occur in the morning from 6am to 11:59am are likely to have increased severity levels and those occurring in the afternoon from 12 pm to 5:59pm are likely to have increased crash sizes (Table 6 and 8). In particular, from the marginal effect estimates, the probability of evident injury increases by the highest margin when the crash occurs in the morning, whereas the probability for three or more vehicle-involved crashes occurring increases highest in the afternoon (Table 9 and 11). Our result slightly deviates from what has been reported in the literature on expressway related bus-involved crashes. Bus crashes on expressways have been identified as likely to be severe when it occurs from late in the night to very early in the morning (11 PM to 5 AM) (Chen and Jovanis, 2000). The reason for the deviation can be attributed to several reasons. First, due to the small land area of South Korea, travel time with express-buses is usually short, and are mostly rife during the day. Statistics from the KEC's crash data revealed that the majority of the bus-involved crashes on expressways occurred during the day, from 6am to 5:59pm (2012 − 2016 main crash data – 1989 out of 2997 bus-involved crashes). However, due to the increased speeds on the expressway, coupled with the many trips in the morning, it is reasonable to observe crashes with increased severities. Also, since traffic volume is slightly reduced in the afternoon, bus drivers may become overly relaxed as they drive. These drivers are likely to indulge in risky driving, which may cause them to run into other vehicles and lead to increased crash sizes. Pai and Saleh (2008) also noted that it is possible to have increased crash severities in the morning periods due to increased speeds and drug use during that period. In comparison to transit bus-involved crashes, Barua and Tay (2010) also noted that crash severity is much higher during the morning off-peak period (6 AM

Table 8
Model estimation results for the period 2012-2016.

Variable	Independent model				Frank copula-based joint model			
	Crash severity		Crash size		Crash severity		Crash size	
	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score	Coeff.	z-score
Temporal characteristics								
Season (baseline: winter)								
Spring (1 if crash occurred from March to May; 0 otherwise)	0.189	3.68	–	–	0.158	3.25	–	–
Summer (1 if crash occurred from June to August; 0 otherwise)	–	–	−0.144	−2.83	–	–	−0.150	−2.95
Time of crash (baseline: midnight to early morning − 12 am to 5:59am)								
Noon to late afternoon (1 if crash occurred from 12 pm to 5:59pm; 0 otherwise)	–	–	0.085	1.82	–	–	0.090	1.92
Drivers characteristics								
Age (baseline: old driver (≥50))								
Young (1 if driver is young < 30; 0 otherwise)	−0.127	−2.45	0.475	9.45	−0.087	−1.66	0.478	9.5
Drivers fault (baseline: wrongful passing, reversing, and headway)								
Negligence (1 if driver was negligent; 0 otherwise)	0.297	5.50	0.499	9.02	0.256	4.84	0.491	8.94
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	−0.325	−4.70	–	–	−0.339	−4.86
Drivers condition (baseline: impaired (alcohol, drugs, tired))								
Normal (1 if driver was normal at the time of crash; 0 otherwise)	−0.528	−6.69	–	–	−0.460	−5.99	–	–
Weather characteristics								
Weather during crash (baseline: fine)								
Inclement weather (1 if weather was cloudy, foggy, rainy, snowy; 0 otherwise)	–	–	−0.116	−2.30	–	–	−0.114	−2.27
Crash characteristics								
Crash type (baseline: vehicle-vehicle crash)								
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	−0.296	−4.58	−0.571	−8.78	−0.348	−4.76	−0.591	−9.09
Crash location and roadway characteristics								
Pavement condition (baseline: normal)								
Slippery (1 if pavement was slippery; 0 otherwise)	−0.277	−3.84	–	–	−0.310	−3.80	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	−0.505	−4.17	0.285	2.81	−0.557	−3.86	0.284	2.81
Roadway segment of crash (baseline: tollgate/tunnel section)								
Mainline (1 if crash occurred on mainline section; 0 otherwise)	0.486	8.58	0.633	7.49	0.490	7.49	0.641	7.60
Ramp (1 if crash occurred at ramp section; 0 otherwise)	–	–	−0.332	−2.95	–	–	−0.319	−2.83
Vertical alignment (baseline: zero-grade)								
Grade (1 if crash occurred on roadway with vertical curves; 0 otherwise)	–	–	−0.124	−2.49	–	–	−0.128	−2.57
Median barrier type (baseline: no median barrier)								
Fixed concrete median barrier (1 if crash occurred on roadway section with fixed concrete barrier of height 81/127 cm; 0 otherwise)	–	–	0.156	2.47	–	–	0.154	2.44
Shoulder barrier type (baseline: guardrail shoulder barrier type)								
Other shoulder barrier types (1 if crash occurred on roadway section with concrete/rock/pipe shoulder barrier type; 0 otherwise)	0.152	2.73	0.133	2.41	0.145	2.80	0.119	2.16
Vehicle characteristics								
Size of bus (baseline: large-sized - 45-seater)								
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	−0.406	−6.82	−0.504	−8.49	−0.387	−6.03	−0.513	−8.63
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	−0.277	−4.24	−0.230	−3.55	−0.243	−3.97	−0.224	−3.49
Traffic characteristics								
Logarithm of AADT	0.142	4.53	0.110	3.40	0.147	4.54	0.107	3.33
Proportion of private cars	−0.934	−2.91	−0.720	−2.25	−0.918	−2.92	−0.689	−2.14
Proportion of trucks	4.298	3.3	5.025	3.82	3.826	3.05	4.903	3.73
Thresholds, log-likelihood at convergence and copula parameter								
Threshold 1	0.715	1.86	2.063	5.37	1.133	3.21	2.579	6.88
Threshold 2	1.508	4.06	2.316	6.21	1.463	3.89	2.275	6.03
Log-likelihood at convergence	−4398.859				−4389.214			
Dependence	Copula dependence parameter (θ)				2.762	13.97		
	Kendal's tau $\tau(\theta)$				0.286			

– 9 AM) relative to those occurring during the morning peak period (9 AM – 12 PM).

Results obtained from running both 2014−2016 and 2012−2016 models revealed that crashes during the spring season are likely to have an increased severity compared to those that occurred during the winter season. The probability of evident injury is seen to be the highest among the three crash severity outcomes (see Tables 10 and 11). Unlike winter seasons where drivers drive slowly when snow falls, the hazards of rain during the spring season, which presents issues such as reduced visibility for drivers and hydroplaning (occurs when the tires of a vehicle loose grip of the pavement and skids off the road atop water puddles) causes an increase in crash counts and severity. Studies have also shown that there is a high chance of roadways developing potholes right after the harsh winter conditions. Also, due to the shift from winter, most animals would return to their normal lives. As such, it is expected to see animals crossing the roadways during this season. If drivers fail to remove distractions as they drive along expressways, they may not be prepared to deal with these spring season-related hazards when they spot them late. This result was consistent with the findings in the literature (Savolainen and Ghosh, 2008; Wang and Prato, 2019; Zhang et al., 2016).

Furthermore, the results from the 2012−2013 models showed that

Table 9
Marginal effects for model estimates (2012-2013).

Variable	Independent model						Frank copula-based joint model					
	Crash severity			Crash size			Crash severity			Crash size		
	0	1	2	0	1	2	0	1	2	0	1	2
Temporal characteristics												
Season (baseline: winter)												
Summer (1 if crash occurred from June to August; 0 otherwise)	0.0655	−0.0519	−0.0136	0.0902	−0.0504	−0.0398	0.0482	−0.0332	−0.0151	0.0575	−0.0279	−0.0296
Time of crash (baseline: midnight to early morning −12 am to 5:59am)												
Morning peak to late morning (1 if crash occurred from 6am to 11:59am; 0 otherwise)	−0.0967	0.0742	0.0225	–	–	–	−0.0743	0.0506	0.0238	–	–	–
Drivers characteristics												
Age (baseline: old driver (≥50))												
Young (1 if driver is young < 30; 0 otherwise)	0.0555	−0.0440	−0.0115	−0.2122	0.0997	0.1126	0.0493	−0.0340	−0.0153	−0.1735	0.0943	0.0792
Drivers fault (baseline: wrongful passing, reversing, and headway)												
Negligence (1 if driver was negligent; 0 otherwise)	−0.1078	0.0817	0.0260	−0.2239	0.1034	0.1205	−0.0889	0.0603	0.0286	−0.2210	0.1227	0.0983
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	–	0.0878	−0.0497	−0.0380	–	–	–	0.0778	−0.0371	−0.0407
Drivers condition (baseline: impaired (alcohol, drugs, tired))												
Normal (1 if driver was normal at the time of crash; 0 otherwise)	0.2302	−0.1581	−0.0721	–	–	–	0.2325	−0.1583	−0.0740	–	–	–
Weather characteristics												
Weather during crash (baseline: fine)												
Inclement weather (1 if weather was cloudy, foggy, rainy, snowy; 0 otherwise)	–	–	–	0.0787	−0.0430	−0.0357	–	–	–	0.0569	−0.0280	−0.0289
Crash characteristics												
Crash type (baseline: vehicle-vehicle crash)												
Vehicle-facility crash (1 if bus crushed into roadway facilities; 0 otherwise)	0.1393	−0.1077	−0.0316	–	–	–	0.1060	−0.0726	−0.0335	–	–	–
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	0.1772	−0.1448	−0.0324	0.2001	−0.1200	−0.0801	0.1222	−0.0873	−0.0350	0.1630	−0.0778	−0.0852
Crash location and roadway characteristics												
Pavement condition (baseline: normal)												
Slippery (1 if pavement was slippery; 0 otherwise)	0.0963	−0.0781	−0.0183	–	–	–	0.0781	−0.0551	−0.0230	–	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	0.1661	−0.1386	−0.0275	–	–	–	0.1394	−0.1044	−0.0350	–	–	–
Roadway segment of crash (baseline: tollgate/tunnel section)												
Mainline (1 if crash occurred on mainline section; 0 otherwise)	−0.1385	0.1118	0.0267	−0.2299	0.1373	0.0926	−0.1200	0.0851	0.0350	−0.2175	0.1074	0.1101
Median barrier type (baseline: no median barrier)												
Fixed concrete median barrier (1 if crash occurred on roadway section with fixed concrete barrier of height 81/127 cm; 0 otherwise)	–	–	–	−0.0855	0.0471	0.0384	–	–	–	−0.0672	0.0329	0.0343
Shoulder barrier type (baseline: guardrail shoulder barrier type)												
Other shoulder barrier types (1 if crash occurred on roadway section with concrete/rock/pipe shoulder barrier type; 0 otherwise)	–	–	–	−0.0785	0.0403	0.0382	–	–	–	−0.0610	0.0314	0.0296
Vehicle characteristics												
Size of bus (baseline: large-sized - 45-seater)												
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	0.1486	−0.1157	−0.0329	0.2520	−0.1331	−0.1189	0.1215	−0.0836	−0.0380	0.2358	−0.1229	−0.1129
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	0.0935	−0.0745	−0.0190	0.1072	−0.0604	−0.0468	0.0715	−0.0495	−0.0220	0.0872	−0.0420	−0.0452
Traffic characteristics												
Proportion of trucks*	−2.4919	1.9513	0.5406	−2.4567	1.3277	1.1290	−2.1137	1.4428	0.6717	−2.0777	1.0292	1.0485

