Is an informed driver a better decision maker? A random parameters heterogeneity-in-means approach to investigate the impact of the connected environment on driving behaviour in safety-critical situations

Anshuman Sharma[a], Zuduo Zheng[a1], Jiwon Kim[a], Ashish Bhaskar[b], Md. Mazharul Haque[b]

[a]*School of Civil Engineering, The University of Queensland, St. Lucia 4072, Brisbane, Australia*
[b]*School of Civil Engineering and Built Environment, Science and Engineering Faculty, Queensland University of Technology (QUT), 2 George St, Brisbane, Qld. 4001, Australia*

Abstract

The motive of sharing the information in a connected environment (CE) is to assist a driver in operational, tactical, and strategic decision making and improving driving task performance. The influence of such information assistance on driver decision making and task performance during safety-critical events is not well understood. Thus, this study focusses on understanding the impact of CE on the acceleration noise and the response time as indicators of task performance and the decision making involved in safety-critical events. To overcome the paucity of CE data, an advanced driving simulator experiment is designed and conducted. Three categories of uninterrupted information are available to drivers in CE scenario, namely continuous information, on-time event-triggered information, and advanced event-triggered information. The safety-critical event designed in the simulator experiment is the leader's hard braking behaviour in car-following regime. In CE scenario, drivers receive an advanced message for this safety-critical event. To model drivers' decision in safety-critical situations, random parameters modelling approaches are adopted to account for the unobserved heterogeneities in drivers' decision. Consequently, a random parameters hazard-based duration model and a random parameters linear regression model—both with heterogeneity in parameter means—are estimated for the response time and the acceleration noise, respectively. Results show that the acceleration noise reduces in CE while the response time can either increase or decrease in CE compared to those in the traditional environment. To better understand this mixed effect on response time, a decision tree analysis is conducted. For human factors, the results demonstrate that young drivers take more advantage of CE relative to the middle-aged or old drivers. Overall, drivers exhibit stable driving behaviour because they have more time to react and thus, are at low risk in safety-critical situations in CE.

Keywords: connected vehicle; connected environment; safety; random parameters; decision tree; decision making; task performance.

1. INTRODUCTION

Connected vehicles are equipped with driver assistance systems to receive, process, and then display information such as the position and speed of a leader, space gap to the leader, and traffic

[1] Corresponding author. Tel.: +61 7 3443 1371
E-mail address: zuduo.zheng@uq.edu.au (Z. Zheng)

incident at the downstream, via vehicle-to-vehicle, vehicle-to-infrastructure, and vehicle-to-everything communications. The motive of sharing the information is to assist a driver in operational (steering, acceleration/deceleration, monitoring, and braking), tactical (lane-change, responding to a sudden event, deliberately driving slow or fast etc.), and strategic (origin-destination points etc.) decision making and improving driving task performance (Talebpour et al., 2016, 2015). A few recent studies have demonstrated using CE empirical data that CE improves the safety margin associated with mandatory lane-changing manoeuvres (Ali et al., 2019, 2018). However, in car-following scenarios, it is still unclear whether such information assistance can indeed improve the driver decision making and the driving task performance. Sharma et al. (2017) highlight that human factors such as age, gender, education, and driving experience influence the decision making and the driving task performance of the traditional vehicles (no information sharing or assistance) and the connected vehicles. However, the nature and the magnitude of such influence are still largely unexplored. Thus, this study focusses on understanding the impact of connected environment and human factors on task performance and the decision making involved in safety-critical events during car-following. We choose acceleration noise and response time as two indicators of task performance and decision making because of the reasons mentioned in the ensuing paragraphs. In this study, a traffic environment consisting of only traditional vehicles is termed as the traditional environment (TE), whereas a traffic environment consisting of only the connected vehicles is termed as the connected environment (CE). Note that the driver has full control of the connected vehicle.

The acceleration noise, first introduced by Herman et al. (1959), is defined as the standard deviation of acceleration. It is calculated for a given road segment or for a given time period. The acceleration noise is a speed variation measure and used to characterise drivers and roadway conditions. A high acceleration noise is observed amongst reckless drivers (Herman et al., 1959), young drivers (Belz and Aultman-Hall, 2011; Ko et al., 2010), faster drivers (Jones and Potts, 1962), distracted drivers (Saifuzzaman et al., 2015), hazardous road conditions (Herman et al., 1959), and small carriage way (Jones and Potts, 1962). In addition, the acceleration noise is a measure of traffic flow quality on freeways and arterials. Drew et al. (1967) first suggested acceleration noise as a level-of-service measure using an energy-acceleration noise model, and a high acceleration noise depicts traffic congestion or drop in level-of-service (Babu and Pattnaik, 1997; Jones and Potts, 1962; Kim et al., 2003; Ko et al., 2006). All these studies underscore the significance of acceleration noise as a measure of microscopic traffic flow characteristics and driver behaviour.

Note that TE forms the backdrop of abovementioned studies on acceleration noise. Contrarily, due to the novelty of CE, the following question is underexplored: How CE and human factors e.g., age, gender, education, etc., influence the acceleration noise? To the best of the authors' knowledge, only one study has analysed the acceleration noise in CE. Farah et al. (2012) in their pioneer work reported no statistically significant difference in the acceleration noise when the V2I communication was available versus when it was unavailable. A total of 35 participants were recruited in that study and the safety messages disseminated were speed advice, upstream accident advice, road conditions ahead, lane keeping advice, and speed limit advice. Note that V2V communication and the corresponding messages such as speed of and spacing to the leader were unavailable.

Sharma et al. (2019b) define the response time as the time taken by a driver to adjust his/her speed against a stimulus, with or without deliberately delaying his/her decision. The response time is a critical driving behaviour parameter that influences traffic flow efficiency and traffic safety. It also provides a robust measure of the driver attention (Scott and Gray, 2008). Minimum response time, i.e., reaction time, as explained in Sharma et al. (2019b), is a fundamental car-following parameter and an integral part of several microscopic traffic flow models (refer to Saifuzzaman and Zheng, (2014) for a review on car-following models and Zheng (2014) for a review on lane-changing models). Previous studies demonstrate that an increase in response time compromises the traffic flow stability in TE and CE (Sun et al., 2018; Talebpour and Mahmassani, 2016; Treiber et al., 2007). In the car-following scenario, increased response time corresponds to low-speeds (Mehmood and Easa, 2009), large headways (Mehmood and Easa, 2009; Schweitzer et al., 1995), older drivers (Mehmood and Easa, 2009; Warshawsky-Livne and Shinar, 2002), female drivers (Mehmood and Easa, 2009), and distracted drivers (Caird et al., 2008; Haque and Washington, 2014). The response time also reflects the delay experienced between the identification of a stimulus that can lead to a potential collision and the application of the control measures such as brake pedal press and accelerator release. Undoubtedly, response time is a key measure of safe manoeuvres and thus a part of a number of surrogate safety measures such as unsafe following condition (Son et al., 2011), potential index for collision with urgent deceleration (Uno et al., 2002), and warning index (Moon et al., 2009).

A limited research has been conducted using driving simulators or connected vehicle test beds to investigate how CE influences the response time. Using driving simulator experiments, researchers report early response (Wu et al., 2018) and shorter response time (Chang and Wei, 2013) to the rear-end collision warning in CE. Researchers also postulate different response time distributions for CE and TE (Ni et al., 2011). A few studies claim that the response time will be short in CE compared to TE (Ni et al., 2011; Talebpour and Mahmassani, 2016) because additional information can enhance driver's decision making. Conversely, studies also suggest a longer response time in CE because of the same reason i.e., informed drivers make better decisions (Olia et al., 2016). These contradictory arguments and limitations of previous studies highlight the need of answering the following question: How CE and human factors, e.g., age, gender, education, etc., influence the response time?

This study answers both the research questions raised above. Paucity of CE empirical data is one of the reasons for limited research on understanding the impact of CE on driver decision making in safety-critical events. Hence, a driving simulator experiment has been carefully designed and conducted to collect the connected vehicle trajectory data necessary for assessing driving task performance and decision-making in the car-following scenario. The trajectory data are collected for TE and CE. In the experiment, two safety-critical events are created, one in the high-speed and another in the low-speed car-following regions. Here, a safety-critical event is defined as the leader's unexpected hard braking behaviour. Such events can lead to rear-end collisions if drivers are unable to anticipate and react to the leader's behaviour. The acceleration noise and the response time of each participant are measured in the two safety-critical events. The acceleration noise and the response time are modelled using random parameters models (Washington et al., 2010) because of the existence of heterogeneity in drivers' car-following behaviour and how they percieve CE (it is likely that different drivers will percieve CE differently).

