The association between crashes and safety-critical events: synthesized evidence from crash reports and naturalistic driving data among commercial truck drivers

Miao Cai[a], Mohammad Ali Alamdar Yazdi[b], Amir Mehdizadeh[c], Qiong Hu[c], Alexander Vinel[c], Karen Davis[d], Fadel Megahed[d], Hong Xian[a], Steven E. Rigdon[a,*]

[a]*Saint Louis University, Saint Louis, MO, 63108*
[b]*Johns Hopkins University, Baltimore, MD, 21218*
[c]*Auburn University, Auburn, AL, 36849*
[d]*Miami University, Oxford, OH, 45056*

Abstract

The past decade has witnessed a continuous growth of naturalistic driving studies (NDSs). In NDSs, safety-critical events (SCEs) are commonly used as the outcome since crashes are very rare. However, the association between SCEs and crashes is not consistent in previous studies and has not been verified among commercial truck drivers. Based on routinely collected kinematic data from 31,881 truck drivers in a large commercial trucking company, this paper investigates the association between four types of SCEs (headways, hard brakes, collision mitigation, and rolling stability) and crashes, as well as injuries and fatalities. Bayesian negative binomial models were performed to examine this association among all drivers and drivers in different business units and driver types. It was found that a unit increase in the number of any type of SCEs per 10,000 miles was associated with 8.4% (95% credible interval CI: 8-8.8%) increase in SCEs per mile and 8.7% (95% CI: 4.8%-13.6%) increase in the number of injuries per mile. The increase was different in different types of SCEs: 3.3% (95% CI: 2.6-4%) for headways, 8.1% (95% CI: 7.5-8.7%) for hard brakes, 50.4% (95% CI: 41.4-60%) for rolling stability, and 22.2% (95% CI: 19.8-24.5%) for collision mitigation. The results are consistent when stratified by different business units and driver types. This study provides statistically strong and robust evidence that SCEs are positively associated with crashes and injuries among commercial truck drivers. NDS and kinematic data routinely collected by trucking or insurance companies provide a promising opportunity for future data analytics research.

Keywords: truck, naturalistic driving studies, safety critical events, crashes, injuries, fatalities

1. Introduction

According to the World Health Organization (WHO), road injuries were the eighth leading cause of death globally in 2018 (The WHO, 2018a). They claimed a total of 1.35 million lives every year worldwide, with 74% of them being men and boys (The WHO, 2018a). Among children and young adults aged five to 29 years old, road traffic injuries are the leading cause of death (The WHO, 2018b). Besides life losses, road injuries also cause up to 50 million non-fatal injuries and around 75.5 million disability-adjusted life years (Staton et al., 2016; The WHO, 2018b). Among different types of vehicles on the road, trucks are often associated with catastrophic consequences, which can be attributable to long routes, intensive schedule, massive size and weight, and potentially hazardous cargoes (Chen and Xie, 2014; Chen et al., 2016; Hanowski et al., 2005; Zheng et al., 2018). ~~Since costs of truck crashes are dramatic and a public concern,~~ truck crash prediction has been a widely studied topic in traffic safety research.

*Corresponding Author
Email addresses: miao.cai@slu.edu (Miao Cai), mza0052@auburn.edu (Mohammad Ali Alamdar Yazdi), azm0127@auburn.edu (Amir Mehdizadeh), qzh0011@auburn.edu (Qiong Hu), alexander.vinel@auburn.edu (Alexander Vinel), davisk4@miamioh.edu (Karen Davis), fmegahed@miamioh.edu (Fadel Megahed), hong.xian@slu.edu (Hong Xian), steve.rigdon@slu.edu (Steven E. Rigdon)

A substantial amount of road crash studies focused on using driver characteristics, traffic, weather, and others to predict the likelihood, frequency, or severity of crashes based on retrospective police report data (Lord and Mannering, 2010; Roshandel et al., 2015; Savolainen et al., 2011; Wang et al., 2013, 2013). However, truck crash data used in these studies ultimately trace back to retrospective reports, and therefore are subject to several limitations (Stern et al., 2019). First, truck crash data reported by drivers and police are post-hoc and therefore subject to recall and information bias. The reports cannot accurately record the information including real-time traffic and weather variables prior to the accidents. Second, truck crash data are extremely sparse compared to non-crashes. Federal Motor Carrier Safety Administration (2018) estimated that fatalities associated with large trucks and buses in 2017 were 0.156 per million traveled miles. In this sense, it is difficult to make statistically valid conclusions on the risk factors of crashes (Guo et al., 2010; Theofilatos et al., 2018). Third, drivers at fault tend not to report or underreport critical factors such as distraction or cell phone use, in order to avoid associated penalties (Dingus et al., 2011; Stern et al., 2019). For the same purpose, those crashes with no human injuries and little financial losses also tend to be underreported.

In view of the limitations in truck crashes data, naturalistic driving studies (NDSs) have been proposed to collect high-resolution objective data and enhance traffic safety. NDSs use sensors and devices to proactively and unobtrusively collect high-frequency kinematic and Global Positioning System (GPS) data in a real-world driving setting (Guo, 2019; Guo and Fang, 2013; Neale et al., 2005). Compared to traditional truck crash studies that are roadway-based or intersection-based case-control studies, NDS data are vehicle-based and are superior in estimating the rate or risk of events since the all non-events are collected. More importantly, NDS has the power of providing precise real-time data on kinematic events, such as acceleration or deceleration based on a pre-specified threshold, which provide the opportunity to investigate the short period prior to crashes or safety events without information bias or reporting bias (Guo et al., 2010). Since NDS data are typically collected every 30 seconds to 5 minutes, the amount of NDS data are generally enormous, which provides both an opportunity and a challenge for data analytics.

The number of NDSs worldwide has been actively increasing in the past 10 years, for example, the 100-Car NDS (Dingus et al., 2006), the second Strategic Highway Research Program (Guo, 2019), and the Europe's UDRIVE NDS (Eenink et al., 2014). These NDSs are sponsored by government departments or research organizations, and the number of participated drivers are typically limited considering the budget of the programs. However, in the recent five years, it is noticeable that an increasing number of vehicle and insurance companies have been routinely collecting real-time NDS and kinematic data, which have been applied in safety surveillance, insurance pricing, and performance evaluation fields. If these routinely collected large-scale high-resolution data are sufficiently valuable to give insights into transportation safety, they hold huge potential for follow-up data analytics research.

This study is done in collaboration with a leading freight shipping truck company in the United States. The name of the company cannot be revealed for confidentiality reasons and we will name it as Company A hereafter. The company provides transportation, delivery, and logistics services to customers and consumers in North America.

It has three business units (dedicated, intermodal, and final-mile) and three driver types (local, regional, and over-the-road). Dedicated drivers serve a single customer with familiar routes, task and work duties. Intermodal drivers work with major rail providers and transport freight container from rail yards to customer locations, which are traditional driving duties. Final-mile drivers receive task from their managers every day and make deliveries to customer locations. Regarding driver types, local drivers transport freight within a 200-mile radius and return home the same day, regional drivers move freights in a region that may include surrounding states, and over-the-road drivers pick up and delivery freights throughout the country, and they are required to on duty for at least two weeks and then take off.

The primary objective of this paper is to investigate the association between crashes and safety-critical events (SCEs) among commercial truck drivers. Detected from dynamic kinematic events, SCEs are special types of accident precursors that have all features of accidents, except that potentially catastrophic outcomes were avoided by last-second evasive maneuver (Dingus et al., 2011; Saleh et al., 2013). These SCEs are around 10 to 15 times as frequent as crashes, and past studies suggested that they were indicative of crashes (Guo et al., 2010). SCEs provide an alternative measure of transportation safety, without having to observe for a prohibitively long time to have a sufficiently large amount of crashes (Guo, 2019). However, previous studies that quantify this association are limited in the number of drivers and no studies has examined this relationship among commercial truck drivers.

This paper contributes to the following aspects. First, NDS data from more than 30,000 commercial truck drivers were collected, and the data cover most areas in the United States and are generally representative of the trucking industry. The large sample size can yield statistically significant results and ~~convincing~~ conclusions. Second, four types of SCEs are considered in this study: headways, hard brakes, collision mitigation, and rolling stability. We show that the magnitude of association between crashes and the four types of SCEs are different. Third, since most of the crashes has no human injuries or fatalities, previous studies that have smaller sample size do not have sufficient cases of injury or fatality. As the number of drivers and miles driven is relatively large, the association between injuries, fatalities, and SCEs is examined, which is a more important issue than crashes since most crashes have no human injuries.

2. Literature review

There has been an increasing number of studies evaluating the association between crashes and surrogate measures since the 1980s, with the general approach being estimating the conversion factor between the two types of events (Cooper, 1984; Evans and Wasielewski, 1983, 1982; Hydén, 1987; Risser, 1985). This topic has become a crucial issue as NDS data sets are becoming increasingly available to researchers in the recent decade. The first large-scale NDS was the 100-Car NDS, including 100 drivers in North Virginia and Washington, D.C., the United States (Neale et al., 2005). The study continuously and naturalistically recorded driving data for 102 participating drivers for one year, resulting in around two million vehicle miles and over 40,000 hours of driving data. Over the study period, 69 crashes, 761 near-crashes, and 8,295 incidents were collected. Based on this 100-Car study, Dingus et al. (2006) found that hard braking events were significantly associated with collisions and near-crashes. Since the number of near-crashes and incidents were significantly larger than crashes, they proposed using near-crashes and incidents as surrogate measures of crashes.