Notes: "-" represents "not applicable"; Crash severity (1 = PDO, 2 = evident injury, 3 = fatal/severe injury); Crash size (1 = one vehicle involved crash, 2 = two vehicles involved in a crash, 3 = three or more vehicles involved in a crash); variables with (*) are continuous, while the others are dummy variables.

Table 10
Marginal effects for model estimates (2014-2016).

Variable	Independent model						Frank copula-based joint model					
	Crash severity			Crash size			Crash severity			Crash size		
	0	1	2	0	1	2	0	1	2	0	1	2
Temporal characteristics												
Season (baseline: winter)												
Spring (1 if crash occurred from March to May; 0 otherwise)	−0.0927	0.0748	0.0179	–	–	–	−0.0949	0.0765	0.0184	–	–	–
Drivers characteristics												
Age (baseline: old driver (≥50))												
Young (1 if driver is young < 30; 0 otherwise)	0.0824	−0.0686	−0.0138	−0.1095	0.0642	0.0453	0.0820	−0.0683	−0.0137	−0.1070	0.0626	0.0444
Drivers fault (baseline: wrongful passing, reversing, and headway)												
Negligence (1 if driver was negligent; 0 otherwise)	–	–	–	−0.0712	0.0416	0.0296	–	–	–	−0.0757	0.0440	0.0316
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	–	0.0759	−0.0475	−0.0284	–	–	–	0.0752	−0.0469	−0.0282
Normal (1 if driver was normal at the time of crash; 0 otherwise)	0.1433	−0.1114	−0.0320	−0.1059	0.0689	0.0370	0.1435	−0.1114	−0.0321	−0.1035	0.0672	0.0364
Crash characteristics												
Crash type (baseline: vehicle-vehicle crash)												
Vehicle-facility crash (1 if bus crushed into roadway facilities; 0 otherwise)	0.2946	−0.2341	−0.0605	0.6636	−0.2912	−0.3724	0.2948	−0.2342	−0.0606	0.6631	−0.2904	−0.3727
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	0.2511	−0.2162	−0.0349	0.4452	−0.2984	−0.1468	0.2502	−0.2154	−0.0348	0.4457	−0.2984	−0.1474
Crash location and roadway characteristics												
Pavement condition (baseline: normal)												
Slippery (1 if pavement was slippery; 0 otherwise)	0.0666	−0.0561	−0.0105	–	–	–	0.0683	−0.0576	−0.0108	–	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	0.1595	−0.1388	−0.0206	−0.1474	0.0777	0.0696	0.1578	−0.1372	−0.0205	−0.1473	0.0775	0.0698
Roadway segment of crash (baseline: tollgate/tunnel section)												
Mainline (1 if crash occurred on mainline section; 0 otherwise)	−0.0951	0.0797	0.0154	−0.1709	0.1085	0.0625	−0.0988	0.0828	0.0160	−0.1756	0.1113	0.0643
Ramp (1 if crash occurred at ramp section; 0 otherwise)	–	–	–	0.1007	−0.0648	−0.0359	–	–	–	0.0937	−0.0599	−0.0337
Shoulder barrier type (baseline: guardrail shoulder barrier type)												
No shoulder barrier (1 if crash occurred on roadway section with no shoulder barrier; 0 otherwise)	0.0477	−0.0397	−0.0080	–	–	–	0.0485	−0.0403	−0.0082	–	–	–
Vehicle characteristics												
Size of bus (baseline: large-sized - 45-seater)												
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	0.0811	−0.0669	−0.0142	0.0660	−0.0398	−0.0262	0.0831	−0.0686	−0.0146	0.0673	−0.0405	−0.0268
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	0.0592	−0.0494	−0.0097	–	–	–	0.0549	−0.0458	−0.0091	–	–	–
Traffic characteristics												
Logarithm of AADT*	−0.0673	0.0556	0.0117	−0.0678	0.0410	0.0268	−0.0680	0.0562	0.0118	−0.0665	0.0401	0.0264
Proportion of private cars*	0.3830	−0.3166	−0.0664	–	–	–	0.3785	−0.3128	−0.0657	–	–	–

Notes: "-" represents "not applicable"; Crash severity (1 = PDO, 2 = evident injury, 3 = fatal/severe injury); Crash size (1 = one vehicle involved crash, 2 = two vehicles involved in a crash, 3 = three or more vehicles involved in a crash); variables with (*) are continuous, while the others are dummy variables.

Table 11
Marginal effects for model estimates (2012-2016).

Variable	Independent model						Frank copula-based joint model					
	Crash severity			Crash size			Crash severity			Crash size		
	0	1	2	0	1	2	0	1	2	0	1	2
Season (baseline: winter)												
Spring (1 if crash occurred from March to May; 0 otherwise)	−0.0679	0.0517	0.0163	–	–	–	−0.0479	0.0356	0.0123	–	–	–
Summer (1 if crash occurred from June to August; 0 otherwise)	–	–	–	0.0527	−0.0255	−0.0272	–	–	–	0.0550	−0.0266	−0.0284
Time of crash (baseline: midnight to early morning −12 am to 5:59am)												
Noon to late afternoon (1 if crash occurred from 12 pm to 5:59pm; 0 otherwise)	–	–	–	−0.0318	0.0148	0.0170	–	–	–	−0.0335	0.0156	0.0180
Drivers characteristics												
Age (baseline: old driver (≥50))												
Young (1 if driver is young < 30; 0 otherwise)	0.0441	−0.0345	−0.0096	−0.1799	0.0769	0.1030	0.0245	−0.0183	−0.0063	−0.1810	0.0772	0.1038
Drivers fault (baseline: wrongful passing, reversing, and headway)												
Negligence (1 if driver was negligent; 0 otherwise)	−0.1077	0.0810	0.0267	−0.1907	0.0781	0.1127	−0.0815	0.0610	0.0206	−0.1875	0.0769	0.1107
Over speeding (1 if driver was over speeding; 0 otherwise)	–	–	–	0.1159	−0.0585	−0.0574	–	–	–	0.1205	−0.0609	−0.0596
Drivers condition (baseline: impaired (alcohol, drugs, tired))												
Normal (1 if driver was normal at the time of crash; 0 otherwise)	0.2003	−0.1388	−0.0615	–	–	–	0.1964	−0.1521	−0.0444	–	–	–
Weather characteristics												
Weather during crash (baseline: fine)												
Inclement weather (1 if weather was cloudy, foggy, rainy, snowy; 0 otherwise)	–	–	–	0.0427	−0.0203	−0.0224	–	–	–	0.0421	−0.0200	−0.0221
Crash characteristics												
Crash type (baseline: vehicle-vehicle crash)												
Vehicle-pedestrian (1 if bus crushed with pedestrian; 0 otherwise)	0.0991	−0.0790	−0.0201	0.1945	−0.1025	−0.0920	0.0922	−0.0701	−0.0221	0.2009	−0.1061	−0.0948
Crash location and roadway characteristics												
Pavement condition (baseline: normal)												
Slippery (1 if pavement was slippery; 0 otherwise)	0.0920	−0.0738	−0.0182	–	–	–	0.0807	−0.0615	−0.0192	–	–	–
Pothole/fallen objects (1 if pavement had potholes or fallen objects; 0 otherwise)	0.1543	−0.1275	−0.0268	−0.1096	0.0451	0.0644	0.1339	−0.1054	−0.0285	−0.1094	0.0450	0.0644
Roadway segment of crash (baseline: tollgate/tunnel section)												
Mainline (1 if crash occurred on mainline section; 0 otherwise)	−0.1598	0.1278	0.0319	−0.2175	0.1125	0.1050	−0.1309	0.1001	0.0308	−0.2203	0.1139	0.1064
Ramp (1 if crash occurred at ramp section; 0 otherwise)	–	–	–	0.1160	−0.0603	−0.0556	–	–	–	0.1120	−0.0580	−0.0540
Vertical alignment (baseline: zero-grade)												
Grade (1 if crash occurred on roadway with vertical curves; 0 otherwise)	–	–	–	0.0458	−0.0219	−0.0239	–	–	–	0.0471	−0.0225	−0.0246
Median barrier type (baseline: no median barrier)												
Fixed concrete median barrier (1 if crash occurred on roadway section with fixed concrete barrier of height 81/127 cm; 0 otherwise)	–	–	–	−0.0575	0.0274	0.0301	–	–	–	−0.0569	0.0271	0.0298
Shoulder barrier type (baseline: guardrail shoulder barrier type)												
Other shoulder barrier types (1 if crash occurred on roadway section with concrete/rock/pipe shoulder barrier type; 0 otherwise)	−0.0546	0.0417	0.0129	−0.0499	0.0227	0.0273	−0.0437	0.0325	0.0112	−0.0447	0.0203	0.0244
Vehicle characteristics												
Size of bus (baseline: large-sized - 45-seater)												
Small-sized (1 if bus was small-sized: 21-seater coach; 0 otherwise)	0.1422	−0.1098	−0.0324	0.1854	−0.0860	−0.0994	0.1151	−0.0866	−0.0285	0.1886	−0.0873	−0.1013
Medium-sized (1 if bus was small-sized: 28-seater coach; 0 otherwise)	0.0941	−0.0745	−0.0197	0.0834	−0.0409	−0.0425	0.0664	−0.0499	−0.0165	0.0814	−0.0398	−0.0416
Traffic characteristics												
Logarithm of AADT*	−0.0499	0.0387	0.0112	−0.0407	0.0192	0.0216	−0.0417	0.0310	0.0107	−0.0399	0.0187	0.0211
Proportion of private cars*	0.3285	−0.2551	−0.0734	0.2672	−0.1258	−0.1414	0.2610	−0.1939	−0.0671	0.2560	−0.1204	−0.1356
Proportion of trucks*	−1.5120	1.1740	0.3379	−1.8650	0.8782	0.9869	−1.0883	0.8086	0.2797	−1.8214	0.8564	0.9650