The remainder of the paper is organised as follows. Section 2 describes the driving simulator experiment design and data collection in detail. Section 3 presents the dataset for analysis. Section 4 discusses the methodology, i.e., an overview of random parameters duration and linear regression models, and Section 5 presents the model estimation and assessment results. Section 6 discusses findings from the model estimation and their practical implications, and suggests future research directions.

2. DATA COLLECTION

The absence of the real-world connected vehicle data is a big challenge for researchers to understand the potential impacts of CE on driver decision making and task performance. In this research, driving simulator experiments are designed and conducted using the CARRS-Q Advanced Driving Simulator to collect the necessary connected vehicle data. Refer to Haque and Washington (2014) for more details on the advanced driving simulator.

2.1 Participants

Seventy-eight eligible participants were recruited. A participant is eligible if he/she is between 18 to 65 years old, holds either a provisional or an open Australian driving licence, has no history of motion sickness or epilepsy, and is not pregnant. Participants received AU $75 as a compensation of their time.

Table 1 Descriptive statistics of all participants.

Driver characteristics	Average	SD	Count	Percentage (%)
Driver's age (years)	30.8	11.7	-	-
Gender				
Male	-	-	50	64.1
Female	-	-	28	35.9
Education				
Primary	-	-	2	2.5
Junior (Grade 10)	-	-	1	1.3
Senior (Grade 12)	-	-	18	23.1
TAFE or Apprenticeship	-	-	9	11.5
University	-	-	48	61.6
Licence type				
Open	-	-	62	79.5
Provisional	-	-	16	20.5

Years of driving	12.2	11.5	-	-
Kilometers driven in a typical year				
0-5,000 km	-	-	10	12.8
5,001-10,000 km	-	-	19	24.4
10,000-15,000 km	-	-	15	19.2
15,001-20,000 km	-	-	18	23.1
20,001-25,000 km	-	-	6	7.7
> 25,000 km	-	-	10	12.8
Crash involvement in last one year				
Involved	-	-	8	10.3
Not involved	-	-	70	89.7
Frequency of driving per week				
Less than 2 times	-	-	5	6.4
2-4 times	-	-	28	35.9
5-6 times	-	-	16	20.5
7-8 times	-	-	7	9.0
More than 8 times	-	-	22	28.2
Prior information about Connected Vehicles				
Yes	-	-	33	42.3
No	-	-	45	57.7
General feeling towards Connected Vehicles Technology[2]				
Very negative	-	-	0	0
Negative	-	-	3	9.1
Neutral	-	-	12	36.4
Positive	-	-	9	27.3
Very positive	-	-	9	27.3

Table 1 provides the descriptive statistics of the 78 participants. From Table 1, the average age of the participants is 30.8 (SD 11.7) years. Out of the 78 participants, 35.9% are the female participants, and the average age of the females and the males is 24.9 (SD 6.7) and 34.1 (SD 12.6) years, respectively. Participants who held an open licence are 62 (79.5%), and 68 (87.2%) participants have driven more than 5000 km in the last one year. Only 33 (42.3%) participants have heard of connected vehicles or connected vehicle technology prior to the experiment, and among them, 36.4% have neutral feelings towards the connected vehicle technology. Importantly, the participants have a diverse background and the data collected from the experiments have a reasonable representativeness as evident from the descriptive statistics. Note that all participants

[2] Only those participants have answered this question who responded “Yes” when asked “Have you heard of Connected Vehicles or Connected Vehicle Technology PRIOR TO today’s experiment.”

adhered to the same testing protocol (refer to (Ali et al., 2018) for more details on the participant testing protocol).

2.2 Experiment design

Two scenarios are designed in this experiment, namely TE and CE. In TE, drivers receive no information, whereas, in CE, drivers receive uninterrupted dissemination of information simulating V2V and V2I communications. Participants are required to drive the simulator car in TE and CE. In each scenario, participants have to follow a platoon of vehicles on a single-lane motorway for 3 km. Figure 1(a) displays the road geometry. Hereon, the simulator car driven by a participant is termed as 'follower', the vehicle immediately in front of the simulator car (the first leader) is termed as 'leader' (see Figure 1(b)), and the platoon of vehicles in front of the simulator car is termed as 'leading cars'. Figure 1(c) depicts the participant's view showing different messages displayed on the windscreen of the simulator car.

Several effective strategies have been carefully implemented to minimise any potential learning effect and to ensure the realism of the participants' driving experience, e.g., the randomised sequence of the drives, different driving environment and surrounding traffic in each drive, a break between the drives, etc. Moreover, CE is cautiously designed after a comprehensive review of the literature on in-vehicle driver assistance systems and the current driving aids provided by major car manufacturers. In the ensuing paragraphs, we describe the CE design.

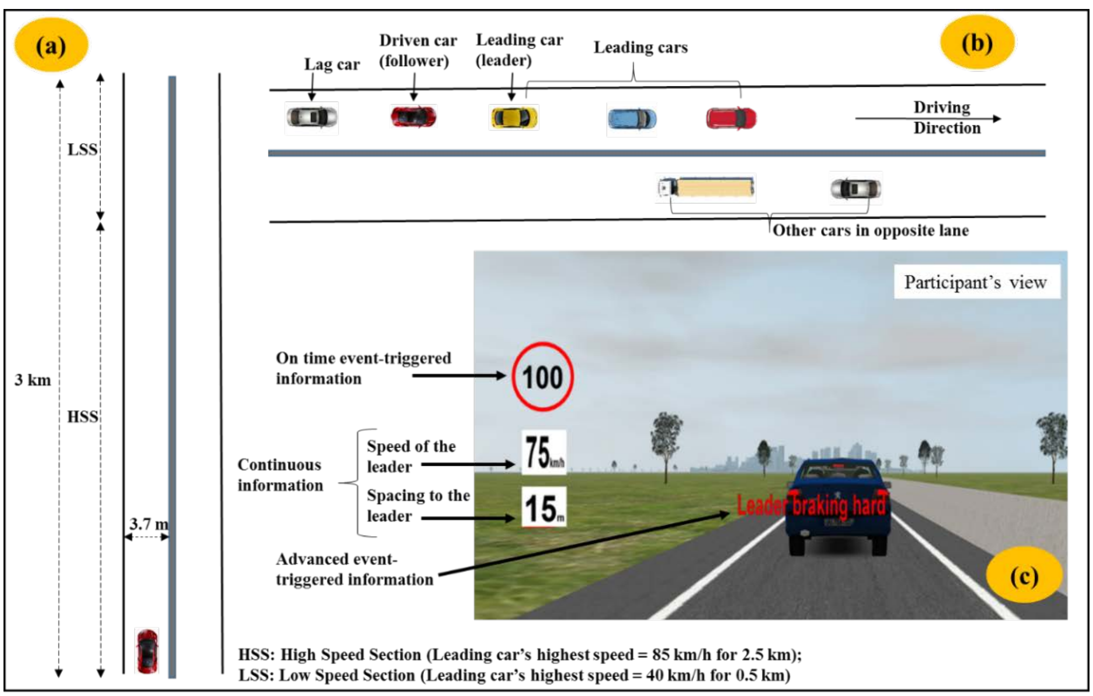

Figure 1 Details of the driving simulator experiment (Sharma et al. (2019a)). (a) The road geometry; (b) The vehicles in the simulator environment; (c) Categories of the information available in the connected scenario.

2.2.1 Connected environment design

Continuous information, on-time event-triggered information, and advanced event-triggered information are the three categories of information delivered to participants (Figure 1(c)). The continuous information is always present on the windscreen, the on-time event-triggered information is delivered to participants depending on the occurrence of specific events, and the advanced event-triggered information is provided to participants 3 seconds before the event occurs (we assume that the connectivity technologies will be smart enough to warn the drivers about the safety-critical events like hard braking in advance). For instance, in Figure 1(b) we can observe the driven car (say f1) is following three leaders (l1->l2->l3). Let's assume such a scenario on motorway in reality. After a few kilometers, the first leader (l3) unexpectedly brakes hard. In TE, vehicles f1, l1, and l2 will brake after observing that the vehicle immediately in front has braked (l1 will brake after observing that l2 has braked). Multivehicle anticipation also plays a role in followers' responses but majorly a follower responds to the vehicle immediately in front. Therefore, if vehicle l3 commences sudden braking at 't' seconds then f1 will commence the braking at around 't + Δt' seconds where Δt represents the time taken by f1 to react to l1's speed reduction. In CE, vehicles can recive information about leaders' actions. Thus, as soon as l3 starts braking, all the vehicles behind (in this case three vehicles) will receive the warning message "leader braking hard" at the same time. For the vehicle l2, the information will be an on-time event triggered infomration; however, for f1 it will be an advanced event triggered information since l1 has not started braking yet.