The rationale for using near-crashes and SCEs as surrogates for crashes is the Heinrich's Triangle, which assumes that less severe events are more frequent than severe events, and the frequency of severe events can diminish as the frequency of less severe events decreases (Guo, 2019). The former assumption is commonly seen in most NDSs and is considered to be reasonable. The association in the latter assumption can be quantitatively tested based on crash and SCEs data in large-scale NDSs, although the causal mechanism is not clear and hard to prove (Guo et al., 2010). The validity of SCEs varies significantly from studies to studies, so researchers are advised to be cautious when applying SCEs as the outcome (Guo, 2019). If the assumptions of Heinrich's Triangle are met, SCEs can provide insights to crash risk and transportation safety when the number of crashes is limited. However, the crucial question prior to using SCEs as a measure of real crashes is whether they are good surrogates of traffic crashes.

Guo et al. (2010) proposed two critical principles for using near-crashes as surrogates for crashes: 1) similar or the same causal mechanisms between crashes, surrogates, and risk factors, 2) a strong association between the frequency of surrogates and crashes. Based on the 100-Car study, they investigated the two principles using a sequential factor analysis, a Poisson regression, and sensitivity analyses. The study concluded that using near-crashes as surrogates for crashes will lead to conservative risk estimates but significantly reduce the variance of estimation. They suggested that using near-crashes as surrogates in small-scale studies will be informative for evaluating the risk of crashes.

Gordon et al. (2011) conducted a preliminary study to validate surrogates for road-departure crashes by spatially merging road geometry, average traffic, crashes, and NDS data. Bayesian seemingly unrelated Poisson models estimated with weighted least squares were used to examine if these predictor variables have the same effects on crashes and surrogates. The found that time to edge crossing and lane-departure warning were two useful surrogates for crashes on rural nonfreeway roads, but lane deviation was a poor surrogate for lane-departure crashes.

Simons-Morton et al. (2012) examined whether elevated gravitational-force predicts crashes and near-crashes among 42 newly licensed teenage drivers in Virginia. The study used the Naturalistic Teenage Driving Study that followed the recruited drivers for 18 months. A logistic regression estimated with generalized estimating equations to account for the within-subject correlation among different months was applied. The study found that the rate of elevated gravitational-force events was positively associated with the rate of crashes and near-crashes (odds ratio: 1.07, 95% confidence interval: 1.02-1.12), with an area under curve value of 0.76.

Wu and Jovanis (2012) proposed a conceptual framework to estimate the crash-to-surrogate ratio $\cancel{\pi}$ and used the 100-Car study to test the framework. The study found that the conditional probability of a crash was increased by 24 times with a lateral acceleration more than 0.7 g, but the probability was decreased by other factors such as the event occurring in daylight and dry pavement. A later study by Wu and Jovanis (2013) developed diagnostic procedures to screen crashes and near misses under NDS settings. The study applied the 100-Car NDS on the proposed framework and identified three conditions to define surrogate events: 1) a maximum lateral acceleration difference of no smaller than 0.4 g, 2) non-intersection related, and 3) maximum lateral acceleration difference no smaller than 0.9 per event or between 0.8 and 0.9 g during night time.

Guo and Fang (2013) attempted to identify risk factors of driving at driver's level and predict high-risk drivers based on the 100-Car study. The study used a negative binomial (NB) regression to examine the potential four risk factors of crashes and near-crashes. They used a K-mean clustering to classify the drivers into high-, moderate-, and low-risk groups based on crash and near-crash rates, and applied two logistic regressions to predict high- or moderate-risk drivers. The results confirmed that critical-incident event rates were significantly associated with individual driving risk. The two logistic regressions achieved AUC values of 0.938 and 0.93. They also highlighted that it was a first-step study and more studies with larger and representative data were needed to confirm the association. A similar study by Wu et al. (2014) also used the 100-Car NDS data set. This study used a Bayesian multivariate Poisson log-normal model to simultaneously account for crash frequency and severity. They also found a significant positive association between crashes, near-crashes, and crash-relevant incidents.

Pande et al. (2017) used linear referencing to link Global Positioning System (GPS) data with roadway features on 39 segments of Highway 101 in California. NB models and random-effects NB models that account for segment-specific variance were used to investigate the relationship between historic crashes and hard braking events. It was found that the freeway segments with high hard braking rates also had higher long-term crash rates, although the other three explanatory variables, average daily traffic, the presence of horizontal curvature, and auxiliary lanes were not statistically significant.

Gitelman et al. (2018) used data collected by in-vehicle data recorders on 3,500 segments of interurban roads in Israel to examine the association between two types of safety-related events (braking and speed alert) and crashes on different road types. NB models were used to account for the over-dispersion in the data, and the covariates included several road infrastructure characteristics. The number of braking events was found to be positively associated with injury crashes on single- and dual-carriageway roads, while the association was not significant on freeways. Counterintuitively, they found that speed alert events were consistently and negatively associated with injury crashes on all road types. They suggested that speed alert events were not a good surrogate for crashes, possibly due to its rough definition.

However, some researchers have been skeptical of NDS. For example, Knipling (2015) challenged the validity of using NDS data and SCEs by arguing that the purpose of traffic safety studies is to identify causes of crash harm, where crash harm is defined as property damage, injury, income lost, and all other consequences of different severities (Zaloshnja and Miller, 2007). However, NDS often use SCEs as surrogates of crashes, but very few or no crashes are involved, let alone human harm. Therefore, it is argued that SCEs are not an appropriate part of the Heinrich's

Triangle and researchers generally cannot derive valid quantitative conclusions on causations of harm based on NDS data sets. Another study by Knipling (2017) specifically targeted Hour-of-Service (HOS) rule research, such as Blanco et al. (2011) and Hanowski et al. (2008), and relevant policy revisions among commercial truck drivers. He argued that HOS studies with a quasi-experiment design were subject to confounding variables, so these studies are limited in demonstrating a causal relationship between HOS and safety outcomes. The paper also argued that NDS lacked external validity since no large truck NDS had examined the causal link between crashes and SCEs. Lastly, the construct validity was doubted since the relationship between driver fatigue, HOS, and SCEs had not been validated.

Table 1: Differences between our study and existing literature regarding sample size, driver type, crash surrogates, region, time frame, and statistical models.

Authors	sample size	Driver type	Crash surrogates	Country or region	Time frame	Statistical models
Dingus 2006	100 cars	general	braking, steering, accelerating	Northern Virginia & Washington, DC, USA	1 year	95% confidence limits modeled as a Poisson distribution
Guo 2010	100 cars	general	near crashes	Northern Virginia & Washington, DC, USA	1 year	sequential factor analysis, Poisson regression
Gorden 2011	78 drivers	general	lateral deviation, lane-departure warning, time to edge crossing	Southeastern Michigan, USA	10 months	seemingly unrelated regression, Poisson regression
Simons-Morton 2012	42 drivers	newly licensed teenagers	elevated gravitational-force	Virginia	18 months	logistic regression using generalized estimating equations
Wu 2012	100 cars	general	lateral and longitudinal acceleration, Event button, Yaw rate, Forward amd rear time-to-collision	Northern Virginia & Washington, DC, USA	1 year	logistic regression
Guo 2013	100 cars	general	critical-incident events	Northern Virginia & Washington, DC, USA	1 year	negative binomial regression
Wu 2014	100 cars	general	safety-related events	Northern Virginia & Washington, DC, USA	1 year	Poisson regression
Pande 2017	33 drivers	general	high magnitude jerks while decelerating	San Luis Obispo, California, USA	10 days	negative binomial regression
Gitelman 2018	64 vehicles	general	braking, speed alerts	Israel	1 year	negative binomial regression
This paper	**31,881 drivers**	**commercial truck drivers**	**headways, hard brakes, collision mitigation, rolling stability**	**National data, USA**	**1 year**	**Bayesian negative binomial regression**

In summary, no consistent conclusion has been reached on the association between the number of crashes and surrogates in NDS (Guo, 2019). As shown in Table 1 and the above review, there are three gaps in previous NDS studies on the association between crashes and SCEs: 1) the number of drivers involved are limited to 33 to 100, which makes it difficult to reach statistically significant conclusions on crashes, injuries, and fatalities, 2) although two separate NDS data sets : the Drowsy Driver Warning System Field Operation Test and the Naturalistic Truck Driving Study have been sponsored by the Federal Motor Carrier Safety Administration (Hickman et al., 2018), no studies have examined the association between crashes and surrogates using NDS data sets that specifically target commercial truck drivers, 3) since previous NDS data sets are limited in the number of drivers and mileage driven, no studies have examined the association between crashes surrogates and human injuries or fatalities.

3. Methodology

3.1. Data description

The data used in this study were initially collected by Company A for routine performance monitoring and driver assistance, not for research purpose ~~like the~~ 100-Car study. From April 1st, 2015 to March 31st, 2016, Company A installed data acquisition systems (DASs) to their trucks in all three business units. These DASs intermittently collected real-time driving ping data every ~~a~~ couple of seconds to 5 minutes. The collected ping data include the exact date and time of the record (year, month, day, hour, minute, and second), GPS (latitudes and longitudes specific to five decimal places), speed, and drivers' anonymized unique ID. The research team excluded drivers who 1) had less than 1,000 pings, 2) cannot be matched in driver demographics table, and 3) were identified as obvious outliers regarding the rates of SCEs. In total, $1,494,678,173$ pings collected from 31,881 commercial truck drivers were included in the current study. Geographical distributions of these included pings are shown in the results below.