Notes: “-” represents “not applicable”; Crash severity (1 = PDO, 2 = evident injury, 3 = fatal/severe injury); Crash size (1 = one vehicle involved crash, 2 = two vehicles involved in a crash, 3 = three or more vehicles involved in a crash); variables with (*) are continuous, while the others are dummy variables.

the probability of observing both fatal/severe crashes and large-sized crashes is reduced in summer relative to crashes that occur in winter. A similar trend is observed in the 2012 – 2016 model regarding the crash size outcome variable. A potential rationale for the negative sign could be attributed to the fact that, unlike the winter season, drivers enjoy the benefit of clear weather during summer. As such, their visibility is increased, and they are able to detect and act appropriately during emergency situations. Crashes during the summer are more likely to result in PDO, with most of them being single-vehicle crashes (see marginal effect estimates in Tables 9 and 11). Our result conforms to the finding in the literature. For example, a recent study from the U.S. identified that the likelihood of observing more severe crashes decreases during summer (Feng et al., 2016), whereas another study from Iran found out that the mortality rate of crashes in summer was 35.8 %, 60.2 % in autumn, and 82.8 % in winter (Mehmandar et al., 2014).

6.4. Driver characteristics and faults

To show how driver characteristics and faults affect crash severity and crash size of express bus-involved crashes, we investigated the influence that drivers' age, condition, and faults have on both crash severity and crash size outcomes. Most importantly, it was identified that, compared to older drivers, younger bus drivers are likely to have reduced crash severities; however, they have a higher propensity of being involved in large-sized crashes. From the significant marginal effects results, it is seen that the probability of young drivers having PDO crashes increases by a higher margin compared to the other two severity outcomes, and the probability of young drivers having three or more vehicles involved in a bus crash increases by the highest margin for both 2012 – 2013 and 2012 – 2016 year groups. Generally, the probability of older bus drivers being involved in more severe crashes is higher compared to younger drivers, as reported by other studies (Kim et al., 2013; Russo et al., 2014; Wali et al., 2018; Yasmin et al., 2014). This result can be attributed to the strong physique and attributes of younger bus drivers relative to older bus drivers who are generally weaker and could be more injured in the event of a crash. Also, the finding that younger drivers are more likely to be involved in large-sized crashes is opposite to what has been identified in previous research that considered private passenger cars (Hong et al., 2019b). This result may be due to the physical differences in express-buses and private cars. Besides, due to youthful exuberance and immaturity of younger bus drivers, they are more likely to be engaged in risky-driving, and commit traffic violations such as over speeding and wrongful over-takings, which is likely to cause many vehicles involved in a crash as suggested by Yagil (1998).

With regard to bus drivers' condition, the study also demonstrated that the probability of observing higher crash severities is reduced in normal bus drivers compared to fatigued and alcohol/drug-impaired bus drivers. From the significant marginal effects results, the probability of having PDO crashes increases the highest for all year groups. This finding is in line with the literature (Wu et al., 2014; Zhou and Chin, 2019), and is instinctive as normal drivers are more likely to drive carefully and defensively given the height of safety consciousness campaigns in South Korea. When driving in a healthy state of mind, it is less likely to have impaired judgment that triggers high-risk driving (Das et al., 2012). The result underscores the importance of enforcing rules to ensure that bus drivers do not drive when fatigued or under the influence of drugs. With regard to crash size, we identified that this group of normal bus drivers is also more likely to be involved in large-sized crashes, particularly, crashes involving two (2) vehicles (marginal effects estimate in Table 10). This variable was only significant in the 2014 – 2016 data. Our result confirms previous studies that identified that normal drivers are less associated with a high risk of having single-vehicle crashes, and increased injury severities compared to their alcohol-impaired counterparts (Bham et al., 2012; Öström and Eriksson, 1993). Alcohol-impaired or fatigued bus drivers are less alert compared to normal bus drivers, increasing the chance of having severe vehicle-facility collisions or hitting the rear of other vehicles.

With regard to bus drivers' faults, the variable for negligent driving behavior was positively significant under almost all crash severity and crash size models for the three-time periods. It is worth noting that a driver drives negligently when he/she drives without reasonable care. It may also include distracted driving, driving through stop signs without stopping, or failing to use the turn signals. The positive coefficients suggest an increased chance of observing increased injury severities and larger crash sizes in express-bus involved crashes when the drivers are negligent. From the marginal effect estimates, we observe that the probability of sustaining evident injury increases higher when a driver is negligent compared to the chance of sustaining fatal/severe crashes, and the probability of having three or more vehicles involved in a single crash increases by the highest margin compared to the other two crash size levels when a bus driver is negligent **(see** Tables 9–11). Bus drivers with these behaviors can cause serious harm to other unsuspecting road users by easily running into them. Driver's traffic violations have been investigated by several studies. In the literature, negligent driving has been found to be a significant cause of fatal crashes (Chang and Yeh, 2006). Negligent driving, being a major cause of multi-vehicle crashes, is also in line with the study documented by Hong, Tamakloe, & Park (2019). To avert these, it is imperative to eliminate all distractions, avoid eating while driving, and staying alert, particularly when driving in adverse weather conditions.

Finally, the variable for over speeding had a negative coefficient, indicating a higher chance of single-vehicle crashes (Tables 6–8). This is also confirmed from the estimates of the marginal effects in Tables 9–11. A possible reason for this finding is that, intuitively, the risk of losing control of the vehicle and crashing into roadway facilities while over speeding on the expressway, especially at curved sections, ramps, and toll booth segments is high. Martensen and Dupont (2013) noted that roadways with high speed limits had increased propensity for small crash sizes. These results are consistent with findings in the literature (Chen et al., 2009; Hong et al., 2019b).

6.5. Weather characteristics

In this category, the variable for inclement weather was significant only in the 2012 – 2013 and 2012 – 2016 crash size models. Compared to bus-involved crashes that occurred during fine weather, a reduced propensity for large crash sizes was observed for crashes that occurred during the rainy, snowy, or cloudy weather. From the marginal effect estimates, we observe that the probability for observing single-vehicle crashes increases the most under these weather conditions. The reduced visibility and friction due to rain or snow make drivers drive more cautiously compared to driving on dry pavements during fine weather conditions and might explain the resultant lower chance of having large-sized crashes. Our result is in tandem with other studies that identified that driving in fine weather was associated with a high propensity for multi-vehicle crashes due to the high risk driving nature of most drivers when weather conditions are good (Hong et al., 2020a).

6.6. Crash characteristics

To demonstrate how different crash mechanisms affect both crash severity and size of bus-involved crashes, we investigated the effects of vehicle-facility and vehicle-pedestrian collision on crash severity and crash size outcomes. Compared to vehicle-vehicle crashes, results displayed in Tables 6–8 suggest a reduction in the probability of experiencing larger crash sizes in both vehicle-pedestrian crashes and vehicle-facility crashes. From the marginal effect estimates, crashes of these kinds are more likely to end up as single-vehicle crashes **(see** Tables 9–11**)**. On expressways that are close to human settlements, drivers sometimes encounter pedestrians. As unsuspecting drivers try to dodge and avoid hitting pedestrians, they may crash into roadway facilities. As such, most of these crash types would likely end up in single-vehicle crashes.

Compared to vehicle-vehicle collisions, both vehicle-facility and vehicle-pedestrian crashes tend to have lower crash severities. Essentially, such crashes are mostly PDO crashes **(see** Tables 9–11**)**. Generally, it is reasonable to expect vehicle-facility crashes to result in PDO crashes, especially those that occur around toll booth sections. South Korean expressways have many automatic toll gate sections. Upon spotting these facilities, drivers are required to slow down and carefully pass through the facility as they make payments. In the process of passing through the narrow toll gate facility, they may end up hitting near-by objects. In line with our results, Hong et al. (2020a, 2020b) identified that heavy vehicles are more likely to crash into facilities near toll gates in South Korea. The authors posit that such crashes are likely to have reduced crash sizes (single-vehicle crash) and result in PDO crashes. On the contrary, vehicles are more likely to speed and interact more with other streams of traffic on mainlines. As such, a crash in such sections is likely to be vehicle-vehicle crashes with higher severities and crash sizes.