The continuous information includes messages about the speed of and the spacing to the leader, the on-time event-triggered information includes speed limit warning, tailgating warning, and the message 'front vehicle accelerating' whenever the leader starts accelerating from a standstill, and the advanced event-triggered information includes the warning message 'leader braking hard.' A combination of both the audio and the visual is adopted for presenting the information that is preferred in previous studies as well (Adell et al., 2011; Fairclough et al., 1997; Ghadiri et al., 2013; Lee et al., 2002; May et al., 1995). The speed limit and tailgating warnings are accompanied with a beep sound, and the remaining warning messages are accompanied with 3 beep sounds.

2.2.2 Vehicle interaction design

The interactions are designed such that drivers undergo all the driving regimes transitions resulting in a dataset containing complete trajectories. A complete trajectory constitutes free-flow regimes, namely free acceleration and cruising at the desired speed, and car-following regimes, namely following the leader at a constant speed, accelerating behind a leader, decelerating behind a leader, and standing behind a leader (Sharma et al., 2018a). Complete trajectories are an important aspect of the trajectory data quality; for more information refer to Sharma et al. (2018a) and Sharma et al. (2018b). Note that this study focusses on car-following regimes only. Figure 2 displays speed profiles of the follower and the leader in the high-speed and the low-speed car-following regions.

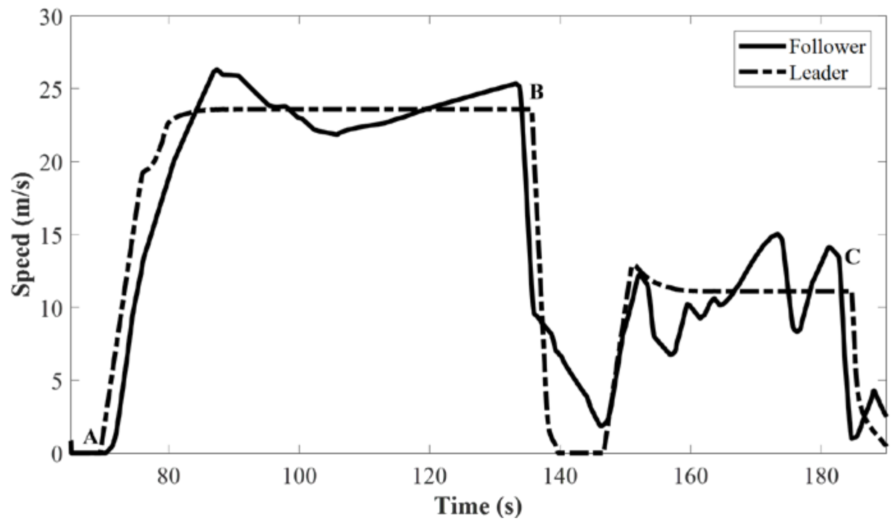

Figure 2 An example of speed profiles of the follower and the leader in CE.

At the beginning of the car-following scenario, the leader and the follower are at standstill. The leader starts accelerating (point A in Figure 2) and continues to accelerate until it attains a speed of 23.5 m/s (85 km/h), and maintains this speed for 50 s. The follower reacts to the leader's stimulus as per his/her discretionary as demonstrated in Figure 2. Next, the leader decelerates hard to mimic the hard braking (point B in Figure 2 represents the beginning of the first safety-critical event) and arrives at a standstill. After 5 s, the leader goes through the same cycle of acceleration, constant speed, hard deceleration (point C in Figure 2 represents the beginning of the second safety-critical event), and standstill, although this time, the constant speed maintained is 11 m/s (40 km/h), much smaller than the previous constant speed in order to create a low-speed car-following region. Note that the vehicle interactions remain the same in TE and CE. Moreover, in CE, advanced event-triggered information ('leader braking hard') is provided to all participants at the start of the two safety-critical events, i.e., at B and C.

3. DATASET FOR ANALYSIS

As mentioned in the introduction section, the acceleration noise and the response time are selected as indicators to demonstrate the impact of CE on driver's task performance and decision making in the car-following scenario. Moreover, we also investigate the impact of traffic flow parameters and human factors on the acceleration noise and the response time. This section describes explanatory variables used in the statistical modelling, and the calculation of dependent variables, namely the acceleration noise and the response time in safety-critical events in high and low-speed car-following regions.

3.1 Leader's hard braking (Safety-critical event)

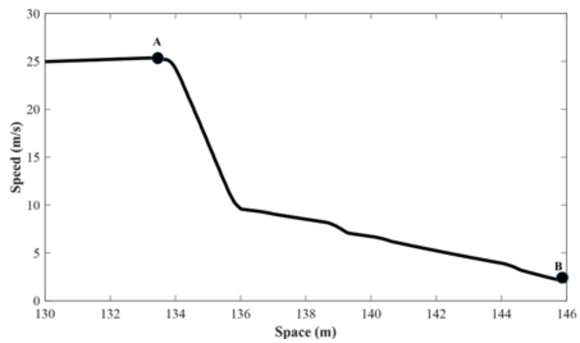

Figure 3 A schematic representation of the follower's response to the leader's hard braking.

The response of the follower commences after noticing the leader's hard deceleration in TE or the message "leader braking hard" in CE. Say the follower response starts at point A and ends at point B as illustrated in Figure 3. All the explanatory traffic flow variables such as speed, spacing, and time headway are measured at the time when the leader starts a hard deceleration in TE or the message "leader braking hard" is delivered in CE. Furthermore, the response variables (acceleration noise and response time) are measured in the region AB. The mathematical formulations of the acceleration noise and the response time are presented in Equations (1) and (2), respectively:

$$Acceleration\ noise\ = \sqrt{\frac{\sum_{i=1}^{n}(v_i - \bar{v})}{n-1}} \tag{1}$$

$$Response\ time = T_{ar} + T_{bp} \tag{2}$$

where v_i is the speed at the i^{th} location between A and B; $\bar{v}$ is the average speed in the section AB, and n is the total number of speed observations in the section AB. T_{ar} is the time difference between when the stimulus is presented (i.e., when the leader's hard deceleration starts in TE or when the message is delivered in CE) and when the accelerator pedal is released (here, ar denotes accelerator release); and T_{bp} is the time difference between when the accelerator pedal is released and when the brake pedal is pressed (here, bp denotes brake pedal press).

Table 2 presents all the explanatory variables considered when estimating statistical models. Age is divided into three categories, namely young work force (YWP), middle-aged workforce (MWP), and old-aged workforce (OWP), based on the service age group that indicates the level of demand for services that target people at different stages in life and how that demand is changing (.idcommunity (2016)).

Table 2 Summary statistics of the explanatory variables included in the models for the acceleration noise and the response time.

Variables and their type	Description and coding	Other
CE (categorical)	1 if connected environment scenario and 0 if baseline scenario	
Speed (continuous)	Measured at A in m/s. Speed is obtained directly from the driving simulator data	Mean = 24.1 m/s; SD = 1.9 m/s (TEHS) Mean = 11.1 m/s; SD = 1.2 m/s (TELS) Mean = 23.6 m/s; SD = 1.8 m/s (CEHS) Mean = 11.2 m/s; SD = 0.9 m/s (CELS)
Spacing (continuous)	Measured at A in m. Spacing is obtained directly from the driving simulator data	Mean = 90.8 m; SD = 77.6 m (TEHS) Mean = 39.1 m; SD = 32.2 m (TELS) Mean = 100.8 m; SD = 63.7 m (CEHS) Mean = 45.9 m; SD = 24.4 m (CELS)
Time headway (continuous)	Measured at A in s. Time headway is calculated as the ratio of spacing over speed	Mean = 3.7 s; SD = 1.9 s (TEHS) Mean = 3.5 s; SD = 1.2 s (TELS) Mean = 4.2 s; SD = 1.7 s (CEHS) Mean = 4.1 s; SD = 0.9 s (CELS)
Low-speed (categorical)	1 if low-speed car-following region and 0 if the high-speed car-following region	1 low-speed section in each TE and CE for every participant (total 78 in each TE and CE)
Females (categorical)	1 if female and 0 if Male	28 female drivers out of 78
Young workforce (YWP) (categorical)	1 if the age is between 18 years to 34 years and 0 otherwise	55 young drivers out of 78
Middle-aged workforce (MWP) (categorical)	1 if the age is between 35 years to 49 years and 0 otherwise	14 middle-aged drivers out of 78
Old workforce (OWP) (categorical)	1 if the age is between 50 years to 65 years and 0 otherwise	9 old drivers out of 78
University (categorical)	1 if university degree and 0 otherwise	48 drivers have a university degree
Experience (categorical)	1 if the driving experience is greater than or equal to 5 years, and 0 otherwise	54 participants with at least 5 years driving experience

TEHS – Traditional environment high-speed section; TELS – Traditional environment low-speed section; CEHS – Connected environment high-speed section; CELS – Connected environment high-speed section.