Apart from the ping data, $480,331$ SCEs from the included drivers were recorded when pre-specified kinematic thresholds were triggered while driving. The collected SCEs data include the exact date and time, latitude and longitude (specific to five decimal places), driver, and type of SCEs. This SCEs data included 184,773 headways, 231,101 hard brakes, 55,345 collision mitigation, and 9,112 rolling stability. Although the detailed kinematic definitions or thresholds were not provided to the research team, non-kinematic definitions of the four types of SCEs are:

- *Headway*: the headway to a lead vehicle is closer than a preset threshold value.
- *Hard brake*: the speed decrease within a unit time is larger than a preset threshold value.
- *Collision mitigation*: an imminent collision with a forward slower moving or standing vehicle or object.
- *Rolling stability*: potential roll-over or loss-of-control situations in various road or weather conditions (dry, wet, snow and ice-covered pavement).

A crash data set that included driver ID, state, city, report time, the number fatalities, and the number of injuries, were collected by retrospective reports from Company A. In total, 35,008 crashes, 241 injuries, and 22 fatalities were collected for included drivers in this table. The company also provided a driver demographics table that includes the age, gender, race, business units, and driver types associated with the driver ID. All data were de-identified prior to the research team had access. The study protocol was approved by the Institutional Review Board of Saint Louis University.

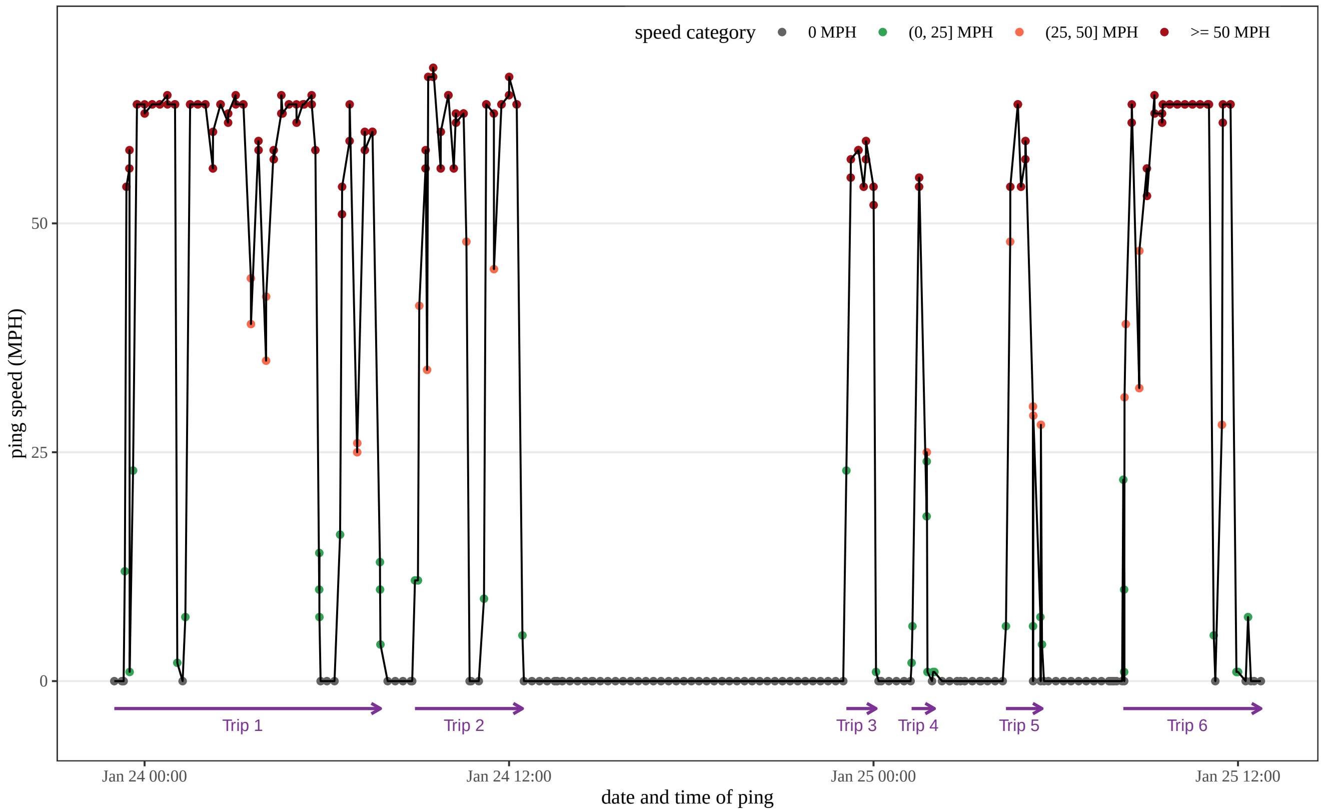


Figure 1: Data aggregation from ping data to trips for a sample of a commercial NDS data set.

3.2. Data aggregation

Since the original ping data were gigantic and difficult to interpret regarding the routing, we aggregated them to trips using the approach adopted by Pande et al. (2017). A trip is defined as a continuous period of driving with no more than 30 minutes' stop. In practice, we sorted the original ping data according to date and time for each driver. Then, if the ping speed data showed that the truck was not moving (the speed of the ping equals zero) for more than 30 minutes, the ping data were separated into two different trips.

This data aggregation process is demonstrated in Figure 1, where the x-axis shows the date and time of pings, and the y-axis presents the speed of the ping (miles per hour, MPH). Each point represented a ping at that date and time, with different colors indicating the real-time speed category. Whenever the truck stopped (the grey points) for at least 30 minutes, the pings were separated into different trips, indicated by the purple arrows in the bottom (Trip 1, Trip 2, ..., Trip 6). The trip time was then calculated by taking the difference between the trip end time and start time.

The traveled distance within a trip was approximately calculated in three steps: 1) sort the ping data according to driver, date, and time, 2) compute the distance between each two nearest ping locations for each driver using the haversine method, which assumes a spherical earth and ignore ellipsoidal effects (Sinnott, 1984), 3) sum up all the distance traveled within a trip for each driver. This algorithm aggregated the original ping data into $18,759,300$ trips, with a total of $2,323,007,042$ miles traveled in $65,706,497$ hours. The median distance of a trip is 2.61 hours and the median of miles per trip was 77.06 miles.

3.3. Statistical models

Since the outcome variable (the number of crashes for each driver) is a count variable, which is a strict non-negative integer, Bayesian NB models are used in this study. Compared with Poisson model commonly used for modeling count outcomes, NB model can adjust for the variance independently from its mean, which handles potential overdispersion or underdispersion issues in the data. Let Y_i denote the number of the outcome variables (crashes, injuries, and fatalities respectively) over a distance of T_i miles for the ith driver. We assume that Y_i has a NB distribution with the mean parameter μ_i and a common auxiliary parameter ϕ. The corresponding probability mass function of Y_i is parameterized as

$$P(Y_i|\mu, \phi) = \binom{y + \phi - 1}{y} \left(\frac{\mu}{\mu + \phi}\right)^y \left(\frac{\phi}{\mu + \phi}\right)^\phi.$$

The mean and variance of Y_i are $E[Y_i] = \mu$ and $V(Y_i) = \mu + \frac{\mu^2}{\phi}$. The inverse of ϕ controls the overdispersion, which is scaled by μ^2. By assuming that the number of SCEs per 10,000 miles has a mathematically multiplicative effect on the logarithm of rate of crashes μ_i, we have the following log-linear Bayesian NB regression

$$\begin{aligned}
Y_i &\sim \text{NB}(T_i \times \mu_i,\ \phi) \\
\log \mu_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_K x_{iK} \\
\beta_0, \beta_1, \cdots, \beta_K &\sim \text{Normal}(0, 10^2) \\
\phi &\sim \text{Exponential}(1),
\end{aligned}$$

where β_0 is the intercept, $\beta_k,\ k = 1, 2, \cdots, K$ is the coefficient of the k-th predictor variable x_k. The total miles driven T_i is considered as an offset variable to account for the mileage difference among the included drivers. The predictor variables of interest were the rates of all SCEs (the number of SCEs per 10,000 miles) and the rates of four different types of SCEs, including headways, hard brakes, collision mitigation, and rolling stability. The covariates were average speed, driver's age and gender, business units, and driver types. Business units included dedicated (DED), intermodal (INT), and final-mile (FIM), with DED being the reference group. Driver types included local (LOC), regional (REG), and over-the-road (OTR), with LOC being the reference group.

As we did not have much prior knowledge on the parameters, weakly informative priors $N(0, 10^2)$ were given to the intercept β_0 and slopes $\beta_1, \beta_2, \cdots, \beta_K$, and an exponential prior with mean parameter 1 was given to the overdispersion parameter ϕ since it was restricted to be positive. To make sure the Markov chain Monte Carlo converged, we set 4,000 iterations for each of the four chains, with the first 2,000 being warm-up iterations. The

Markov chains were considered as converged when the Gelman-Rubin diagnostic $\hat{R}$ was less than 1.1 for each variable (Gelman et al., 1992). The posterior mean and 95% credible intervals (CIs) of the incidence rate ratios, IRRs ($\exp(\beta)$) were reported. The interpretation of the IRRs in this Bayesian NB model is: as the number of SCEs per 10,000 miles increases by one unit, the number of crashes per mile will be multiplied by $\exp(\beta)$ times.