However, our result concerning vehicle-pedestrian crashes was not in line with past research. In the literature, pedestrians have been identified to suffer fatal injury severities, especially when a bus is involved in a vehicle-pedestrian crash (Aziz et al., 2013). Our result may be partly due to the low proportion of vehicle-pedestrian crashes compared to the other types of crashes or due to the crash severity nomenclature used by the KEC. From the total dataset, 638 out of 2997 crashes were vehicle-pedestrian crashes. This is reasonable since pedestrians do not have enough access to expressways compared to other roadway types. Of the 638 crashes, only 25 were fatal/severe, 110 were evident injury crashes, and 503 were PDO crashes. Since the nomenclature for fatal/severe crashes is the only level that captures death, the results show that most of the vehicle-pedestrian crashes in South Korea do not result in fatal/severe crashes. The vehicle may have ended up hitting the pedestrian slightly or may have caused evident injuries. Many of such crashes end up as PDO or evident injury crashes. It is worth noting here that, in South Korea, the KEC's crash severity classification factors both the cost of damage and loss of lives in its nomenclature. As such, the overall crash severity resulting from a vehicle-pedestrian crash should be lower compared to the crash severities occurring as a result of vehicle-vehicle crashes on expressways if there is no death involved.

6.7. Crash location and roadway characteristics

In terms of crash location and roadway characteristics, the estimates in Tables 6–8 show that a wide variety of variables influence both crash severity and crash size. Concerning the crash location factors, we identified that crashes that occur on mainlines have a high propensity for higher severities and larger crash sizes relative to those that occur at toll gate/tunnel sections. It is worth noting that the variable for crashes at the tollgate and tunnel sections was combined due to the small number of observations. In Tables 9 and 11, the marginal effect estimates show that the probability of sustaining evident injury severity increases highest compared to the other two crash levels, and the probability of observing crashes involving three or more vehicles when the crash occurs on the mainline segment increases by the highest margin when the crash occurs on the mainline relative to when the crash occurs at the toll gate/tunnel sections. The result obtained is expected as there is increased interaction between vehicles on expressway mainlines. Also, In South Korea, there are strict regulations, such as lane-changing prohibitions when using tunnels. In addition, tunnels are well lit and have numerous facilities that make the driver aware of the change of environment. In contrast to the high-speed limits and reduced prohibitions on mainlines, it is reasonable to observe lower severity and smaller-sized crashes in tunnels. Also, at toll gates, vehicles have to slow down and make payments as they pass through and make payments. At these segments, vehicles mainly hit the facilities in the narrow paths, as against mainline segments where vehicle-vehicle collisions are highly probable. In line with our result, Hong et al. (2020a,2020b) also noted that crashes occurring at toll gate sections are likely to be PDO and single-vehicle crashes.

Our results also showed that, compared to crashes occurring at toll booth/tunnels, those crashes occurring at ramp sections have a high propensity for being low-sized crashes. The probability of observing a single-vehicle crash when a bus crashes at the ramp section increases by 0.1120 (2012 − 2016 data) and 0.0937 (2014 − 2016 data). Again, these crashes are likely to be vehicles hitting nearby facilities as they join the ramp sections while driving at high speeds. Studies in the literature found that crashes at these sections are likely to be single-vehicle crashes (Hong et al., 2020a, 2019b).

Regarding road conditions, this study's results depict that, relative to crashes that occur on roadways with normal pavement conditions, those that occur on slippery roads, or roads with potholes/fallen objects on the roadways consistently resulted in lower crash severities. However, when there are potholes/fallen objects on the roadway, there is an increased chance of observing large-sized crashes (Tables 6–8). The marginal effect estimates show that the chance of observing PDO crashes increases by the highest margin when the crash is a result of potholes/fallen objects. Also, when there are potholes/fallen objects on the roadway, the probability of observing a bus-involved crash involving three or more vehicles increases by the highest margin (**see** Table 11). A possible explanation for this observation is that, due to the high-speed limit on expressways, unsuspecting bus drivers may lose control of their vehicles out of panic as they try to dodge potholes/fallen objects on the roadway. They may end up hitting other vehicles, leading to PDO and multi-vehicle crashes or larger sizes. As pointed out in the literature, when the pavement becomes slippery due to factors such as ice or spilled chemicals, there is a high chance that the friction between the vehicle tires can be reduced. Hence, the braking mechanism of vehicles is affected – increasing the chance of road skids or hitting other vehicles/facilities (Duncan et al., 1998). Other studies found that highly severe crashes are expected when drivers drive on roadways with good pavement conditions. Drivers have a high tendency of being over relaxed and indulge in risky driving when on roads with good pavement conditions. This behavior has been identified to lead to more severe crashes (Hong et al., 2020a; Yu et al., 2019).

From Tables 6 and 8, we identified that crashes that occurred at roadway segments with shoulder barriers made of concrete, rock, or pipes were likely to have larger crash sizes compared to segments with shoulder guardrails. These results are confirmed by the increased probabilities shown in the significant marginal effect estimates in Tables 9 and 11. These results are in tandem with a recent study that identified that single-vehicle crashes are more likely to occur when a vehicle crashes into a guardrail (Hong et al., 2020b). Also, relative to roadway segments with shoulder guardrails, our study identified that, while the likelihood of observing a severe bus-involved crash is lower at roadway segments without shoulder barriers, those with concrete/rock/pipe shoulders have a higher chance of being severe. From the estimates of the marginal effects in Table 10, the probability of observing a PDO crash is highest at roadway segments without shoulder barriers, and the chance of observing evident injury crashes when a bus crashes into a concrete or rock shoulder barrier is the highest among the three crash severity levels. This result is intuitive and is in line with previous studies. Even though guardrail systems are made and installed to reduce the severity of crashes, several studies have identified that the chance of observing fatal and severe crashes increases when vehicles hit them (Dissanayake and Roy, 2014; Li et al., 2018). As noted by Padmanaban et al. (2010), hitting rigid roadway facilities at high speeds could damage the vehicle and injure its occupants. Therefore, it is reasonable to have PDO crashes at roadway sections without guardrails.

Further, we observed that compared to roadway segments without median barriers, those with fixed concrete median barriers are likely to be associated with large crash sizes. In particular, hitting fixed concrete wall median barriers of heights 81/127 cm was likely to result in crashes of large-sizes. From the marginal effects estimates in Tables 9 and 11, the probability of observing three or more vehicle-involved crash outcomes increases highest when a vehicle crashes into fixed median barriers. This type of barrier had been introduced with the main purpose of preventing cross-median crashes on divided expressways; however, interestingly, studies have documented that median-related crash counts have increased after its installation on roadway sections (Hu and Donnell, 2010). Crashes of these kinds are mostly caused by negligence, as established in the literature (Hong et al., 2020a). This is also true as drivers become more relaxed and comfortable as they drive on median-divided expressways with straight and good pavement conditions. This increases the tendency to be negligent and indulge in risky driving. Once the leading vehicle crashes with the median barrier, other unsuspecting drivers may run into it, leading to multi-vehicle crashes. Moreover, as shown in this study and backed by literature, negligent driving is likely to be associated with crashes involving multiple vehicles.

It is also interesting to know that, express-bus crashes occurring on roadways with grade/slopes are more likely to result in small crash sizes compared to crashes that occur on roadways with zero grades (slope-less roadways). The finding is true as drivers are less careful and are tempted into indulging in risky behaviors when driving on straight and flat mainline sections. Due to this, they are likely to over speed and cause multi-vehicle crashes. Alternatively, it is more likely to drive carefully on sloppy roads. Crashes occurring in such circumstances are likely to be mere crashes involving only one vehicle. This variable was only statistically significant in the 2012 − 2016 model (Table 8), and the marginal effects model showed that the crashes are likely to be single-vehicle crashes if the road has positive/negative grades (Table 11). The results highlight the importance of providing adequate road safety signs at recommended distances before drivers approach sloppy roadway segments. Security cameras installed at straight roadway segments would also be needed to check unscrupulous driving on flat mainline sections.

6.8. Vehicle characteristics

Relative to large-sized buses, small-sized and medium-sized buses are likely to have reduced severities and crash sizes in the event of a crash; however, that of medium-sized buses was slightly lower compared to small-sized buses. From the marginal effect estimates, the probability of observing a PDO crash when a small bus-involved crash occurs increases by 0.1151, whereas that of a medium-sized bus increases by 0.0664 (see Table 11). The result depicts that the severity of express-bus crashes increases with the increasing size of bus. The crash size component was significant only in the data for 2012 − 2016/ 2012 − 2013 (both small and medium-sized bus) and 2014 − 2016 (small-sized bus only). In general, similar to the effect on crash severity, the trend shows that crash size also increases with the increasing size of bus. While the probability of observing a single-vehicle involved crash increases by 0.1886 for a small-sized bus, that of a medium-sized bus increases by 0.0814 (see Table 11). These findings are intuitive and consistent with the literature (Prato and Kaplan, 2014; Zhou et al., 2020). First, as noted in the literature (Zheng et al., 2018), as the size of the vehicles increases, its gross weight also increases – causing a decrease in stability and maneuverability in times of emergency. Compared to other vehicles, the size and gross weight of express-buses cause them to have larger potential energy, which renders them difficult to handle in case of emergencies, especially on expressways where the speed limit is higher compared to other roadways. The high potential energy increases their stopping distance and renders them highly likely to crash into many vehicles in case of a sudden brake failure or the driver commits a violation (Wali et al., 2018). In the event of a crash or rollover, many people would be exposed to injuries compared to small and medium-sized buses.

6.9. Traffic characteristics

From Table 8, three exposure variables, namely traffic volume (log of AADT), truck proportion, and private car proportion, were found to be significantly associated with express-bus involved crash severity and crash size. Under the crash severity and crash size categories, traffic volume was positively significant in both the 2014 − 2016 and 2012 − 2016 data, whereas the variable for the truck proportion was also positively significant in both the 2012 − 2013 and 2012 − 2016 data. The positive coefficients observed denote that an increase in AADT and the proportion of trucks on the expressways results in an increase in both crash severity and crash size of bus-involved crashes. Besides, the estimates of the marginal effects show that, on average, the probability of sustaining evident injuries increases by the highest magnitude, followed by fatal/severe crashes in case the AADT on the expressway increases by one unit (see Tables 10 and 11). The computed marginal effect estimates also show that the proportion of trucks on the expressways have the highest influence on the probability of increased crash severity and crash size (**see** Tables 9 and 11). On the other hand, the study identified an opposite trend regarding the proportion of private cars. The chance for observing PDO crashes and single-vehicle crashes is highest when the proportion of private cars increases.