Note that the acceleration noise and the response time are measured in both the safety-critical events (one occurred in the high-speed region and another in the low-speed region) and in both TE and CE for all 78 participants. Thus, in total, there are 312 observations (78 × 2 × 2) of both the acceleration noise and the response time.

4. METHODOLOGY

4.1 Random parameters duration model for response time

This section explains the concept of random parameters hazard-based duration modelling approach for modelling the response time. The same philosophy is applicable to other modelling frameworks such as count models (Anastasopoulos and Mannering, 2009; Rusli et al., 2018), choice models (Behnood and Mannering, 2017; Hensher et al., 2015; Zheng et al., 2016), linear regression models (Lu and Xin, 2018) and duration models (Anastasopoulos and Mannering, 2015; Hasan et al., 2013; Tavassoli Hojati et al., 2013).

Duration models are best suited when it is required to model the elapsed time until the occurrence of an event. These models can be estimated using least-squares regression; however, estimation techniques that are based on survival or hazard functions provide additional insights into the underlying duration process (Washington et al., 2010). Recently, duration models (hazard or survival) have been the preferred choice of researchers when modelling the response time, i.e., the time taken by the follower to react to the stimulus (Bella and Silvestri, 2017; Choudhary and Velaga, 2017; Fu et al., 2016; Haque and Washington, 2014). This study also estimates a duration model, more specifically, a survival model with random parameters for the response time.

The survival function (Equation (3)) provides the probability of time T (also called survival time or duration variable) being greater than or equal to some specified time t (Washington et al., 2010).

$$S(t) = P(T \geq t) \tag{3}$$

In the present case, the duration variable T is the response time measured as per Equation (2); and if we assume t equal to 1.5 s, then, $S(t)$ provides the probability of $T \geq 1.5$ s.

There are two approaches to incorporate the effect of covariates on the survival function, namely the proportional hazard and the accelerated failure time (AFT). AFT approach assumes that covariates rescale (accelerate/decelerate) time directly in a baseline survival function as depicted in Equation (4):

$$S(t/X_1) = S_0[te^{\beta_1 X_1}] \tag{4}$$

where S_0 is the underlying survival function, X_1 is a vector of covariates, and β_1 is a vector of estimable parameters. AFT also assumes that the log of survival time varies linearly against the covariates as shown in Equation (5):

$$\ln(T) = \beta_1 X_1 + \varepsilon \tag{5}$$

The AFT formulations (Equation (4) and (5)) allow an intuitive and straightforward understanding of how covariates affect survival time.

When estimating Equation (4), a distributional assumption is imposed on the survival function. Based on the type of distributional assumptions, we can classify the survival models into parametric, semi-parametric, and non-parametric. Refer to Washington et al. (2010) for a comparison of the three types of survival models. Using the parametric approach, the survival function has a specific distribution including lognormal, exponential, Weibull, and log-logistic, and it is easy to incorporate the effect of covariates. Therefore, this study estimates a parametric survival model.

The choice of a distribution in parametric models, as mentioned by Haque and Washington (2014), is critical because it influences the effectiveness and the unbiasedness of the estimated parameters. The response time values calculated in this study follows a lognormal distribution. The literature also documents that the response time follows a lognormal distribution (Koppa, 2000; Rakha et al., 2007). For the demonstration purpose, Figure 4 displays the survival curve for response time values corresponding to the low-speed safety-critical event in TE and the fitted lognormal distribution. Moreover, the Anderson-Darling test results confirm that the best-fit distribution for the response time is the lognormal distribution (test statistic = 0.18, critical value = 0.74, p-value = 0.91). Therefore, in this study, the underlying distribution for the baseline survival function is the lognormal distribution. The parametric survival and density functions for the lognormal distribution are provided in Equations (6) and (7), respectively:

$$S(t) = 1 - \Phi\left\{\frac{ln(t) - \mu'}{\sigma'}\right\} \tag{6}$$

$$f(t) = \frac{1}{t\sigma'\sqrt{2\pi}} exp\left[\frac{-1}{2(\sigma')^2}\{ln(t) - \mu'\}^2\right] \tag{7}$$

where $\Phi\{\}$ is the standard normal cumulative distribution, $\mu' = \beta_1 X_1$, and σ' is the standard deviation. The lognormal distribution exhibits nonmonotonic hazard rates (initially increasing and then decreasing).

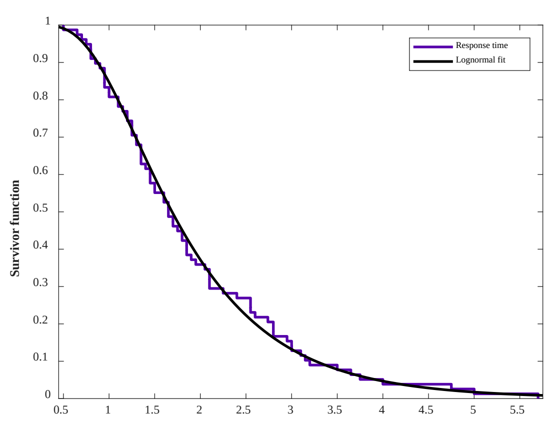

Figure 4 Distribution of response time with the fitted lognormal distribution.

The formulations in Equations (4) to (7) assume that β_1 is constant across individuals i.e., the effect of X_1 on T (or the dependent variable) is the same for all individuals. For the random parameters counterpart of the parametric AFT model with a lognormal distribution, β_1 is replaced by β_{1i} in Equations (4) to (7). For instance, the random parameters formulation for Equation (5) can be written as provided in Equation (8):

$$\ln(T) = \beta_{1i}X_1 + \varepsilon \tag{8}$$

where β_{1i} is the individual specific parameter that varies across individuals with mean μ and scale parameter σ according to a prespecified distribution. This signifies that the effect of X_1 on Y varies across individuals. Rest of the notations have the same meaning as explained before.

The random parameters model accounts for the unobserved factors that affect individual's sensitivity, e.g., consider the effect of CE on the drivers' driving behaviour. Different individuals can perceive CE differently, which would result in different sensitivities towards this environment as demonstrated in recent studies; e.g., Ali et al. (2018) and Ali et al. (2019) showed that different individuals opted for different gaps and gap time, respectively, when assisted with lane change advisory messages during mandatory lane changing. More recently, Sharma et al. (2019a) modelled different levels of driver compliance behaviour against the information received in CE. Factors such as gender, education, trust in technology, and driver aggressiveness, which are not commonly collected, can certainly influence an individual's sensitivity to CE. The random parameters model explicitly incorporates this unobserved heterogeneity–effect of variables unknown to the analyst or left out during data collection–by allowing estimated parameters in the model to vary across individuals. Ignoring unobserved heterogeneity leads to model misspecification, and biased and inefficient estimates of model parameters, which in turn lead to erroneous inferences and predictions. For a comprehensive discussion on random parameters model, see Mannering et al. (2016).

The heterogeneity itself can be a function of explanatory variables, commonly known as heterogeneity-in-means approach, as shown in Equation (9):

$$\beta_{1i} = \mu + \Theta Z_1 + \sigma r_i \tag{9}$$

where Z_1 is a vector of explanatory variables (a potential source of heterogeneity such as age, gender, and education) that influences the mean, Θ captures the influence of Z_1 on β_{1i}, and r_i is a random variable defined as per the specified distribution of β_{1i}.

4.2 Random parameters linear regression model for acceleration noise

A simple linear regression model is provided in Equation (10):

$$Y = \beta_0 + \beta_1 X_1 + \varepsilon \tag{10}$$

where Y is a vector of the observed responses of individuals (dependent variable) and for this study it represents log-transformed acceleration noise values, β_0 is a constant term representing the expected mean of Y when $X_1 = 0$, and other variables are explained before. Note that β_1 is constant, i.e., the effect of X_1 on Y is the same for all individuals. The random parameters formulation for the above model is provided in Equation (11):

$$Y = \beta_0 + \beta_{1i} X_1 + \varepsilon \tag{11}$$

The heterogeneity in the mean of β_{1i} is modelled using the same way as shown in Equation (9) and the interpretation of β_{1i} is also the same as explained in the previous section.