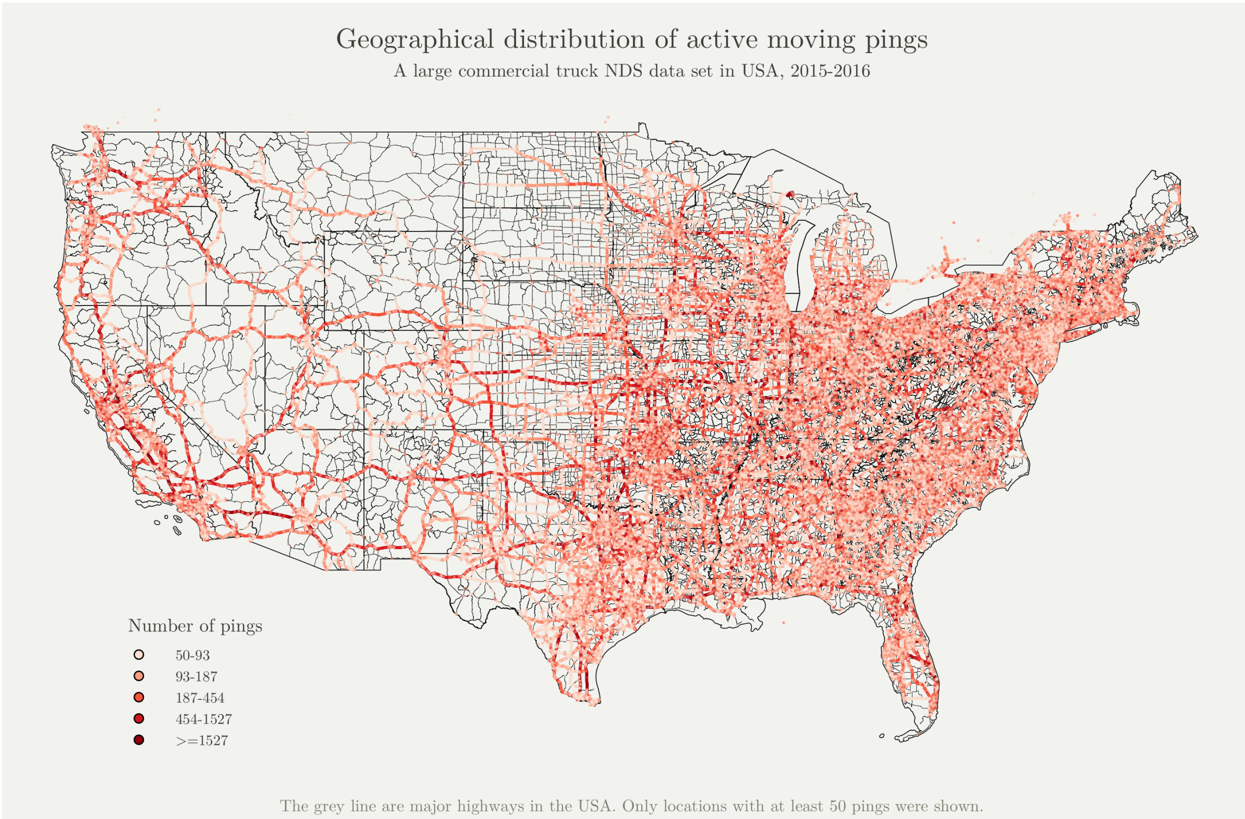
We used Pareto smoothed importance-sampling leave-one-out (PSIS-LOO) cross-validation (CV) to check the goodness-of-fit of and compare different models (Vehtari et al., 2017, 2015). Instead of exact CV that refits the model with different subsamples, the PSIS-LOO uses fast, efficient, and stable importance sampling weights to approximate leave-one-out CV (Gelfand, 1996; Gelfand et al., 1992). It estimates the expected log predicted density (ELPD), estimate number of parameters, and the LOO Information Criterion (LOOIC) for a new data set. Compared with other statistics such as Widely Applicable Information Criterion (WAIC), Deviance information criterion (DIC), and other variants (Spiegelhalter et al., 2002; Watanabe, 2010), PSIS-LOO is both fast and stable in computing. Apart from PSIS-LOO, we also used posterior predictive checks to examine the prediction accuracy (Gelman et al., 2013). The interpretation of these goodness-of-fit and model comparison statistics will be explained in the Results.

All data cleaning, statistical modeling, and visualization were performed in the statistical computing environment `R 3.6.0` (R Core Team, 2019). The haversine method distance was computed using the `distHaversine()` function in `geosphere` package (Hijmans, 2019). The Bayesian NB model was conducted using the `stan_glm()` function in `rstanarm` package (Goodrich et al., 2018). The PSIS-LOO statistics were computed using the `loo()` function in `loo` package (Vehtari et al., 2019). A sample of the original ping data, full data for Bayesian NB regression models, as well as R code for data aggregation, Bayesian NB regression models, model comparison, and diagnostic statistics can be found in the supplementary materials.

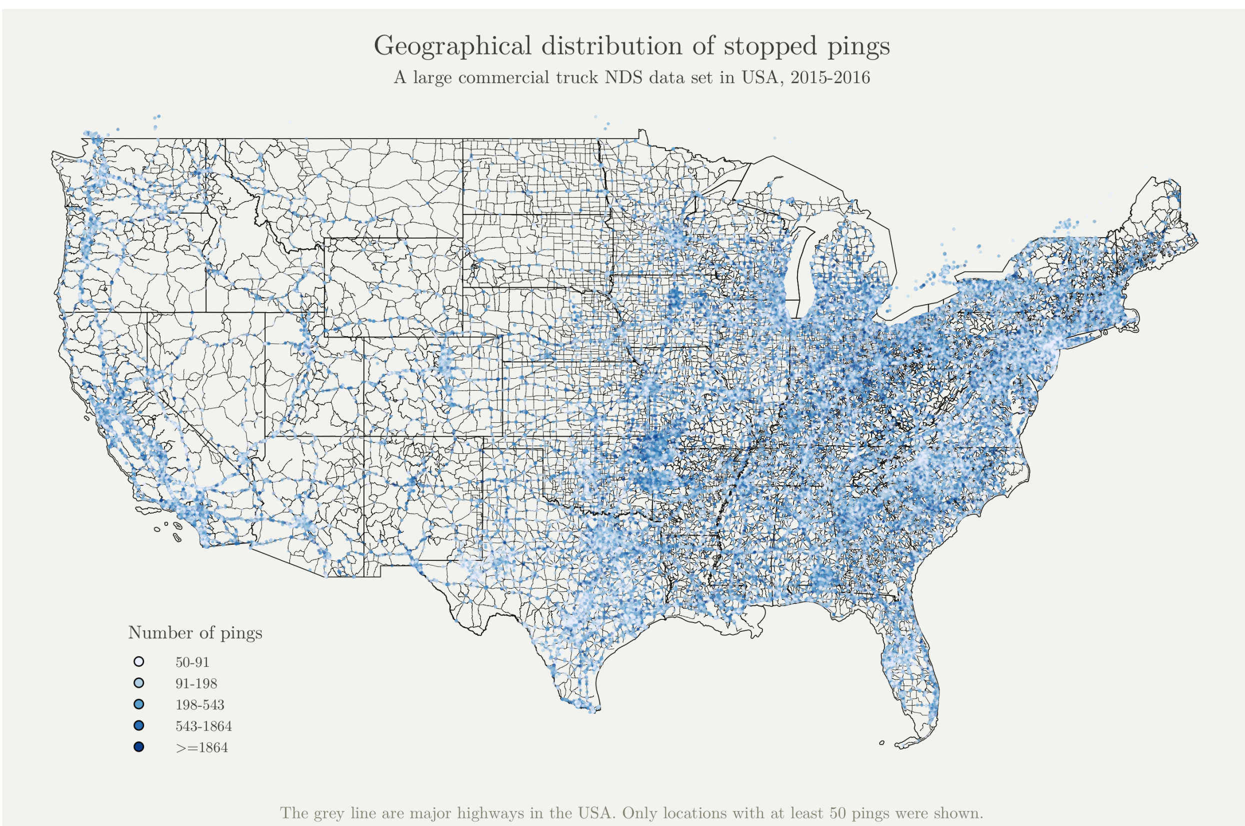
4. Results

4.1. Sample description

Figure 2 below shows the geographical point patterns of active pings that have a speed of more than 0 MPH (Figure 2a) and inactive pings that have a speed of 0 MPH (Figure 2b). Since plotting the original gigantic ping data may overwhelm the entire map, we rounded all latitudes and longitudes to two decimal places, which are approximately as large as normal-size parking lot, and then displayed only locations with more than 50 pings. The red points are the locations with at least 50 actively moving pings and blue dots are the locations with at least 50 stopping pings, with darker colors (dark red and dark blue) indicating higher number of pings. The background grey lines are the major roads in the United States (United States Geological Survey, 2014) and the solid black lines are state boundaries. The geographical point patterns suggest that most of the trucking transportation is in middle and eastern parts, as well as California and Seattle. The active and inactive pings are generally consistent, but active pings are much more concentrated than inactive pings on the major Midwest major roads.