The explanation of the result from the AADT variable is as follows. It is noteworthy that, due to the relatively small land size of South Korea, many people travel between cities by express buses, and the proportion of express buses that use the expressways is high. As such, the likelihood of having many casualties in express-bus involved crashes is expected in the event of a crash. Another plausible explanation for this result is that, on the expressways where speed limits are high, an increase in traffic volume increases the interaction of buses with other vehicles on the roadway, which results in higher crash probability and severity. In particular, Gårder (2006) identified that high AADT is positively associated with severe crashes, especially on roadways with broad shoulders. Also, consistent with the literature (Yu and Abdel-Aty, 2013), the increased exposure to mixed traffic flow increases the chance of hitting or running into many other vehicles in times of crashes, leading to larger crash sizes. Finally, unlike private cars, trucks are mostly associated with longer stopping distances and high speeds. Also, they have dangerous blind spots on their sides. As such, a high proportion of trucks is likely to induce crashes involving many vehicles with increased severity levels (Chang and Mannering, 1999; Prato and Kaplan, 2014).

7. Discussion and policy recommendations

The ultimate goal of this study was to analyze and to give insights into the potential effects of factors influencing the severity and size of bus-involved crashes on expressways using a more robust statistical approach that accounts for the dependence between both crash outcome variables. In addition, we examined the overall temporal stability of the model estimates across different time periods. With our goal in mind, we retrieved bus-involved crashes from the database of all crashes that occurred in South Korea from 2012 to 2016 and used it for the analysis. The main findings from Section 6 are further discussed in this section, and policy recommendations would be suggested thereafter for practice.

The findings of the overall temporal instability of the bus-involved crash models showed that the effect of the factors in the models varied from time to time. Even though the marginal effect of variables in the crash severity and crash size models mostly followed the same trend, most indicator variables were almost twice the value when compared to other year's results. For example, in comparing the estimates of the marginal effects for 2012 − 2013 and 2014 − 2016 copula-based model results, the indicator for drivers age less than 30, small-sized bus, a roadway with potholes/fallen objects, and vehicle-facility collision had almost doubled effects in the 2014 – 2016 crash severity model. More specifically, the probabilities for PDO crashes were almost doubled in 2014 – 2016 compared to that of 2012 – 2013. A similar trend was observed for crash sizes regarding vehicle-pedestrian crashes. The probability of having single-vehicle crashes in 2014 – 2016 (0.4457) more than doubled compared to that of 2012 – 2013 (0.1630), and the probability of observing larger crash sizes for indicators such as negligent driver, and drivers' age less than 30 decreased significantly in 2014 – 2016 compared to 2012 – 2013. The reduction in crash severities and sizes may be as a result of the increased road safety awareness program introduced by GOK within that period. Similar studies also found major differences in marginal effects when different year group data were studied and compared (Behnood and Mannering, 2019; Islam and Mannering, 2020).

Estimating the dependence between injury severity and crash size of bus-involved crashes on expressways showed interesting results. First, it was shown that both outcome variables have a positive relationship. Generally, in most variables, the results show that, when there is an increase in crash severity, there is a resultant increase in crash size (for variables such as log of AADT, truck proportion, negligent driving, crashes on mainline segments, and hitting concrete shoulder barriers). To further reduce the severity of crashes on expressways, it is necessary to continue with the intensive driver sensitization and to put in more effort in reducing the size of crashes by checking negligent driver behavior, especially on wide mainline segments of expressways. For the safety of travelers using buses, it would be necessary to check to find appropriate times for when trucks can use the expressways in order to reduce their interaction with the normal traffic stream.

A critical look at the driver's characteristics and faults shows that the leading cause of increased severity and crash size is negligent driving. This variable had been consistently positively significant in all three-time periods studied in this research. As for crash size, it has been shown that relative to older drivers, younger drivers are more likely to have multiple vehicle crashes. These findings aligned with existing knowledge that stressed that aberrant driving had the tendency of causing crashes with high severities (Hong et al., 2020a; Wang and Prato, 2019; Zhu and Srinivasan, 2011). While some studies link the increase of these violations to stress (Lagarde et al., 2004), others blamed the risk-taking trait of large vehicle drivers as having the most substantial impact on traffic violations (Rowden et al., 2011). According to Hong et al. (2019a), drivers are likely to violate traffic rules when they drive under the influence of drugs/alcohol, when fatigued, or when they are ill. Their study also demonstrated that large vehicle driver infringements such as improper passing and safe distance violations were likely to increase the probability of crashes. Crashes resulting from these violations are likely to result in multi-vehicle crashes. This finding is alarming in that if it is not checked, the number of severe crashes would continue to increase. As a countermeasure, drivers, especially those below the age of 30, should be educated on the need to stay attentive while driving and follow the driving regulations. Lawmakers have to enact laws to check negligent driving, publicize them, and ensure that they are enforced. There should be cameras installed on the highway to detect over speeding and road traffic rule violations, such as using mobile phones while driving. Culpable drivers should be heavily fined as low fines may not convince people to desist from that habit.

From the roadway perspective, it was not surprising to discover that crash severity increases when the crash occurs on mainline segments. This finding is plausible since some drivers tend to indulge in risky driving when on these segments (Hong et al., 2020a; Yu et al., 2019). This trend can be reduced by putting in measures to check negligent and risky driving behavior on such segments, as noted in the previous paragraph. The most alarming observation has to do with the increase in crash sizes when vehicles hit concrete/fixed median and shoulder barriers. We also observed an increase in severity when vehicles hit concrete shoulder barriers. In particular, the area around the shoulder is supposed to provide an additional margin of safety and not to increase crash severities; however, if hitting the shoulder barrier increases the severity of crashes in such segments, then it would be necessary to identify those hotspots and reconsider the type of barriers suitable for such sections in order to reduce the impact when vehicles hit them.

Regarding traffic characteristics, both traffic volume and truck proportion had a positive effect on both the severity and size of bus-involved crashes. In particular, the variable for truck proportion was identified to have the highest impact on both outcome variables in this study. The reason for this trend is the increased interaction between trucks and the general traffic during the day when many trips using express buses are common. As mentioned in the manuscript, this issue can be checked by introducing truck prohibitions on certain lanes of the expressways at certain times of the day. The provision of enough safety signs posted on the roadway, and driver education concerning how to drive safely when around heavy trucks, will reduce the increased severity levels of crashes. An investigation of this nature was conducted in a study in Florida and Houston. Trucks were prohibited from using the left lane within certain times of the day. Even though full compliance was not achieved, some positive gains in terms of accident rates were achieved (Harwood et al., 2003). In order to ensure full compliance, the knowledge of artificial intelligence could be harnessed by installing sensors and cameras on restricted lanes to detect and fine defaulting drivers. Other measures that can be employed to check the interaction between express buses and trucks for the long term include the provision of dedicated routes for trucks and the provision of clearly visible advance warning signs for all drivers.

A look at the temporal features shows that bus-involved crash severity increases during the morning period from 6 a.m. to 11:59 am. As identified by Hao et al. (2016), this observation may be due to the increased speeds or risky driving behavior as drivers are in a hurry to reach their destinations on time. Again, there is the need for speed control, especially in the morning. This can be achieved by installing speed cameras on all segments of expressways (considering shorter distances between each speed camera) as a crash countermeasure. If culprits are severely punished, it will serve as a deterrent to other drivers and will reduce road crashes during this period in the long run.

The seasonal factor also has an impact on the severity of crashes. Among the seasons studied in this research, we identified that crashes occurring during the spring season led to a significant increase in crash severity. Compared to the winter season, many more crashes occur during the spring season due to many reasons. During this period, there are a lot of rain showers, which reduces visibility on the roadway. Also, due to the harsh winter conditions, roadways may develop potholes. Moreover, there may be animals crossing due to the shift to sunny weather. These hazards pose a threat to the driver in case they are not spotted on time. As a countermeasure, drivers need to be educated through the media to focus on driving by removing distractions as much as possible. This can help them spot hazards and react in time. Drivers can be encouraged to report potholes on the roadways to local government agencies to get them fixed. Finally, authorities have to make sure roadways are checked frequently for repairs to reduce the severity of crashes.

Based on the results of the study, specific policy recommendations are suggested to improve express bus transport safety:

- The interventions championed by the GoK through the implementation of the 7th National Transport Safety Master Plan (2012 – 2016) led to considerable reductions in severe crashes when the results from the 2012 – 2013 and 2014 – 2016 data are compared. However, there is the need to further heighten interventions and educational programs targeted at younger drivers on the need to drive on expressways cautiously.

- The evidence of vehicle-pedestrian crashes on the expressway is indicative of the fact that there are human activities near some of the expressways. Providing pedestrian barriers and pedestrian accommodations (grade-separated crossing) at vantage points and making regulations to protect pedestrians would be necessary.
- The finding that exposure parameters, particularly the proportion of trucks, have the highest probability of causing increased crash severity and crash size outcomes signifies increased interaction between express buses and other trucks. This can be mitigated by enacting and enforcing regulations that control the operation time and choice of lanes for trucks. Also, educating both truck and express bus drivers to be fully aware of blind spots and how to accommodate other vehicles on the roadway is crucial.
- Advanced technology can be employed through the installation of roadway cameras on straight mainline segments to check roadway violations. Regulations should be put in place to enforce regular servicing of vehicles, especially right after winter, due to the adverse effects the winter season can have on vehicles. Drivers found culpable of speed violations, and other negligent violations and faults should be severely punished to serve as a deterrent to others.
- There is a need to install adequate warning signs and to ensure the visibility of road markings.
- Roadway reconfiguration is needed at ramps, tollbooth sections, and sharply curved segments to accommodate for the wide turning swing and maneuverings of long express buses. This countermeasure can reduce the increased propensity of single-vehicle PDO crashes.
- More research is needed to determine appropriate roadway median and shoulder barrier types for shoulder/median crash hotspot segments
- Frequent roadway repairs of potholes would be necessary, especially right after the winter season.