4.3 Estimating random parameters models

The random parameters models are estimated using the maximum simulated likelihood technique. The number and the approach of drawing parameter values are important elements of model estimation using simulation. In this regard, Halton sequence approach is an efficient way to obtain accurate approximations of the parameters using simulation (Bhat, 2003). Hence, in this study, 1000 Halton draws are executed when estimating the random parameters model as recommended in Bhat (2003) and adopted in Behnood and Mannering (2017) and Zheng et al. (2016).

A likelihood ratio test is carried out to evaluate whether the random parameters model is statistically superior to its fixed parameter counterpart and constant only models. The likelihood ratio test statistic is chi-square distributed (χ^2 statistic), and calculated as shown in Equation (12):

$$-2[LL(\hat{\beta}_s) - LL(\hat{\beta}_c)] \tag{12}$$

where the log likelihood at the convergence of the simpler model is $LL(\hat{\beta}_s)$ and the log likelihood at the convergence of the complex model is $LL(\hat{\beta}_c)$. In addition, the Akaike Information Criterion (AIC) criteria as a goodness of fit measure is considered since it penalises the addition of variable in the model. AIC values are calculated for the random parameters model, the fixed parameters model, and the constant only model. Moreover, the model with the minimum AIC is the best model among these candidate models.

The log transformed acceleration noise is the dependent variable in the random parameters linear regression model and it satisfies the linear model assumptions whereas response time is the dependent variable in the random parameters duration model. A large number of models are estimated by considering all the variables listed in Table 2 as independent variables and then following the standard statistical model building process. Note that the models reported in this paper are parsimonious models and are the best performing models in terms of the logical soundness, and based on the log-likelihood ratio test and AIC criteria.

5. MODELS ESTIMATION AND RESULTS

5.1 Random parameters lognormal survival AFT model for response time: Estimation and Evaluation results

Table 3 shows the estimation results for the random parameters survival model fitted to the response time data. The model is estimated by specifying a lognormal distribution for the baseline survival function. Moreover, the parameter for CE is assumed as random and normally distributed. The non-random parameters are time headway, OWP, and an interaction term (Low-speed × CE). Other potential candidates for random and non-random parameters are also considered when estimating the model. Table 3 presents the best performing model in terms of the logical soundness and goodness of fit measures. Equations (13) and (14) detail the mathematical formulation of the response time function.

$$\ln(Response\ time) = 0.09 + \beta_c \times CE + 0.05 \times time\ headway + 0.16 \times OWP + 0.25 \times (Low\ Speed \times CE) \tag{13}$$

$$\beta_c = 0.29 - 0.19 \times YWP + 0.28 \times n \tag{14}$$

where n is the standard normally distributed random variable. Table 4 presents a comparison of the models estimated. The random parameters model is better than the fixed parameter and the constant only models based on the likelihood ratio test and AIC values.

Table 3 Random parameters AFT model estimates of response time including heterogeneity in the random parameter.

Variable	**Parameter estimate ($\boldsymbol{\beta}$)**	**Standard error**	**p-value**	$\boldsymbol{exp(\beta)}$
Random parameters				
CE (mean)	0.29	0.094	0.002	1.336
CE (standard deviation)	0.28	0.038	<0.001	1.323
Heterogeneity-in-mean of random parameter (CE)				
YWP	-0.19	0.096	0.043	0.827
Non-random parameters				
Constant	0.09	0.051	0.092	1.094
Time headway	0.05	0.003	<0.001	1.051
OWP	0.16	0.091	0.083	1.173
Low-speed × CE	0.25	0.066	<0.001	1.284
Standard deviation for survival distribution	0.46	0.019	<0.001	1.584
Model statistics				
Number of observations	312 (78×4)			
Log-likelihood at convergence	-223.78			

Table 4 Models comparison statistics for response time models.

Models compared (H_0 = simpler model is better)	$\boldsymbol{\chi^2}$ **statistic**	**p-value**
Constant vs. Fixed parameters model	72.86	<0.001
Constant vs. Random parameters model	86.36	<0.001
Fixed vs. Random parameters model	13.5	0.001
AIC values		

Constant only model	538.2
Fixed parameters model	473.3
Random parameters model	463.6

5.1.1 Effect of random parameter on response time

The mean and standard deviation of the coefficient of CE are statistically significant at 90% confidence level. Figure 5(a) displays the distribution of CE's coefficient. It can be observed from Figure 5(a) that 85 % of the generated CE coefficients are positive. Note that if a fixed parameter model is estimated, the CE will produce a positive and statistically significant parameter, which would signify response time only increases in CE. Such conclusion would be misleading. Conversely, the random parameters model estimation reveals that the effect of CE on response time is non-uniform not only in magnitude but also in nature, i.e., there is a class of drivers whose response time increases in CE and another class of drivers whose response time decreases in CE. However, for the current sample, response time increases for the majority of drivers in CE. Furthermore, to better illustrate our finding, we develop the probability density and survival curves of response time in TE and CE, as shown in Figures 5(b) and (c), respectively. These plots are developed using the mathematical formulations as reported in Equations (6) and (7), while the values of other explanatory variables are fixed, e.g., time headway is 2 s, and OWP and the interaction term are set to 1. Specifically, three curves are plotted for CE and one for TE. The three curves for CE represent drivers having CE's coefficient equal to (i) the mean of CE's coefficient (CE in Figures 5(b) and 5(c)); (ii) the upper limit of CE's coefficient i.e., mean + 1.645 × standard deviation (CE_UL in Figures 5(b) and 5(c)); and (iii) the lower limit of CE's coefficient, i.e., mean - 1.645 × standard deviation (CE_LL in Figures 5(b) and 5(c)), respectively. The upper and lower limits are based on the 90% confidence interval.

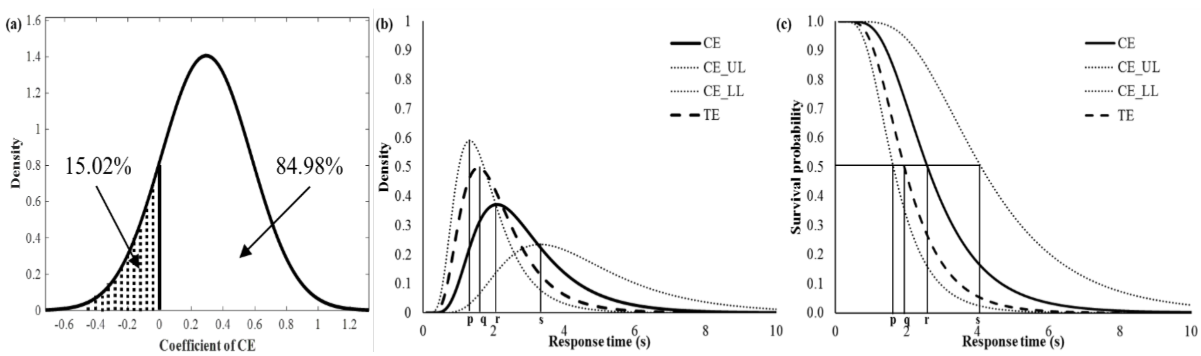

Figure 5 (a) Distribution of coefficient of CE; (b) Probability density plots of response time for TE, CE_LL, CE, and CE_UL; and (c) Survival plots of RT for TE, CE_LL, CE, and CE_UL.

Note that a probability density curve answers the question: What is the probability that response time is equal to t s? On the contrary, a survival curve answers the question: What is the probability that response time ≥ t s? The inflection points, namely p, q, r, and s on the curves of CE_LL, TE, CE, and CE_UL, respectively, (see Figure 5(b)) represent that the probability of observing a certain response time increases up to these points and then decreases. The response time at inflection points p, q, r, and s are 1.3, 1.6, 2.1, and 3.3 s, respectively. In addition, the probability that the response time is equal to 1 s for CE_LL, TE, CE, and CE_UL is 0.49, 0.31, 0.10, and 0.01, respectively, and the probability that the response time is equal to 4 s for CE_LL, TE, CE, and CE_UL is 0.03, 0.06, 0.14, and 0.22, respectively.