(a) Active pings



(b) Inactive pings

Figure 2: Geographical point patterns of active and inactive pings in a large commercial NDS data set.

Among all the included 31,881 drivers, 29,296 (91.89%) were male, 1,585 (4.97%) were female, and 1,000 (3.14%) were unspecified. The mean age of the drivers were 44.48 years with a standard deviation of 11.72 years. The range of the ages were from 20 to 82 years. Regarding business units, 16,182 (50.76%) drivers served in the DED unit, in which the drivers work for a single customer and drive on familiar routes. 9,780 (30.68%) drivers were in ITM, where the drivers conduct traditional truck driving duties and transport freight containers from rail yards to local locations. 5,919 (18.57%) drivers were in FIM and made deliveries to customer locations. As for driver types, 13,399 (42.03%) were LOC drivers who transport freight within 200 miles and return home on the same day. 15,729 (49.34%) were REG drivers who move freight within a region or interstate and return home on a weekly or bi-weekly basis. Only 2,753 (8.64%) were OTR drivers who deliver and pick up freight throughout the country, and they are required to be on the road for at least two weeks and then consider taking days off or continue working.

4.2. Bayesian NB models for all included drivers

Table 2 below shows the estimates of posterior IRRs and their CIs in the Bayesian NB models for all the included drivers. We conducted models for all SCEs combined (Pooled model), four types of SCEs as four variables in one model (Four SCEs), and four SCEs separately in four models (columns 4-7). In the Bayesian setting, parameters are considered as random variales that have probabilistic distributions instead of unknown fixed values, so no P-values are reported here. Instead, we reported 95% posterior credible intervals (CIs), which can be interpreted as there is 0.95 probability that the parameters are within the interval. All the IRRs of SCEs in the six models were greater than one and none of the associated CIs included one, indicating statistically strong evidence that the rates of SCEs were positively associated with the rates of crashes. One unit increase in the number of any type of SCEs per 10,000 miles was associated with 8.4% (95% CI: 8-8.8%) increase in the number of crashes per mile. Specifically, as the number of headways, hard brakes, rolling stability, and collision mitigation per 10,000 miles increased by one unit, the number of crashes per mile were increased by 3.3% (95% CI: 2.6-4%), 8.1% (95% CI: 7.5-8.7%), 50.4% (95% CI: 41.4-60%), and 22.2% (95% CI: 19.8-24.5%) respectively, controlling for other covariates.

4.3. Bayesian NB models stratified by business units and driver types

Since the driving behaviors and routing patterns vary significantly among different business units and driver types, we also conducted Bayesian NB models for drivers in different business units and types, and the results are shown below in Table 3. The posterior IRRs and CIs of four SCEs are consistent with those in Table 2. All four type of SCEs were positively associated with the number of crashes per mile. None of the CIs included one except for headways in the DED-OTR unit. These stratified results indicate strong evidence that SCEs were positively associated with crashes in different business units and driver types. Among the four types of SCEs, rolling stability had the highest IRRs, followed by collision mitigation, hard brake, and headway.

Table 2: Bayesian NB regressions with the rate of SCEs predicting crashes, non-stratified models

variables	Pooled model	Four SCEs	Headways	Hard brakes	Rolling stability	Collision mitigation
All SCEs	1.084 (1.080, 1.088)					
Headways		1.033 (1.026, 1.040)	1.077 (1.069, 1.085)			
Hard brakes		1.081 (1.075, 1.087)		1.109 (1.102, 1.116)		
Rolling stability		1.504 (1.414, 1.600)			2.147 (2.015, 2.295)	
Collision mitigation		1.222 (1.198, 1.245)				1.343 (1.316, 1.369)
Age	0.992 (0.990, 0.993)	0.992 (0.991, 0.993)	0.989 (0.988, 0.990)	0.991 (0.989, 0.992)	0.989 (0.988, 0.991)	0.990 (0.988, 0.991)
Mean speed	0.979 (0.976, 0.982)	0.982 (0.979, 0.985)	0.971 (0.968, 0.973)	0.980 (0.977, 0.983)	0.973 (0.970, 0.976)	0.975 (0.973, 0.978)
Gender: male	0.817 (0.756, 0.886)	0.808 (0.754, 0.867)	0.848 (0.785, 0.919)	0.823 (0.762, 0.887)	0.845 (0.787, 0.909)	0.826 (0.770, 0.891)
Gender: unknown	0.975 (0.785, 1.199)	0.954 (0.777, 1.149)	1.097 (0.896, 1.347)	1.096 (0.884, 1.349)	1.018 (0.842, 1.239)	1.058 (0.870, 1.299)
(Intercept)	0.054 (0.047, 0.062)	0.048 (0.042, 0.054)	0.090 (0.079, 0.103)	0.057 (0.050, 0.066)	0.082 (0.072, 0.093)	0.073 (0.064, 0.083)
Business unit: ITM	0.698 (0.670, 0.727)	0.717 (0.690, 0.745)	0.706 (0.679, 0.735)	0.701 (0.672, 0.730)	0.735 (0.706, 0.765)	0.729 (0.700, 0.758)
Businees unit: FIM	0.907 (0.861, 0.954)	0.897 (0.852, 0.943)	0.925 (0.882, 0.971)	0.904 (0.865, 0.948)	0.922 (0.880, 0.967)	0.901 (0.859, 0.942)
Type: OTR	1.071 (0.994, 1.151)	1.094 (1.022, 1.174)	1.053 (0.981, 1.131)	1.064 (0.994, 1.140)	1.067 (0.990, 1.144)	1.106 (1.030, 1.182)
Type: REG	1.003 (0.957, 1.045)	1.012 (0.969, 1.057)	0.971 (0.928, 1.015)	0.994 (0.950, 1.037)	0.973 (0.932, 1.016)	0.984 (0.943, 1.028)
Fit statistics:						
sample size	31828	31828	31828	31828	31828	31828
elpd_loo	-39985.2 (236.5)	-39770.2 (233.5)	-40792.7 (238.9)	-40315.5 (237.2)	-40710.1 (237.8)	-40503.2 (239.4)
p_loo	18.1 (1.1)	30 (2.4)	19.8 (1.9)	18.2 (1.2)	15.9 (0.8)	16.1 (1)
looic	79970.4 (472.9)	79540.5 (467.1)	81585.4 (477.8)	80631 (474.5)	81420.1 (475.7)	81006.5 (478.7)

The SCEs were measured as the number of events per 10,000 miles driven. IRRs and associated 95% credible intervals were reported.

Table 3: Bayesian NB regressions with the rate of SCEs predicting crashes, stratified by business units and driver types

variables	DCS-LOC	DCS-OTR	DCS-REG	ITM-LOC	ITM-REG	FIM-OTR	FIM-REG
Headways	1.026 (1.011, 1.042)	1.001 (0.993, 1.010)	1.048 (1.032, 1.067)	1.026 (1.012, 1.042)	1.060 (1.038, 1.082)	1.082 (1.020, 1.149)	1.050 (1.031, 1.068)
Hard brakes	1.069 (1.057, 1.080)	1.241 (1.194, 1.293)	1.163 (1.140, 1.188)	1.047 (1.040, 1.054)	1.114 (1.093, 1.138)	1.086 (1.049, 1.131)	1.183 (1.154, 1.211)
Rolling stability	1.528 (1.367, 1.733)	1.648 (1.269, 2.229)	1.676 (1.467, 1.951)	1.419 (1.284, 1.578)	2.477 (1.590, 3.717)	4.320 (2.210, 9.522)	1.175 (1.039, 1.369)
Collision mitigation	1.163 (1.127, 1.203)	1.318 (1.132, 1.540)	1.362 (1.292, 1.440)	1.212 (1.174, 1.252)	1.577 (1.422, 1.766)	1.134 (0.952, 1.353)	1.170 (1.121, 1.234)
Age	0.992 (0.989, 0.995)	0.988 (0.982, 0.993)	0.993 (0.990, 0.996)	0.995 (0.993, 0.998)	0.986 (0.982, 0.990)	0.999 (0.989, 1.010)	0.997 (0.993, 10000)
Mean speed	0.976 (0.970, 0.983)	1.016 (1.005, 1.027)	0.968 (0.962, 0.974)	0.994 (0.987, 10000)	10000 (0.988, 1.012)	0.973 (0.958, 0.988)	0.983 (0.973, 0.994)
Gender: male	0.883 (0.702, 1.083)	0.868 (0.631, 1.227)	0.844 (0.716, 0.997)	0.749 (0.650, 0.862)	0.841 (0.691, 1.029)	0.675 (0.433, 1.027)	0.751 (0.634, 0.893)
Gender: unknown	1.065 (0.589, 1.908)	1.378 (0.706, 2.617)	0.576 (0.325, 0.980)	1.287 (0.774, 2.079)	0.194 (0.044, 0.626)		0.816 (0.571, 1.158)
(Intercept)	0.055 (0.040, 0.076)	0.015 (0.008, 0.027)	0.062 (0.046, 0.084)	0.026 (0.020, 0.033)	0.021 (0.013, 0.033)	0.047 (0.021, 0.102)	0.033 (0.022, 0.049)
Fit statistics:							
sample size	6950	1797	7405	6429	3339	943	4963
elpd_loo	-9300.8 (125.3)	-2416.9 (52.2)	-9799.1 (112.1)	-7624.2 (90.8)	-3912.6 (70.5)	-1139.8 (40)	-5293.9 (85.4)
p_loo	30.6 (5)	14.3 (2.5)	20.4 (2.6)	17.4 (2.3)	13.9 (1.7)	11 (1.6)	19.4 (2.9)
looic	18601.6 (250.6)	4833.8 (104.5)	19598.1 (224.2)	15248.3 (181.6)	7825.2 (141)	2279.5 (79.9)	10587.7 (170.8)

The SCEs were measured as the number of events per 10,000 miles driven. IRR and associated 95% credible intervals were reported.