8. Summary and conclusions

The study focused on jointly modeling and exploring the factors that influence crash severity and crash size outcomes of express bus-involved crashes. With three possible outcomes, each for both crash severity and crash size (number of vehicles involved in the crash), a wide range of factors that could have an impact on express bus-involved crashes were considered for the analysis. While studies have deeply examined factors associated with crash severity and crash size outcomes of private cars and trucks, not much has been done with regards to buses, especially those that ply the expressways. In general, existing studies have predominantly utilized separate models to investigate crash severity and the number of vehicles involved in crashes. Given the importance of these outcome variables in transport safety, it is necessary to study them in detail using robust methodological frameworks to derive more appropriate model estimates.

This study contributes to the current body of literature on express-bus safety in diverse ways. First, owing to the potential correlation between both outcome variables, a copula-based regression built upon the ORP model is implemented. This model considers the dependence between both outcome variables while accommodating for the endogeneity due to common unobserved and omitted factors, measurement errors, or from simultaneous causality. Also, the temporal stability of factors influencing the outcome variables was evaluated to help guide in making appropriate countermeasures. Note that this is very necessary as independent variables may have different effects on outcome variables in different year periods. Failing to test for temporal stability of factors and proceeding to run the model may yield inaccurate countermeasures. To the best of our knowledge, this is the first study to apply a robust approach to establish the dependence between both crash outcomes and to derive appropriate countermeasures that can be adopted to help minimize the number and severity of bus-involved crashes on expressways.

The primary objectives of this study were:

- To investigate the factors affecting crash severity and crash size of express bus-involved crashes that occurred on expressways while addressing endogeneity due to correlations between both outcome variables.
- To investigate the dependency between bus-involved crash severity and crash size.
- To analyze the temporal stability of factors impacting bus-involved crash severity and crash size.

To achieve the aim of the study, the empirical analysis used to achieve the objectives of this study involved the estimation of models using six different copula structures. By comparing copula structures, the goodness-of-fit results show that Frank copula was more appropriate for modeling the express bus-involved crash data. Also, comparing both the copula-based model and the independent ordered model reinforced the significance of accommodating for the dependence between the crash severity and crash size outcomes of express bus-involved crashes. From the results, we posit that accounting for the correlations between both outcome variables leads to more appropriate model estimation results. The estimation results identify several crucial exogenous factors contributing to express bus-involved crash outcomes.

The key finding drawn from this study had to do with the fact that a significantly strong dependency between crash severity and crash size of express bus-involved crashes exists. We demonstrated that the factors affecting both outcome variables were temporarily unstable; hence, conducting the investigation by splitting the data was warranted. We identified that, for express buses, an increase in AADT and the proportion of trucks on the road would yield a high increase in both crash severity and crash sizes. Also, crash severity and size will increase considerably when an express bus crushes into concrete shoulder barriers or median barriers, when the crash occurs on mainline segments, or when it occurs on roadway segments with potholes. Negligent driving is also a key contributing factor to increased crash severity and crash size.

The results demonstrate that, even though government interventions have helped to reduce crash severities and sizes to a large extent, more is needed to further reduce the level of carnage on the expressways. It is anticipated that the study reported in this paper sheds light on the selection of appropriate methodologies for traffic crash analysis. Also, the associated effects of crash-risk factors on the two outcome variables and the countermeasures identified in this study can be utilized by traffic engineers to improve express bus safety.

A major limitation worth mentioning is the absence of vehicle-level crash severity information since the KEC's approach to determining the crash severity aggregates the severity of all vehicles involved in a single crash. Due to this, it was not possible to model the severity of the occupants of only bus crashes. As the data employed for this research was based on KEC reported crashes, the likelihood of biased or under-reported observations is high. In some observations, it may be challenging to determine the exact cause of the crash, the driver at fault, or the vehicle's exact speed at the time of the crash. These data limitations can affect the validity of the findings. Also, the copula-based framework used in this study assumed a fixed copula parameter. The results obtained in our study also showed a superior fit compared to the independent model; however, the implication of assuming a fixed parameter could slightly affect model efficiency gains as the model assumes that the dependencies between the two response variables are constant. The fixed parameter copula model helped us achieve the primary objective of this study, which involves jointly investigating bus-involved crash severity and crash size and characterizing the dependence between them. This is novel to the bus-involved crash safety study area. This article compares and matches with others who applied the fixed-parameter copula methods and showed that it is also valid and provided more statistically valid results (Bhat and Eluru, 2009; Martey and Attoh-Okine, 2018a, 2018b; Rana et al., 2010; Sener and Bhat, 2011; Spissu et al., 2009; Wali et al., 2018). In future work, we would parameterize the copula function and compare the results to those obtained using the current methodology. Finally, our study succeeded in demonstrating temporal instability in factors affecting the severity and size of bus-involved crashes on expressways. In the future, it would be interesting to identify which subsets of crash populations that show temporal stability/instability. It would also be worthwhile to employ machine learning algorithms to identify important rules that show a set of factors leading to bus-involved crashes, especially at mainline sections where crashes are usually severe.

CRediT authorship contribution statement

Reuben Tamakloe: Conceptualization, Methodology, Software, Formal analysis, Validation, Writing - original draft. **Jungyeol Hong:** Formal analysis, Writing - review & editing, Visualization, Validation, Data curation. **Dongjoo Park:** Supervision, Investigation.

Declaration of Competing Interest

The authors report no declarations of interest.

Acknowledgement

This work was supported by the 2019 sabbatical year research grant of theUniversity of Seoul, South Korea.

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.aap.2020.105736.