From the survival curves, the probability that the response time ≥ 2 s is 0.32, 0.47, 0.71, and 0.94 for CE_LL, TE, CE, and CE_UL, respectively. Furthermore, survival curves can be used to calculate the k^{th} percentile ratio or k-PR ratio. Equation (15) provides the formulation of k-PR ratio (Q_k):

$$Q_k = \frac{k^{th} percentile\ of\ the\ survival\ distribution\ for\ the\ control\ group}{k^{th} percentile\ of\ the\ survival\ distribution\ for\ the\ treatment\ group} \tag{15}$$

In this study, the treatment group can be CE, CE_UL, or CE_LL, whereas the control group is TE. This study computes $Q_{0.5}$ (k = 0.5), which is also called as the median ratio, the mostly preferred percentile ratio in the literature (Barrett et al., 2012; Michiels et al., 2005; Spruance et al., 2004). From Figure 5(c), the response times at the 50th percentile for CE_LL, TE, CE, and CE_UL are equal to 1.6, 1.9, 2.6, and 4.1 s, respectively. The median ratios for CE_LL, CE, and CE_UL are 1.2, 0.73, and 0.46. A median ratio > 1, as in treatment group CE_LL, indicates that the median response time (the time at which, on average, 50% of the drivers have reacted) decreases in CE, and a median ratio <1, as in the treatment groups CE and CE_UL, indicates that the median response time increases in CE. These observations confirm that, as compared to TE, the response time can increase or decrease in CE.

Table 3 shows that the heterogeneity in CE's impact on response time is related to age groups. A simulation based approach is adopted to understand the impact of heterogeneity in CE on response time. The response time values in CE are calculated for 1000 values of n for both YWP = 1 and YWP = 0 (middle-aged and old drivers), while keeping the other explanatory variables constant. Figure 6 illustrates the influence of YWP on response time in CE. The mean response time is short for young drivers (1.6 s) as compared to that for middle aged and old drivers (2.2 s) in CE. The response time is mostly low for young drivers, although there is a possibility that for some values of n, response time can be high for young drivers as compared to that for middle-aged and old drivers. The possible reasons behind this behavioural difference are discussed in Section 6.

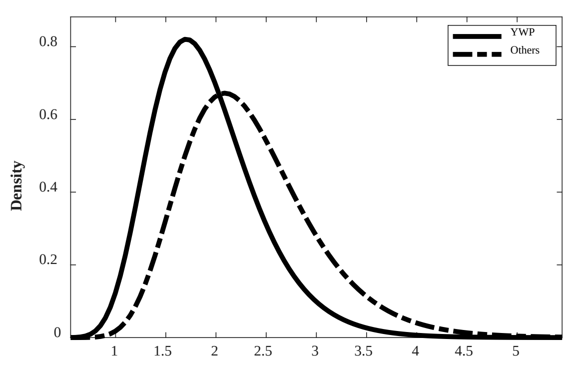

Figure 6 Distributions of response time for young drivers vs. others (middle-aged and old drivers).

5.1.2 Effects of non-random parameters on response time

The coefficients of time headway, OWP, and the interaction term (Low-speed × CE) are positive and statistically significant. To gain insights into the effects of time headway, OWP, and the interaction term on response time, an intuitive way is to calculate the exponents of coefficients that translate to a percentage change in response time from a unit increase for continuous variable (e.g., time headway) and a change from zero to one for indicator variables (e.g., OWP). Table 3 presents the exponents of coefficients and the corresponding change in response time. Note that, in AFT models, the percentage change in the dependent variable corresponding to an explanatory variable is calculated at the sample means of the other variables. Based on Table 3, a 1 s increase in time headway is likely to increase the response time by 5%. One possible reason is that at large time headways drivers have more time to react to the leader's action. Moreover, response time is expected to increase by 17% for old drivers as compared to others i.e., middle-aged and young drivers. Lastly, about 28% increase in response time is more likely for the low-speed car-following region as compared to the high-speed region in CE.

A decision tree based classification approach is utilised to understand the factors behind an increase or a decrease in the response time in CE (refer to Section 5.3).

5.2 Random parameters linear regression model for acceleration noise: Estimation and Evaluation results

Tables 5 presents the model estimation results for acceleration noise and Table 6 presents the random parameters distribution and heterogeneity estimates of random parameters in acceleration noise model, respectively. The coefficients of CE and spacing are treated as random parameters following a normal distribution. The normal distribution is assumed because of the uncertainty in the sign of the coefficients of CE and spacing. The coefficient of CE is assumed as random because of the reasons stated in Section 4. Furthermore, the existence of inter-driver heterogeneity is well established in traffic flow literature (Ossen and Hoogendoorn, 2007, 2011); hence, it is justified to assume the coefficient of spacing as random.

This study assumes 90% confidence level to assess the significance of the explanatory variables. The estimated model is mathematically presented in Equations (16), (17), and (18):

$$\ln(acceleration\ Noise) = 2.088 + \beta_c \times CE + \beta_{sp} \times spacing - 0.982 \times low\ speed - 0.104 \times university - 0.001 \times (spacing \times females) \quad (16)$$

$$\beta_c = -0.502 - 0.254 \times YWP + 0.141 \times n \quad (17)$$

$$\beta_{sp} = -0.008 + 0.001 \times n \quad (18)$$

where n is a standard normally distributed random variable. Table 7 reports the model evaluation results based on the likelihood ratio test. Based on the likelihood ratio test and AIC values, the random parameters model is better as compared to the fixed parameter and the constant only models.

Table 5 Random parameters linear regression model estimates of acceleration noise.

Variable	**Parameter estimate (β)**	**Standard error**	**p-value**	$exp(\beta)$
Random parameters				
Spacing (mean)	-0.008	0.000	<0.001	1.008
CE (mean)	-0.502	0.089	<0.001	1.652
Non-random parameters				
Constant	2.088	0.091	<0.001	
Low-speed	-0.982	0.064	<0.001	2.667
University	-0.104	0.057	<0.001	1.109
Spacing × Females	-0.001	0.000	0.066	1.001
Model statistics				
Number of observations	312 (78 participants × 4 groups)			
Log-likelihood at convergence	-217.05			

Table 6 Distribution estimates and heterogeneities in the random parameters.

Random parameters (normal distribution is assumed for each parameter)	**Parameter estimate (β)**	**Standard error**	**p-value**
Spacing (Standard deviation)	0.001	0.000	<0.001
CE (Standard deviation)	0.141	0.038	<0.001
Heterogeneity in the mean of random parameter (CE)			
YWP	-0.254	0.084	0.003

Table 7 Model comparison statistics for acceleration noise models.

Models compared in log likelihood ratio test(H_0 = parsimonious model is better)	χ^2 **statistic**	**p-value**
Constant vs. Fixed parameters model	300.4	<0.001
Constant vs. Random parameters model	316.4	<0.001
Fixed vs. Random parameters model	16.02	0.001
AIC values		
Constant only model	755.2	
Fixed parameters model	464.5	
Random parameters model	454.1	

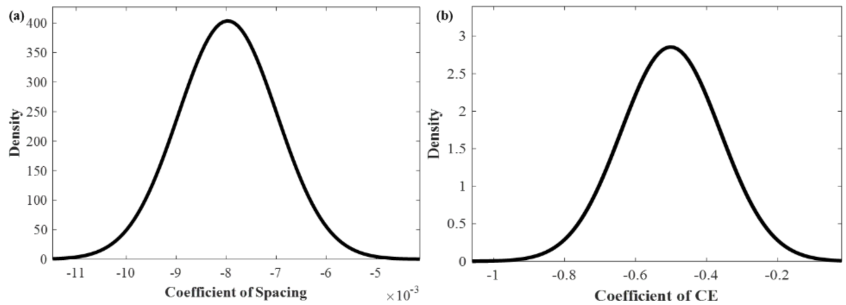

Figure 7 Distributions of coefficients of (a) spacing, and (b) CE.

5.1.1 Effect of random parameters on acceleration noise

The mean and the standard deviation (Tables 5 and 6) are significant for both the random parameters. To comprehend the impact of CE and spacing on acceleration noise the distributions of the coefficients are plotted in Figure 7 (using the mean and the standard deviation of random parameters). The variation in the nature (positive or negative) and the magnitude of coefficients in these distributions reflect the heterogeneity among participants as captured by the random coefficients.

For CE, a statistically significant standard deviation of the distribution justifies that different individuals perceive CE differently. Based on Figure 7 (b), all the coefficients of CE are negative, which implies that acceleration noise reduces in CE. On average, the acceleration noise reduces by 65.2% ($exp(0.502) = 1.652$) in CE relative to that in TE. This shows when the participants are assisted with messages about the surrounding traffic (in this case the leader), they press the brake pedal more smoothly. In addition, since the drivers are informed, they make better and more timely decisions, and are at low risk in response to the sudden braking by the leader. For spacing, a statistically significant standard deviation of the distribution confirms the existence of inter-driver heterogeneity in this parameter. Moreover, similar to CE, all the coefficients of spacing are negative (Figure 7 (a)) implying that as spacing increases acceleration noise decreases. On average, a unit increase in spacing is expected to decrease the acceleration noise by 0.8%. This is because a larger spacing tends to give drivers more time to decide the response against the leader's action, which is reflected in low acceleration noise values.