4.4. Bayesian NB models for injuries and fatalities

Table 4 below presents the results of Bayesian NB models predicting the number of injuries and the number of fatalities respectively. Compared with the models for crashes, the result for injuries and fatalities are less conclusive since the number of injuries and fatalities is small. One unit increase in the number of any type of SCEs per 10,000 miles was associated with 8.7% (95% CI: 4.8%-13.6%) increase in the number of injuries per mile. When stratified into four different types of SCEs, all 95% CIs of IRRs included one, which indicated weak evidence for modeling injuries or fatalities, although the posterior means were positive. In the two models using the number of fatalities as the outcome variable (column 4 and 5), all 95% CIs of IRRs included one and the CIs were very wide, which suggested insufficient sample size in the number of fatalities to yield statistically significant results.

Table 4: Bayesian NB regressions with the rate of SCEs predicting crashes

variables	Injuries: pooled	Injuries: four SCEs	Fatalities: pooled	Fatalities: four SCEs
All SCEs	1.087 (1.048, 1.136)		0.973 (0.791, 1.149)	
Headways		1.061 (0.961, 1.181)		0.955 (0.592, 1.478)
Hard brakes		1.080 (0.995, 1.177)		0.957 (0.652, 1.387)
Rolling stability		1.773 (0.684, 5.439)		1.631 (0.043, 102.782)
Collision mitigation		1.174 (0.987, 1.535)		0.866 (0.200, 3.632)
Age	0.987 (0.970, 1.004)	0.986 (0.969, 1.004)	0.966 (0.912, 1.020)	0.965 (0.906, 1.030)
Mean speed	0.967 (0.929, 1.007)	0.970 (0.931, 1.009)	0.915 (0.797, 1.049)	0.910 (0.778, 1.050)
Gender: male	0.825 (0.301, 2.149)	0.800 (0.298, 2.176)	1.770 (0.074, 54.444)	1.953 (0.062, 80.045)
Gender: unknown	1.022 (0.094, 8.499)	0.993 (0.092, 9.338)	0.093 (0.000, 76.248)	0.093 (0.000, 115.564)
Business unit: ITM	0.459 (0.265, 0.788)	0.467 (0.280, 0.789)	0.354 (0.068, 1.573)	0.341 (0.044, 2.057)
Businees unit: FIM	0.710 (0.352, 1.420)	0.675 (0.330, 1.321)	1.576 (0.209, 10.438)	1.536 (0.140, 13.475)
Type: OTR	0.785 (0.321, 1.942)	0.801 (0.306, 1.955)	0.410 (0.022, 5.402)	0.388 (0.014, 6.205)
Type: REG	0.472 (0.265, 0.821)	0.463 (0.263, 0.820)	0.389 (0.064, 1.970)	0.379 (0.050, 2.214)
Intercept	0.013 (0.002, 0.070)	0.012 (0.002, 0.064)	0.008 (0.000, 1.855)	0.011 (0.000, 5.179)
Fit statistics:				
sample size	31828	31828	31828	31828
elpd_loo	-1134.5 (80.8)	-1137.3 (81.1)	-182.4 (37.9)	-182.4 (37.9)
p_loo	13.9 (3.6)	16.4 (4)	11.3 (3.2)	11.3 (3.2)
looic	2269.1 (161.5)	2274.6 (162.1)	364.7 (75.7)	364.7 (75.7)

The SCEs were measured as the number of events per 10,000 miles driven. IRR and associated 95% credible intervals were reported.

4.5. Diagnostics statistics and model selection

All the models have Pareto k diagnostic statistics of less than 0.7, which suggest no signal for model misspecification. The estimated effective number of parameters (`p_loo` in Tables 2, 3, 4), which were similar to the total number of parameters in the models. These two results suggest NB models were reasonably specified models given the large number of observations in this study (Vehtari et al., 2017, 2015). The LOOIC in the tables can be used to compare different models, with lower values indicating better models. In Table 2, the "Four SCEs" model has the lowest LOOIC (79,540.5) among the six models, although the standard errors in the bracket suggested that the "Four SCEs" model was not significantly better than the "Pooled" model.

In addition, we also checked the model fit by comparing the observed data to 100 replicated data sets generated from the parameter posterior distributions (Gelman and Hill, 2006). For each simulated data set, we computed the proportion of zero crashes and compared them to the observed proportion in the original data. Figure 3 and 4 below

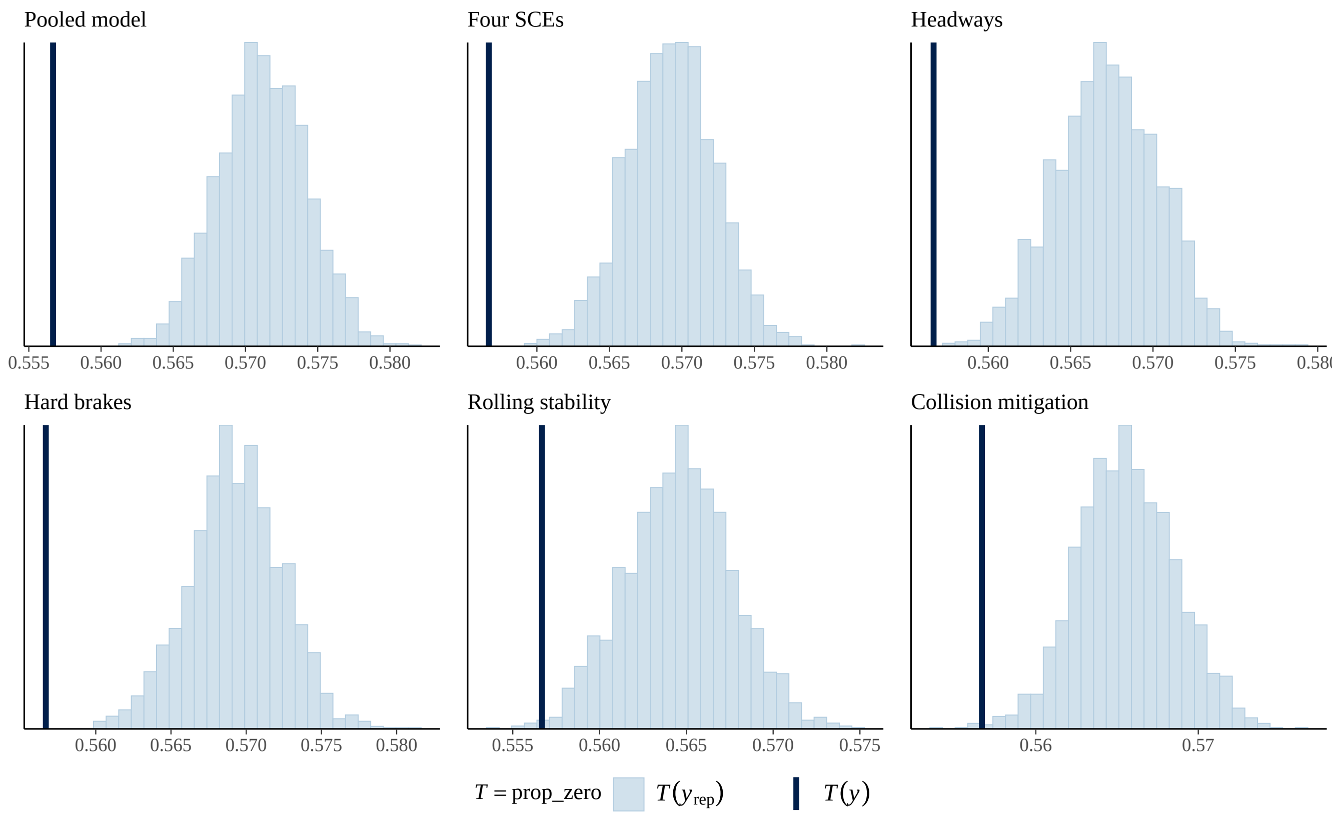


Figure 3: PPC with zero count test statistic for Bayesian negative binomial models for all drivers

present the posterior predictive check for models in Table 2 and 3. The black solid vertical lines are the observed proportion of zero crashes in observed data, while the light blue histograms show the distribution of proportion of zero crashes in the 100 simulated data sets generated from the parameter posterior distributions. Although the observed proportion was away from the distribution of proportions in simulated data sets in Figure 3, the difference between the observed proportion and the center of simulated distribution was less than 1.5%. By contrast, in Figure 4 below, the observed proportion of zero crashes located almost exactly at the center of the simulated distributions when stratified by different business units and driver types, which suggested that stratified models in Table 3 have better prediction accuracy and model fit than non-stratified models.