References

Adler, M.W., Ahrend, R., 2017. Traffic Safety in South Korea: Understanding the Vulnerability of Elderly Pedestrians. Paris. https://doi.org/10.1787/22fca93d-en.
Aidoo, E.N., Ackaah, W., 2019. A generalized ordered logit analysis of risk factors associated with driver injury severity. J. Public Health (Bangkok). https://doi.org/10.1007/s10389-019-01135-8.
Al-Bdairi, N.S.S., Behnood, A., Hernandez, S., 2020. Temporal stability of driver injury severities in animal-vehicle collisions: a random parameters with heterogeneity in means (and variances) approach. Anal. Methods Accid. Res. 26, 100120. https://doi.org/10.1016/j.amar.2020.100120.
Ayuso, M., Bermúdez, L., Santolino, M., 2016. Copula-based regression modeling of bivariate severity of temporary disability and permanent motor injuries. Accid. Anal. Prev. 89, 142–150. https://doi.org/10.1016/j.aap.2016.01.008.
Aziz, H.M.A., Ukkusuri, S.V., Hasan, S., 2013. Exploring the determinants of pedestrian–vehicle crash severity in New York City. Accid. Anal. Prev. 50, 1298–1309. https://doi.org/10.1016/j.aap.2012.09.034.
Barua, U., Tay, R., 2010. Severity of urban transit bus crashes in Bangladesh. J. Adv. Transp. 44 (1), 34–41. https://doi.org/10.1002/atr.104.
Bedford, T., Cooke, R.M., 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. Ann. Math. Artif. Intell. 32, 245–268. https://doi.org/10.1023/A:1016725902970.
Behnood, A., Mannering, F.L., 2015. The temporal stability of factors affecting driver-injury severities in single-vehicle crashes: some empirical evidence. Anal. Methods Accid. Res. 8, 7–32. https://doi.org/10.1016/j.amar.2015.08.001.
Behnood, A., Mannering, F.L., 2016. An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. Anal. Methods Accid. Res. 12, 1–17. https://doi.org/10.1016/j.amar.2016.07.002.
Behnood, A., Mannering, F., 2019. Time-of-day variations and temporal instability of factors affecting injury severities in large-truck crashes. Anal. Methods Accid. Res. 23, 100102. https://doi.org/10.1016/j.amar.2019.100102.
Bham, G.H., Javvadi, B.S., Manepalli, U.R.R., 2012. Multinomial logistic regression model for single-vehicle and multivehicle collisions on urban U.S. Highways in Arkansas. J. Transp. Eng. 138 (6), 786–797. https://doi.org/10.1061/(ASCE)TE.1943-5436.0000370.
Bhat, C.R., Eluru, N., 2009. A copula-based approach to accommodate residential self-selection effects in travel behavior modeling. Transp. Res. Part B Methodol. 43 (7), 749–765. https://doi.org/10.1016/j.trb.2009.02.001.
Bhat, C.R., Sener, I.N., 2009. A copula-based closed-form binary logit choice model for accommodating spatial correlation across observational units. J. Geogr. Syst. 11 (3), 243–272. https://doi.org/10.1007/s10109-009-0077-9.
Bhowmik, T., Rahman, M., Yasmin, S., Eluru, N., 2019. Alternative model structures for multivariate crash frequency analysis: comparing simulation-based multivariate model with copula-based multivariate model, in: transportation research board 98th annual meeting. Washington DC, United States.
Broyles, R.W., Narine, L., Clarke, S.R., Baker, D.R., 2003. Factors associated with the likelihood of injury resulting from collisions between four-wheel drive vehicles and passenger cars. Accid. Anal. Prev. 35 (5), 677–681. https://doi.org/10.1016/S0001-4575(02)00046-5.
Chang, L.-Y., Chien, J.-T., 2013. Analysis of driver injury severity in truck-involved accidents using a non-parametric classification tree model. Saf. Sci. 51 (1), 17–22. https://doi.org/10.1016/j.ssci.2012.06.017.
Chang, L.-Y., Mannering, F., 1999. Analysis of injury severity and vehicle occupancy in truck- and non-truck-involved accidents. Accid. Anal. Prev. 31 (5), 579–592. https://doi.org/10.1016/S0001-4575(99)00014-7.
Chang, H.-L., Yeh, T.-H., 2006. Risk factors to driver fatalities in single-vehicle crashes: comparisons between non-motorcycle drivers and motorcyclists. J. Transp. Eng. 132 (3), 227–236. https://doi.org/10.1061/(ASCE)0733-947X(2006)132:3(227).
Chen, W.-H., Jovanis, P.P., 2000. Method for identifying factors contributing to driver-injury severity in traffic crashes. Transp. Res. Rec. J. Transp. Res. Board 1717 (1), 1–9. https://doi.org/10.3141/1717-01.
Chen, H.Y., Ivers, R.Q., Martiniuk, A.L.C., Boufous, S., Senserrick, T., Woodward, M., Stevenson, M., Williamson, A., Norton, R., 2009. Risk and type of crash among young drivers by rurality of residence: findings from the DRIVE Study. Accid. Anal. Prev. 41 (4), 676–682. https://doi.org/10.1016/j.aap.2009.03.005.
Christoforou, Z., Cohen, S., Karlaftis, M.G., 2010. Vehicle occupant injury severity on highways: an empirical investigation. Accid. Anal. Prev. 42 (6), 1606–1620. https://doi.org/10.1016/j.aap.2010.03.019.
Das, A., Gjerde, H., Gopalan, S.S., Normann, P.T., 2012. Alcohol, drugs, and road traffic crashes in India: a systematic review. Traffic Inj. Prev. 13 (6), 544–553. https://doi.org/10.1080/15389588.2012.663518.
Dissanayake, S., Roy, U., 2014. Crash severity analysis of single vehicle run-off-Road crashes. J. Transp. Technol. 04 (01), 1–10. https://doi.org/10.4236/jtts.2014.41001.
Duncan, C.S., Khattak, A.J., Council, F.M., 1998. Applying the ordered probit model to injury severity in truck-passenger Car rear-End collisions. Transp. Res. Rec. J. Transp. Res. Board 1635 (1), 63–71. https://doi.org/10.3141/1635-09.
Eluru, N., Paleti, R., Pendyala, R.M., Bhat, C.R., 2010. Modeling multiple vehicle occupant injury severity: a copula-based multivariate approach. Transp. Res. Rec. 2165, 1–11.
Feng, S., Li, Z., Ci, Y., Zhang, G., 2016. Risk factors affecting fatal bus accident severity: their impact on different types of bus drivers. Accid. Anal. Prev. 86, 29–39. https://doi.org/10.1016/j.aap.2015.09.025.
Gårder, P., 2006. Segment characteristics and severity of head-on crashes on two-lane rural highways in Maine. Accid. Anal. Prev. 38 (4), 652–661. https://doi.org/10.1016/j.aap.2005.12.009.
Goh, K., Currie, G., Sarvi, M., Logan, D., 2014. Factors affecting the probability of bus drivers being at-fault in bus-involved accidents. Accid. Anal. Prev. 66, 20–26. https://doi.org/10.1016/j.aap.2013.12.022.
Hao, W., Kamga, C., Wan, D., 2016. The effect of time of day on driver's injury severity at highway-rail grade crossings in the United States. J. Traffic Transp. Eng. (English Ed. 3 (1), 37–50. https://doi.org/10.1016/j.jtte.2015.10.006.
Harwood, D.W., Potts, J.B., Torbic, D.J., Glauz, W.D., 2003. Highway/Heavy Vehicle Interaction, Commercial Truck and Bus Safety Synthesis Program No. 3 – TRB. Transportation Research Board, Washington DC.
Hasler, B.P., Forbes, E.E., Franzen, P.L., 2014. Time-of-day differences and short-term stability of the neural response to monetary reward: a pilot study. Psychiatry Res. Neuroimaging 224 (1), 22–27. https://doi.org/10.1016/j.pscychresns.2014.07.005.
Hernández-Alava, M., Pudney, S., 2016. Bicop: a command for fitting bivariate ordinal regressions with residual dependence characterized by a copula function and normal mixture marginals. Stata J. Promot. Commun. Stat. Stata 16 (1), 159–184. https://doi.org/10.1177/1536867X1601600114.
Hong, J., Park, J., Lee, G., Park, D., 2019a. Endogenous commercial driver's traffic violations and freight truck-involved crashes on mainlines of expressway. Accid. Anal. Prev. 131 (August 2018), 327–335. https://doi.org/10.1016/j.aap.2019.07.026.
Hong, J., Tamakloe, R., Park, D., 2019b. A comprehensive analysis of multi-vehicle crashes on expressways: a double hurdle approach. Sustainability 11 (10), 2782. https://doi.org/10.3390/su11102782.
Hong, J., Tamakloe, R., Park, D., Choi, Y., 2019c. Estimating incident duration considering the unobserved heterogeneity of risk factors for trucks transporting HAZMAT on expressways. Transp. Res. Rec. https://doi.org/10.1177/0361198119827925.
Hong, J., Tamakloe, R., Park, D., 2020a. Discovering insightful rules among truck crash characteristics using apriori algorithm. J. Adv. Transp. 2020, 1–16. https://doi.org/10.1155/2020/4323816.
Hong, J., Tamakloe, R., Park, D., 2020b. Application of association rules mining algorithm for hazardous materials transportation crashes on expressway. Accid. Anal. Prev. 142, 105497. https://doi.org/10.1016/j.aap.2020.105497.
Hu, W., Donnell, E.T., 2010. Median barrier crash severity: some new insights. Accid. Anal. Prev. 42 (6), 1697–1704. https://doi.org/10.1016/j.aap.2010.04.009.
Islam, M., Mannering, F., 2020. A temporal analysis of driver-injury severities in crashes involving aggressive and non-aggressive driving. Anal. Methods Accid. Res. 27, 100128. https://doi.org/10.1016/j.amar.2020.100128.
Kim, Sangyoup, Choi, J., Kim, M., Kim, Sunggyu, 2012. Determination of accident modification factors for the median bus lanes on urban arterials. Int. J. Urban Sci. 16 (1), 99–113. https://doi.org/10.1080/12265934.2012.662588.
Kim, J.-K., Ulfarsson, G.F., Kim, S., Shankar, V.N., 2013. Driver-injury severity in single-vehicle crashes in California: a mixed logit analysis of heterogeneity due to age and gender. Accid. Anal. Prev. 50, 1073–1081. https://doi.org/10.1016/j.aap.2012.08.011.
Kockelman, K.M., Kweon, Y.-J., 2002. Driver injury severity: an application of ordered probit models. Accid. Anal. Prev. 34 (3), 313–321. https://doi.org/10.1016/S0001-