Table 6 shows that heterogeneity in CE is partially related to age groups. A simulation based approach is adopted to understand the impact of heterogeneity, i.e., YWP on acceleration noise. The acceleration noise values in CE are calculated for 1000 values of n for both YWP = 1 and YWP = 0, i.e., others (middle-aged and old drivers), while keeping other explanatory variables constant. We simulate both the intra-age group heterogeneity (i.e., heterogeneity within YWP or within others) and the inter-age group heterogeneity (i.e., heterogeneity between YWP and others).

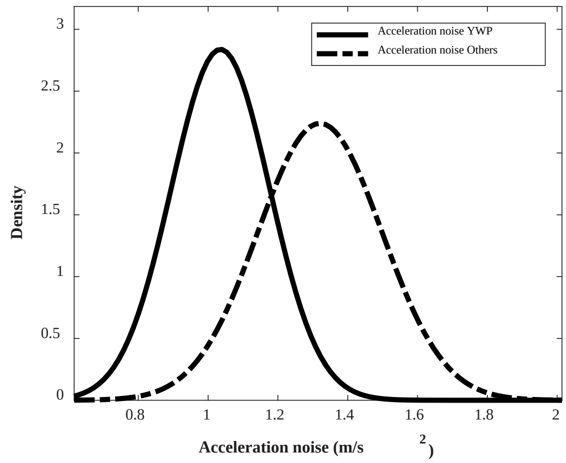

Figure 8 Distributions of acceleration noise for young drivers and others (middle-aged and old drivers).

Figure 8 illustrates the distributions of acceleration noise for young drivers and others. The mean acceleration noise is low for young drivers (1.03 m/s^2) as compared to middle-aged and old drivers (1.35 m/s^2) in CE. The acceleration noise for young drivers is mostly low although there is a possibility that for some values of n, acceleration noise for young drivers can be high as compared to old drivers. Overall, we can conclude that CE assists young drivers to make better and timely driving decisions, and thus are at low risk.

5.1.2 Effect of non-random parameters on acceleration noise

Low-speed section is negatively associated with acceleration noise, indicating that low-speed sections are more likely to result in low acceleration noise compared to high-speed sections. More specifically, acceleration noise decreases by 66.9 % (Table 5) in the low-speed car-following region as compared to the high-speed region. In addition, drivers with a university degree are more likely to have low acceleration noise relative to those without a university degree. More specifically, acceleration noise decreases by 10.9 % for drivers with a university degree as compared to their counterpart.

The coefficient of the interaction term of spacing and female is negative and statistically significant. Hence, the decrease in acceleration noise for maintaining a large spacing differs for males and females. More specifically, acceleration noise tends to decrease by 0.1 % for females relative to males when the spacing is fixed at its median value (more details in Section 6).

5.3 Understanding the factors behind an increase or a decrease in response time in CE as compared to TE

Section 5.1 describes the possibility of two classes of drivers: one with an increased response time and one with a decreased response time in CE relative to TE. The next natural step is to identify the factors behind the two different drivers' responses. Specifically, the aim is to obtain specific conditions derived from traffic flow parameters and human factors that can assist in classifying drivers into the two identified groups. To this aim, several popular classification algorithms are first assessed, and then the one with the best performance is employed to classify drivers.

Classification, in machine learning, is a supervised learning approach in which an algorithm learns from the data fed to it (i.e., training data) and then classifies a new observation based on the learning. Several classification algorithms are available in the literature. Refer to Kotsiantis et al. (2007; 2006) for a comprehensive review of important classification algorithms. This study assesses several major classification algorithms and then selects the one with a good accuracy and a high classification interpretability. The algorithms considered are decision tree, support vector machines, *k*-nearest neighbour, neural network, and ensemble algorithms (refer to Table 9 for the algorithms and their subtypes considered). These are well established machine learning algorithms, hence, to save space a description of these algorithms is omitted. The first criterion, i.e., accuracy depends on the type of data, classes, factors, etc. This study calculates and compares the four-fold cross-validation accuracy (hereon validation accuracy) of the selected classification algorithms and chooses the one with the highest validation accuracy. The validation accuracy is defined as in Equation (19):

$$Validation\ accuracy = 100\ \times (1 - \frac{1}{4}\sum_{i=1}^{4} MSE_i) \tag{19}$$

where MSE_i represents the mean squared error of i^{th} cross-validation dataset between the observed class and the predicted class of drivers in the i^{th} dataset. Note that all the selected classification algorithms are trained and validated using the inbuilt classification learner app in MATLAB (MathWorks, 2017). Table 8 lists the features (also called explanatory or predictor variables) used to classify drivers with increased or decreased response time in CE. The features *speed ratio* and *spacing ratio* quantify the change in speed and spacing in CE as compared to TE, respectively. Other features are already explained in Table 2. In total, 8 features are considered at the beginning of the classification. Table 9 details the validation accuracies for each classification algorithm using a different number of features. The highest validation accuracy for most of the algorithms is achieved when 6 features are used and these features are speed ratio, spacing ratio, speed section, age, gender, and university. The algorithms with good accuracies (the first criterion) are decision tree (fine tree), support vector machines (Fine Gaussian SVM), *k*-nearest neighbour (Fine KNN), ensembles (boosted trees and random forest), and neural network. Moreover, the decision tree based classification has the best interpretability among all the algorithms (Kotsiantis et al. (2007; 2006)). Also, decision tree analysis has been applied previous in transportation literature (Haque et al., 2016; Mirchandani and Head, 2001). Hence, decision tree is employed in this study because of its high accuracy and interpretability.

Table 8 A description of features or predictors used in testing the classification algorithms.

Features	Definition/Coding
Speed ratio	Speed in CE divided by speed in TE
Spacing ratio	Spacing in CE divided by speed in TE
Speed section in CE Age Gender University Experience	Same as defined in Table 2

Licence type	
Response variable / Class	
Response time	Increased response time = 1; Decreased response time = 2

Using the decision tree classification, one can obtain the feature importance ranking (also called predictor/variable importance ranking). This ranking can assist in identifying the most important traffic flow parameters and the human factors contributing to a driver's decision to increase or decrease the response time in CE. In this study, the feature importance ranking based on decision tree is obtained as per the methodology described in MATLAB (Mathworks, 2017). The decreasing order of feature importance (Figure 9) is speed section, spacing ratio, speed ratio, gender, education, and age. To check the reliability of the ranking order obtained from the decision tree, we also examine the feature importance order generated from the random forest algorithm and the neural network. Overall, the ranking order is consistent across the three algorithms.

Figure 9 depicts that the traffic flow parameters influence more as compared to the human factors in deciding whether the driver will increase or decrease the response time in CE or not. The most important deciding factor is speed section, i.e., whether a driver is following the leader at a high-speed or at a low-speed car-following region. Moreover, the ranking shows gender is the most important feature among the human factors.

According to the final classification tree (Figure 10), in CE, if a driver is along a low-speed road section and if the spacing is greater than that in TE, then it is more likely that the driver's response time will increase in CE. This is consistent with the findings reported in the previous section. Moreover, when driving along the low-speed section, if the spacing is smaller and the speed is higher as compared to TE, then the response time is more likely to decreases in CE. Moving towards the left branch of the root node (Figure 10), i.e., the high-speed section, in CE's high-speed car-following region, if the speed is higher than that in TE the response time is more likely to decrease due to high risk. Furthermore, the human factors are more likely to cause an increase or a decrease in the response time only if the driver is on the high-speed section. Thus, if the drivers are along the high-speed section and driving at a lower speed as compared to TE, the response time is expected to decrease for females in CE. In addition, based on the decision tree analysis, no clear conclusion can be made on how university and age influence an increase or a decrease in the response time in CE.

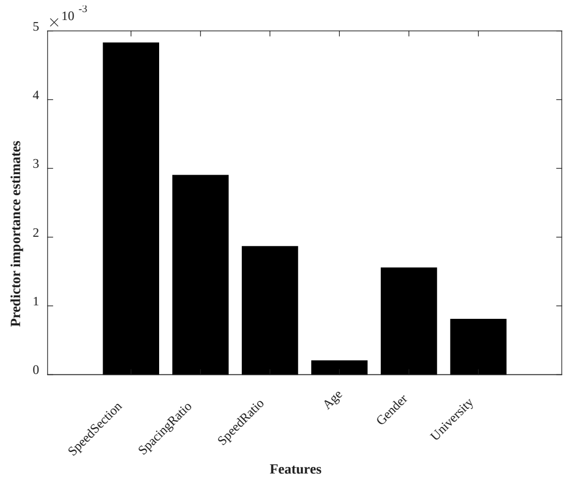

Figure 9 Feature importance ranking based on decision tree analysis.