5. Discussion

In line with previous studies on the association between crashes and SCE in NDS (Gitelman et al., 2018; Gordon et al., 2011; Guo and Fang, 2013; Guo et al., 2010; Pande et al., 2017; Simons-Morton et al., 2012; Wu and Jovanis, 2012), this study provides statistically significant and robust evidence that there is positive association between SCEs and crashes among commercial truck drivers in the United States.

The current study contributes to the existing literature in three aspects. First, this paper overcome the small sample size issues in previous crashes and crash surrogates papers, which typically include no more than 100 drivers or vehicles (Gitelman et al., 2018; Guo et al., 2010). Our study included 31,881 commercial truck drivers, and the

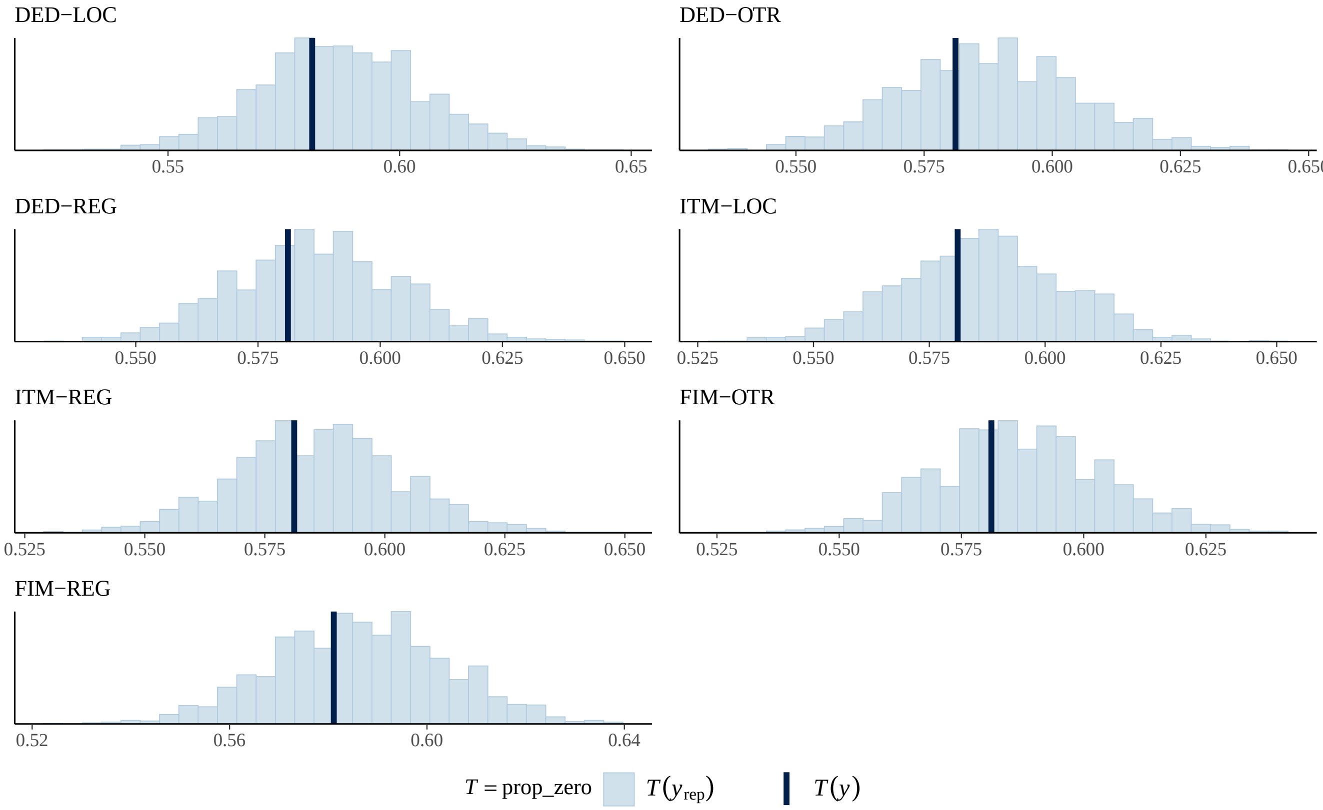


Figure 4: PPC with zero count test statistic for Bayesian negative binomial models, stratified by business unit and driver types

large sample size allows us to investigate the association between four different types of SCEs and crashes, as well as stratified analyses across business units and driver types. Second, commercial trucks are the major causes of catastrophic accidents and transportation risk, but evidence on the association between crashes and crash surrogates among truck drivers is scare. Our study gives insights to this less studied field using a nationwide large-scale sample. Third, this paper explored the association between SCEs and human injuries, which is investigated in previous papers yet important question. We found that as the number of SCEs per 10,000 miles increase by one unit, the number of injuries per mile will increase by 8.7% (95% CI: 4.8%-13.6%).

Among the four types of SCEs studies in this paper, the most frequent one is hard brakes (n = 231,101), followed by headways (n = 184,773), collision mitigation (n = 55,345), and rolling stability (n = 9,112). The number of hard brakes is around 25 times higher than that of rolling stability. Although the number of hard brakes and headways are much more than collision mitigation and rolling stability. In statistical analyses, we did not find that hard brakes and headways were over-defined. One unit increase in the number of hard brakes or headways per 10,000 miles were consistently associated with increase in crashes, and this association still hold when stratified by business units and driver types. The IRRs of hard brakes and headways were smaller than those of collision mitigation and rolling stability, which could be explained by the relatively high frequencies.

Although our study shows a positive association between SCEs and crashes, as well as injuries, we cannot conclude that the second assumption in Heinrich's Triangle is true, i.e., the frequency of severe events can diminish as the frequency of less severe events decreases (Guo et al., 2010). This study used a cross-sectional observatory design and no experiments were involved. Therefore, we should not make any causal statements on this association. Exploring the causal relationship between SCEs and crashes requires well-designed experiments or state-of-start causal inference models for observatory data, which is beyond the scope of this study and could be a direction that future studies can focus on.

Our results are based on truck driving data from a single trucking company, but we argue that the generalizability of the results to other trucking companies is high. First, Company A is a leading trucking and transportation company in the United States, and their business divisions, routing, and scheduling are not particularly special compared to other trucking companies. Second, the commercial trucking industry has a high turn-over rate (Johnson et al., 2010), which means the truck drivers actively switch between companies. Therefore, different trucking companies share a fair amount of drivers and are generally similar regarding drivers. Third, a substantial number of drivers in each of the three business units and three driver types have been included in our study. The number of truck drivers involved is at least 30 times as high as those in other relevant studies, and the traveling pings cover the entire country and most of the major roads. Nonetheless, we still suggest the researchers always investigate the relationship between SCEs and crashes or injuries prior to applying SCEs as surrogates of crashes in risk prediction.

This study has several limitations. First, since the exact time of the crashes were not recorded, we are not able to find out which and how many SCEs directly cause crashes, which is a crucial question in Heinrich's Triangle. Second, although we have a relatively large number of drivers and high mileage driven, the number of injuries or fatalities is not sufficiently large for stratified analyses. The statistical evidence on the number of injuries is strong only when we combine four types of SCEs, while the evidence on fatalities is either of the two models is not convincing given the small number of fatalities. Third, we do not have data on traffic or weather variables, which are important predictors of crashes. Fourth, since the crash data in this study were reported by the drivers, it is likely that the drivers were underreporting crashes, especially for non-injury or non-fatality injuries.

6. Summary and conclusion

Based on routinely collected NDS and kinematic data from 31,881 truck drivers in a large commercial trucking company, this study investigated the association between SCEs (headways, hard brakes, collision mitigation, and rolling stability) and crashes, injuries, and fatalities using Bayesian negative binomial regression models. We found that one unit increase in the number of any type of SCEs per 10,000 miles was associated with 8.4% (95% CI: 8-8.8%) increase in SCEs per mile and 8.7% (95% CI: 4.8%-13.6%) increase in the number of injuries per mile. The increase was different in different types of SCEs: 3.3% (95% CI: 2.6-4%) for headways, 8.1% (95% CI: 7.5-8.7%) for hard brakes, 50.4% (95% CI: 41.4-60%) for rolling stability, and 22.2% (95% CI: 19.8-24.5%) for collision mitigation. The results are consistent when stratified by different business units and driver types. This work provides statistically strong and robust evidence that SCEs are positively associated with crashes and injuries among commercial truck drivers.

7. Supplementary materials

In an effort to promote open and reproducible research, we have released our source code and aggregated data at a website hosted by GitHub Pages. Interested reader can visit it at https://blinded-research.github.io/crashSCE/. The original ping data include sensitive latitude and longitude information and cannot be made open. The website contains the following information:

A) Ping data and trip aggregation

- Ping data demonstration
- Aggregating ping data into trips
- Ping and trip data visualization

B) Statistical modeling

- Aggregated driver-level data
- Bayesian Negative binomial regression using `rstanarm`
- Model comparison and diagnostics using `loo`

Acknowledgment

This work was supported in part by: the National Science Foundation (CMMI-1635927 and CMMI-1634992); the Ohio Supercomputer Center (PMIU0138 and PMIU0162); the American Society of Safety Professionals (ASSP) Foundation; the University of Cincinnati Education and Research Center Pilot Research Project Training Program; the Transportation Informatics Tier I University Transportation Center (TransInfo); a Google Cloud Platform research grant for data management; and a Dark Sky grant for extended API access (i.e., they increased the number of possible queries per day).