4575(01)00028-8.
Lagarde, E., Chastang, J.-F., Gueguen, A., Coeuret-Pellicer, M., Chiron, M., Lafont, S., 2004. Emotional stress and traffic accidents. Epidemiology 15 (6), 762–766. https://doi.org/10.1097/01.ede.0000142140.77194.ad.
Laman, H., Yasmin, S., Eluru, N., 2018. Joint modeling of traffic incident duration components (Reporting, response, and clearance time): a copula-based approach. Transp. Res. Rec. J. Transp. Res. Board 2672 (30), 76–89. https://doi.org/10.1177/0361198118801355.
Lee, J.-Y., Chung, J.-H., Son, B., 2008. Analysis of traffic accident size for Korean highway using structural equation models. Accid. Anal. Prev. 40 (6), 1955–1963. https://doi.org/10.1016/j.aap.2008.08.006.
Leone, M.J., Fernandez Slezak, D., Golombek, D., Sigman, M., 2017. Time to decide: diurnal variations on the speed and quality of human decisions. Cognition 158, 44–55. https://doi.org/10.1016/j.cognition.2016.10.007.
Li, N., Park, B.B., Lambert, J.H., 2018. Effect of guardrail on reducing fatal and severe injuries on freeways: real-world crash data analysis and performance assessment. J. Transp. Saf. Secur. 10 (5), 455–470. https://doi.org/10.1080/19439962.2017.1297970.
Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Anal. Methods Accid. Res. 17, 1–13. https://doi.org/10.1016/j.amar.2017.10.002.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16. https://doi.org/10.1016/j.amar.2016.04.001.
Martensen, H., Dupont, E., 2013. Comparing single vehicle and multivehicle fatal road crashes: a joint analysis of road conditions, time variables and driver characteristics. Accid. Anal. Prev. 60, 466–471. https://doi.org/10.1016/j.aap.2013.03.005.
Martey, E.N., Attoh-Okine, N., 2018a. Bivariate severity analysis of train derailments using copula-based regression models. ASCE-ASME J. Risk Uncertain. Eng. Syst. Part A Civ. Eng. 4 (4), 04018034. https://doi.org/10.1061/AJRUA6.0000982.
Martey, E.N., Attoh-Okine, N., 2018b. Modeling tamping recovery of track geometry using the copula-based approach. Arch. Proc. Inst. Mech. Eng. Part F J. Rail Rapid Transit 1989-1996 232 (8), 2079–2096. https://doi.org/10.1177/0954409718757556.
Martey, E.N., Attoh-Okine, N., 2019. Analysis of train derailment severity using vine copula quantile regression modeling. Transp. Res. Part C Emerg. Technol. 105, 485–503. https://doi.org/10.1016/j.trc.2019.06.015.
Mehmandar, M., Soori, H., Amiri, M., Norouzirad, R., Khabzkhoob, M., 2014. Risk factors for fatal and nonfatal road crashes in Iran. Iran. Red Crescent Med. J. 16, 8. https://doi.org/10.5812/ircmj.10016.
Melchiori, M.R., 2003. Which Archimedean Copula is the Right One? SSRN Electron. J. 1–20. https://doi.org/10.2139/ssrn.1123135.
Ministry of Land Infrastructure and Transport, 2020. Enforcement Rules of the Passenger Car Act. [WWW Document]. URL http://www.law.go.kr/lsSc.do?menuId=0&p1=&subMenu=1&nwYn=1&section=&tabNo=&query (accessed 4.5.20). .
Nashad, T., Yasmin, S., Eluru, N., Lee, J., Abdel-Aty, M.A., 2016. Joint modeling of pedestrian and bicycle crashes: copula-based approach. Transp. Res. Rec. J. Transp. Res. Board 2601 (1), 119–127. https://doi.org/10.3141/2601-14.
Nelson, R.B., 2006. An Introduction to Copulas, second. ed. Springer, New York.
OECD/ITF, 2016. Halving the Number of Road Deaths in. Lessons from other Countries., Korea.
Öström, M., Eriksson, A., 1993. Single-vehicle crashes and alcohol: a retrospective study of passenger car fatalities in Northern Sweden. Accid. Anal. Prev. 25 (2), 171–176. https://doi.org/10.1016/0001-4575(93)90057-4.
Padmanaban, J., RAJARAMAN, R., STADTER, G., NARAYAN, S., RAMESH, B., 2010. Analysis of In-depth Crash Data on Indian National Highways and Impact of Road Design on Crashes and Injury Severity. Berichte Der Bundesanstalt Fuer Strassenwesen. Unterr, Fahrzeugtechnik.
Pai, C.-W., Saleh, W., 2008. Modelling motorcyclist injury severity by various crash types at T-junctions in the UK. Saf. Sci. 46 (8), 1234–1247. https://doi.org/10.1016/j.ssci.2007.07.005.
Park, H.-C., Joo, Y.-J., Kho, S.-Y., Kim, D.-K., Park, B.-J., 2019. Injury severity of bus–Pedestrian crashes in South Korea considering the effects of regional and company factors. Sustainability 11 (11), 3169. https://doi.org/10.3390/su11113169.
Peng, Q., Sun, T., Zhu, P., Han, F., Matsika, E., Robinson, M., 2019. Occupant injury risk assessment and protective measures in frontal collision of a bus. Int. J. Crashworthiness 24 (4), 429–441. https://doi.org/10.1080/13588265.2018.1472722.
Prato, C.G., Kaplan, S., 2014. Bus accident severity and passenger injury: evidence from Denmark. Eur. Transp. Res. Rev. 6 (1), 17–30. https://doi.org/10.1007/s12544-013-0107-z.
Rana, T.A., Sikder, S., Pinjari, A.R., 2010. Copula-based method for addressing endogeneity in models of severity of traffic crash injuries: application to two-vehicle crashes. Transp. Res. Rec. J. Transp. Res. Board 2147 (1), 75–87. https://doi.org/10.3141/2147-10.
Rowden, P., Matthews, G., Watson, B., Biggs, H., 2011. The relative impact of work-related stress, life stress and driving environment stress on driving outcomes. Accid. Anal. Prev. 43 (4), 1332–1340. https://doi.org/10.1016/j.aap.2011.02.004.
Russo, B.J., Savolainen, P.T., Schneider, W.H., Anastasopoulos, P.C., 2014. Comparison of factors affecting injury severity in angle collisions by fault status using a random parameters bivariate ordered probit model. Anal. Methods Accid. Res. 2, 21–29. https://doi.org/10.1016/j.amar.2014.03.001.
Sam, E.F., Daniels, S., Brijs, K., Brijs, T., Wets, G., 2018. Modelling public bus/minibus transport accident severity in Ghana. Accid. Anal. Prev. 119, 114–121. https://doi.org/10.1016/j.aap.2018.07.008.
Savolainen, P., Ghosh, I., 2008. Examination of factors affecting driver injury severity in Michigan's single-vehicle–Deer crashes. Transp. Res. Rec. J. Transp. Res. Board 2078 (1), 17–25. https://doi.org/10.3141/2078-03.
Sener, I.N., Bhat, C.R., 2011. A copula-based sample selection model of telecommuting choice and frequency. Environ. Plan. A Econ. Sp. 43 (1), 126–145. https://doi.org/10.1068/a43133.
Sklar, M., 1959. Fonctions de repartition an dimensions et leurs marges. Publ. inst. Stat. univ. Paris 8, 229–231.
Spissu, E., Pinjari, A.R., Pendyala, R.M., Bhat, C.R., 2009. A copula-based joint multinomial discrete–continuous model of vehicle type choice and miles of travel. Transportation (Amst). 36 (4), 403–422. https://doi.org/10.1007/s11116-009-9208-x.
Sul, J., Lee, J., Kang, D.S., Young, L.W., Shim, J.I., Myeong, M., Huh, E., Lim, J.K., 2014. Korea's 95% Reduction in Child Traffic Fatalities: Policies and Achievements. The Korean Transport Institute, Goyang.
Tirtha, S.D., Yasmin, S., Eluru, N., 2020. Modeling of incident type and incident duration using data from multiple years. Anal. Methods Accid. Res. 28, 100132. https://doi.org/10.1016/j.amar.2020.100132.
Wali, B., Khattak, A.J., Xu, J., 2018. Contributory fault and level of personal injury to drivers involved in head-on collisions: application of copula-based bivariate ordinal models. Accid. Anal. Prev. 110, 101–114. https://doi.org/10.1016/j.aap.2017.10.018.
Wang, Y., Prato, C.G., 2019. Determinants of injury severity for truck crashes on mountain expressways in China: a case-study with a partial proportional odds model. Saf. Sci. 117, 100–107. https://doi.org/10.1016/j.ssci.2019.04.011.
Wang, K., Yasmin, S., Konduri, K.C., Eluru, N., Ivan, J.N., 2015. Copula-based joint model of injury severity and vehicle damage in two-vehicle crashes. Transp. Res. Rec. J. Transp. Res. Board 2514 (1), 158–166. https://doi.org/10.3141/2514-17.
Wang, K., Bhowmik, T., Yasmin, S., Zhao, S., Eluru, N., Jackson, E., 2019. Multivariate copula temporal modeling of intersection crash consequence metrics: a joint estimation of injury severity, crash type, vehicle damage and driver error. Accid. Anal. Prev. 125, 188–197. https://doi.org/10.1016/j.aap.2019.01.036.
Washington, S.P., Karlaftis, M.G., Mannering, F., 2003. Statistical and Econometric Methods for Transportation Data Analysis. Chapman and Hall/CRC, Washington DC.
Wu, Q., Chen, F., Zhang, G., Liu, X.C., Wang, H., Bogus, S.M., 2014. Mixed logit model-based driver injury severity investigations in single- and multi-vehicle crashes on rural two-lane highways. Accid. Anal. Prev. 72, 105–115. https://doi.org/10.1016/j.aap.2014.06.014.
Yagil, D., 1998. Gender and age-related differences in attitudes toward traffic laws and traffic violations. Transp. Res. Part F Traffic Psychol. Behav. 1 (2), 123–135. https://doi.org/10.1016/S1369-8478(98)00010-2.
Yasmin, S., Eluru, N., 2013. Evaluating alternate discrete outcome frameworks for modeling crash injury severity. Accid. Anal. Prev. 59, 506–521. https://doi.org/10.1016/j.aap.2013.06.040.
Yasmin, S., Eluru, N., Pinjari, A.R., Tay, R., 2014. Examining driver injury severity in two vehicle crashes – A copula based approach. Accid. Anal. Prev. 66, 120–135. https://doi.org/10.1016/j.aap.2014.01.018.
Yasmin, S., Momtaz, S.U., Nashad, T., Eluru, N., 2018. A multivariate copula-based macro-level crash count model. Transp. Res. Rec. J. Transp. Res. Board 2672 (30), 64–75. https://doi.org/10.1177/0361198118801348.
Ye, X., Pendyala, R.M., Al-Rukaibi, F.S., Konduri, K.C., 2008. Joint Model of Accident Type and Severity for Two-Vehicle Crashes, in: Transportation Research Board 87th Annual Meeting. Transportation Research Board, Washington DC, United States.
Yu, R., Abdel-Aty, M., 2013. Multi-level Bayesian analyses for single- and multi-vehicle freeway crashes. Accid. Anal. Prev. 58, 97–105. https://doi.org/10.1016/j.aap.2013.04.025.
Yu, S., Jia, Y., Sun, D., 2019. Identifying factors that influence the patterns of road crashes using association rules: a case study from Wisconsin, United States. Sustain. 11 (7), 1–14. https://doi.org/10.3390/su11071925.
Zhang, C., Guo, X., Wang, Y., Li, Y., Zhang, N., 2016. What leads to severe mountainous freeway crashes in southeast of China? Teh. Vjesn. - Tech. Gaz. 23, 6. https://doi.org/10.17559/TV-20150516034036.
Zheng, Z., Lu, P., Lantz, B., 2018. Commercial truck crash injury severity analysis using gradient boosting data mining model. J. Safety Res. 65, 115–124. https://doi.org/10.1016/j.jsr.2018.03.002.
Zhou, M., Chin, H.C., 2019. Factors affecting the injury severity of out-of-control single-vehicle crashes in Singapore. Accid. Anal. Prev. 124, 104–112. https://doi.org/10.1016/j.aap.2019.01.009.
Zhou, H., Yuan, C., Dong, N., Wong, S.C., Xu, P., 2020. Severity of passenger injuries on public buses: a comparative analysis of collision injuries and non-collision injuries. J. Safety Res. https://doi.org/10.1016/j.jsr.2020.04.003.
Zhu, X., Srinivasan, S., 2011. Modeling occupant-level injury severity: an application to large-truck crashes. Accid. Anal. Prev. 43 (4), 1427–1437. https://doi.org/10.1016/j.aap.2011.02.021.