References

Blanco, M., Hanowski, R.J., Olson, R.L., Morgan, J.F., Soccolich, S.A., Wu, S.-C., others, 2011. The impact of driving, non-driving work, and rest breaks on driving performance in commercial vehicle operations.

Chen, C., Xie, Y., 2014. The impacts of multiple rest-break periods on commercial truck driver's crash risk. Journal of safety research 48, 87–93.

Chen, G.X., Fang, Y., Guo, F., Hanowski, R.J., 2016. The influence of daily sleep patterns of commercial truck drivers on driving performance. Accident analysis & prevention 91, 55–63.

Cooper, P., 1984. Experience with traffic conflicts in canada with emphasis on “post encroachment time” techniques, in: International Calibration Study of Traffic Conflict Techniques. Springer, pp. 75–96.

Dingus, T.A., Hanowski, R.J., Klauer, S.G., 2011. Estimating crash risk. Ergonomics in Design 19, 8–12.

Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M.A., Hankey, J., Ramsey, D., Gupta, S., others, 2006. The 100-car naturalistic driving study. Phase 2: Results of the 100-car field experiment. United States. Department of Transportation. National Highway Traffic Safety.

Eenink, R., Barnard, Y., Baumann, M., Augros, X., Utesch, F., 2014. UDRIVE: The european naturalistic driving study, in: Proceedings of Transport Research Arena. IFSTTAR.

Evans, L., Wasielewski, P., 1983. Risky driving related to driver and vehicle characteristics. Accident Analysis & Prevention 15, 121–136.

Evans, L., Wasielewski, P., 1982. Do accident-involved drivers exhibit riskier everyday driving behavior? Accident Analysis & Prevention 14, 57–64.

Federal Motor Carrier Safety Administration, 2018. Large Truck and Bus Crash Facts 2017.

Gelfand, A.E., 1996. Model determination using sampling-based methods. Markov chain Monte Carlo in practice 145–161.

Gelfand, A.E., Dey, D.K., Chang, H., 1992. Model determination using predictive distributions with implementation via sampling-based methods. STANFORD UNIV CA DEPT OF STATISTICS.

Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B., 2013. Bayesian data analysis. Chapman; Hall/CRC.

Gelman, A., Hill, J., 2006. Data analysis using regression and multilevel/hierarchical models. Cambridge university press.

Gelman, A., Rubin, D.B., others, 1992. Inference from iterative simulation using multiple sequences. Statistical science 7, 457–472.

Gitelman, V., Bekhor, S., Doveh, E., Pesahov, F., Carmel, R., Morik, S., 2018. Exploring relationships between driving events identified by in-vehicle data recorders, infrastructure characteristics and road crashes. Transportation research part C: emerging technologies 91, 156–175.

Goodrich, B., Gabry, J., Ali, I., Brilleman, S., 2018. Rstanarm: Bayesian applied regression modeling via Stan.

Gordon, T.J., Kostyniuk, L.P., Green, P.E., Barnes, M.A., Blower, D., Blankespoor, A.D., Bogard, S.E., 2011. Analysis of crash rates and surrogate events: Unified approach. Transportation research record 2237, 1–9.

Guo, F., 2019. Statistical methods for naturalistic driving studies. Annual review of statistics and its application 6, 309–328.

Guo, F., Fang, Y., 2013. Individual driver risk assessment using naturalistic driving data. Accident Analysis & Prevention 61, 3–9.

Guo, F., Klauer, S.G., Hankey, J.M., Dingus, T.A., 2010. Near crashes as crash surrogate for naturalistic driving studies. Transportation Research Record 2147, 66–74.

Hanowski, R.J., Olson, R., Bocanegra, J., Hickman, J.S., 2008. Analysis of risk as a function of driving-hour: Assessment of driving-hours 1 through 11. US Department. of Transportation, Federal Motor Carrier Safety Administration.

Hanowski, R.J., Perez, M.A., Dingus, T.A., 2005. Driver distraction in long-haul truck drivers. Transportation Research Part F: Traffic Psychology and Behaviour 8, 441–458.

Hickman, J.S., Hanowski, R.J., Bocanegra, J., 2018. A synthetic approach to compare the large truck crash causation study and naturalistic driving data. Accident Analysis & Prevention 112, 11–14.

Hijmans, R.J., 2019. Geosphere: Spherical trigonometry.

Hydén, C., 1987. The development of a method for traffic safety evaluation: The swedish traffic conflicts technique. Bulletin Lund Institute of Technology, Department.

Johnson, J.C., Bristow, D.N., McClure, D.J., Schneider, K.C., 2010. Long-distance truck drivers discuss their occupation. Journal of Transportation Law, Logistics & Policy 77.

Knipling, R.R., 2017. Threats to scientific validity in truck driver hours-of-service studies.

Knipling, R.R., 2015. Naturalistic driving events: No harm, no foul, no validity.

Lord, D., Mannering, F., 2010. The statistical analysis of crash-frequency data: A review and assessment of methodological alternatives. Transportation research part A: policy and practice 44, 291–305.

Neale, V.L., Dingus, T.A., Klauer, S.G., Sudweeks, J., Goodman, M., 2005. An overview of the 100-car naturalistic study and findings. National Highway Traffic Safety Administration, Paper 5, 0400.

Pande, A., Chand, S., Saxena, N., Dixit, V., Loy, J., Wolshon, B., Kent, J.D., 2017. A preliminary investigation of the relationships between historical crash and naturalistic driving. Accident Analysis & Prevention 101, 107–116.

R Core Team, 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Risser, R., 1985. Behavior in traffic conflict situations. Accident Analysis & Prevention 17, 179–197.

Roshandel, S., Zheng, Z., Washington, S., 2015. Impact of real-time traffic characteristics on freeway crash occurrence: Systematic review and meta-analysis. Accident Analysis & Prevention 79, 198–211.

Saleh, J.H., Saltmarsh, E.A., Favaro, F.M., Brevault, L., 2013. Accident precursors, near misses, and warning signs: Critical review and formal definitions within the framework of discrete event systems. Reliability Engineering

& System Safety 114, 148–154.

Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: A review and assessment of methodological alternatives. Accident Analysis & Prevention 43, 1666–1676.

Simons-Morton, B.G., Zhang, Z., Jackson, J.C., Albert, P.S., 2012. Do elevated gravitational-force events while driving predict crashes and near crashes? American journal of epidemiology 175, 1075–1079.

Sinnott, R., 1984. Virtues of the haversine. Sky and telescope.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the royal statistical society: Series b (statistical methodology) 64, 583–639.

Staton, C., Vissoci, J., Gong, E., Toomey, N., Wafula, R., Abdelgadir, J., Zhou, Y., Liu, C., Pei, F., Zick, B., others, 2016. Road traffic injury prevention initiatives: A systematic review and metasummary of effectiveness in low and middle income countries. PLoS One 11, e0144971.

Stern, H.S., Blower, D., Cohen, M.L., Czeisler, C.A., Dinges, D.F., Greenhouse, J.B., Guo, F., Hanowski, R.J., Hartenbaum, N.P., Krueger, G.P., others, 2019. Data and methods for studying commercial motor vehicle driver fatigue, highway safety and long-term driver health. Accident Analysis & Prevention 126, 37–42.

Theofilatos, A., Yannis, G., Kopelias, P., Papadimitriou, F., 2018. Impact of real-time traffic characteristics on crash occurrence: Preliminary results of the case of rare events. Accident Analysis & Prevention.

The WHO, 2018a. The top 10 causes of death.

The WHO, 2018b. Global status report on road safety 2018. World Health Organization.

United States Geological Survey, 2014. USGS Small-scale Dataset - 1:1,000,000-Scale Major Roads of the United States 201403 Shapefile.

Vehtari, A., Gabry, J., Yao, Y., Gelman, A., 2019. Loo: Efficient leave-one-out cross-validation and waic for bayesian models.

Vehtari, A., Gelman, A., Gabry, J., 2017. Practical bayesian model evaluation using leave-one-out cross-validation and waic. Statistics and computing 27, 1413–1432.

Vehtari, A., Gelman, A., Gabry, J., 2015. Pareto smoothed importance sampling. arXiv preprint arXiv:1507.02646.

Wang, C., Quddus, M.A., Ison, S.G., 2013. The effect of traffic and road characteristics on road safety: A review and future research direction. Safety science 57, 264–275.

Watanabe, S., 2010. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. Journal of Machine Learning Research 11, 3571–3594.

Wu, K.-F., Aguero-Valverde, J., Jovanis, P.P., 2014. Using naturalistic driving data to explore the association between traffic safety-related events and crash risk at driver level. Accident Analysis & Prevention 72, 210–218.

Wu, K.-F., Jovanis, P.P., 2013. Defining and screening crash surrogate events using naturalistic driving data. Accident Analysis & Prevention 61, 10–22.

Wu, K.-F., Jovanis, P.P., 2012. Crashes and crash-surrogate events: Exploratory modeling with naturalistic driving data. Accident Analysis & Prevention 45, 507–516.

Zaloshnja, E., Miller, T., 2007. Unit costs of medium and heavy trucks. Report No. FMCSA-RRA-07-034). Washington, DC: Federal Motor Carrier Safety.

Zheng, Z., Lu, P., Lantz, B., 2018. Commercial truck crash injury severity analysis using gradient boosting data mining model. Journal of safety research 65, 115–124.