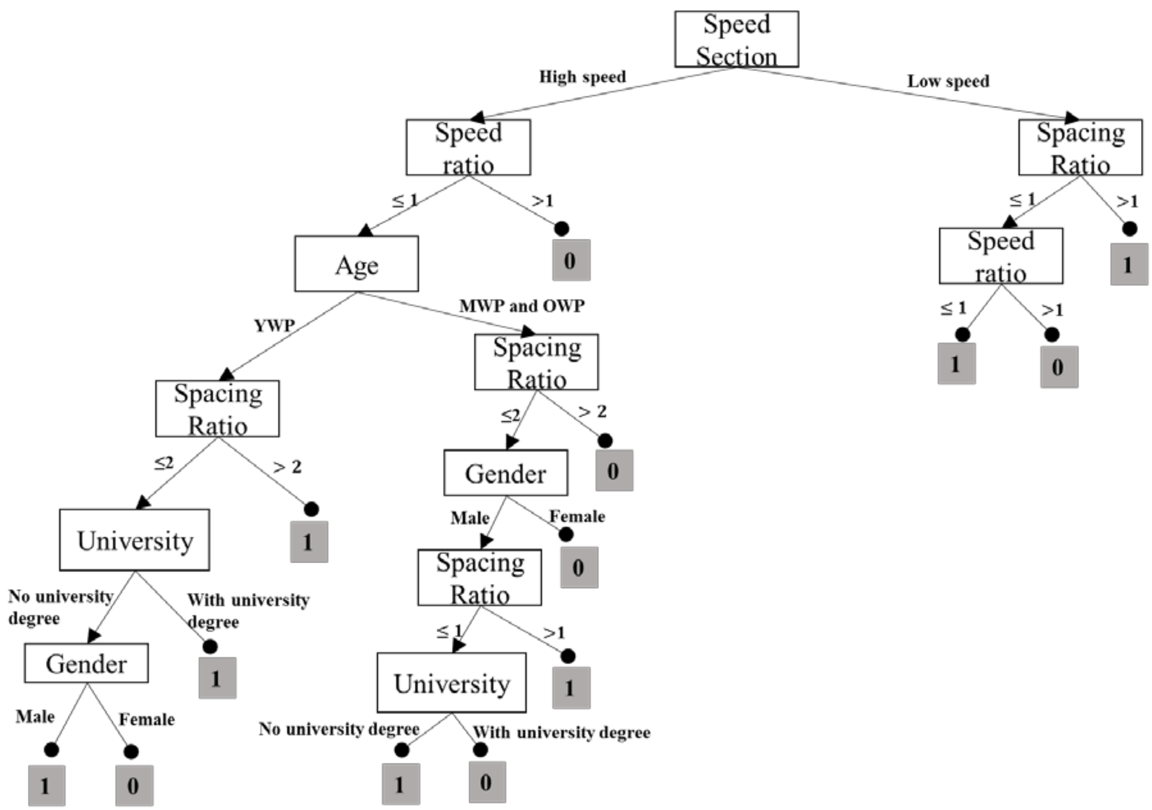

Figure 10 Decision tree of response time (1 and 0 represent increased and decreased response time, respectively, in CE as compared to TE).

Table 9 A comparison of different classification algorithm performance with different number of features.

Classifier	Sub type	Validation accuracy with all features (%)	Highest validation accuracy with 7 features (%)	Highest validation accuracy with 6 features (%)	Highest Validation accuracy with 5 features (%)
Decision Trees (Classification tree) (Breiman, 2017; Martinez and Martinez, 2007)	Fine Tree	68.6	68.6	72.5	66.7
	Medium Tree	68.6	68.6	72.5	66.7
	Coarse Tree	66.7	67.9	66.7	66.7
Discriminant Analysis (Martinez and Martinez, 2007; McLachlan, 2004)	Linear Discriminant	66.0	67.3	68.6	67.3
	Quadratic Discriminant	64.7	66.7	69.2	65.4
Support vector machines (SVM) (Duda et al., 2012; Hearst et al., 1998; Martinez and Martinez, 2007)	Linear SVM	70.5	70.5	67.3	66.0
	Quadratic SVM	64.1	68.6	69.2	67.3
	Cubic SVM	67.7	64.7	69.9	65.4
	Fine Gaussian SVM	66.0	68.6	71.2	67.3
	Medium Gaussian SVM	67.3	66.7	69.2	65.4
	Coarse Gaussian SVM	66.0	66.0	66.7	66.7
***k*-Nearest Neighbour (KNN)**(Duda et al., 2012; Martinez and Martinez, 2007)	Fine KNN	73.1	73.1	74.4	70.5
	Medium KNN	61.5	60.9	71.2	68.6
	Coarse KNN	66.0	66.0	66.0	66.0
	Cosine KNN	62.8	62.8	69.2	71.2
	Cubic KNN	59.0	59.0	70.5	68.6
	Weighted KNN	65.4	66.7	70.5	64.7
Ensembles (Liaw and Wiener, 2002; Martinez and Martinez, 2007)	Boosted Trees	68.6	73.1	72.4	67.9
	Random forest	74.4	73.1	73.7	66.7
Neural Network (Demuth and Beale, 1993; Duda et al., 2012)		76.9	77.0	78.7	78.7

6. DISCUSSION AND CONCLUSIONS

In this study, high quality vehicle trajectory data are collected for TE and CE using high-fidelity driving simulator experiments to understand the potential impacts of CE on driving behaviour (task performance and decision-making) in the car-following scenarios. The acceleration noise and the response time are the two driving behaviour parameters considered in this study. Random parameters linear regression model for the acceleration noise and random parameters lognormal survival AFT model for the response time are estimated. Below we discuss the findings from the modelling results. First, we discuss the influence of traffic flow parameters such as spacing and speed on the acceleration noise and the response time. Next, the focus shifts to the influence of human factors such as age, gender, and education on the acceleration noise and the response time. Finally, a detailed discussion is provided regarding the CE's impact on the two driving behaviour parameters.

6.1 Impact of traffic flow parameters on the acceleration noise and the response time

The random parameters linear regression model for the acceleration noise reveals that low-speed car-following region leads to the low acceleration noise relative to the high-speed car-following region during a safety-critical event in car-following scenarios. A possible reason behind such a behaviour is that drivers tend to press the brake pedal hard when facing the safety-critical event at high-speeds due to high risk of collision. This is consistent with the previous findings that faster drivers have high acceleration noise (Jones and Potts, 1962). A high acceleration noise also causes high discomfort (Babu and Pattnaik, 1997). In addition, as spacing increases it is likely that the acceleration noise decreases since drivers have more space and time to respond to safety-critical events. Note that the time headway is an established indicator of traffic safety (van Winsum and Heino, 1996; Vogel, 2003), and it is directly proportional to spacing. Thus, a large spacing leads to a large time headway thereby reducing the risk of rear-end collisions in safety-critical events during car-following.

Next, the random parameters duration model for the response time reveals that as the time headway increases response time also increases. This is because with a large time headway drivers do not need to rush to take an action after noticing the safety-critical event. A series of experiments by van Winsum and his colleagues (van Winsum, 1998; van Winsum and Brouwer, 1997; van Winsum and Heino, 1996) demonstrated that participants move the foot more slowly to the maximum brake position when following at large time headways, which confirms our argument. Findings from these experiments also suggested that such behaviour was due to the uncertainty about the required braking response at large time headways.

6.2 Impact of human factors on the acceleration noise and the response time

The acceleration noise model estimation results indicate that females depict a low acceleration noise thereby they are at less risk compared to males. These findings are in accordance with Saad et al. (2018). One of the major reasons for such behavioural differences during a safety-critical event is the difference in the risk-taking behaviour of males and females. According to previous studies, male drivers engage in risky driving behaviour more frequently than female drivers (Harris et al., 2006; Oltedal and Rundmo, 2006; Rhodes and Pivik, 2011). Next, drivers with a university degree demonstrate a low acceleration noise against those without a university degree, indicating that higher education may provide a strong cognitive component of safety and comfort attitudes. Another example that supports this argument is that higher-educated drivers are reported to be more likely to wear a seat-belt all the time as compared to less-educated drivers (Shinar et al., 2001).

As for the human factors considered in the response time model, we found that the old drivers take more time to respond to safety-critical events as compared to the young and the middle-aged drivers. Previous studies have also noticed slow responses of old drivers (Broen and Chiang, 1996; Edwards et al., 2003; Mehmood and Easa, 2009; Warshawsky-Livne and Shinar, 2002). A possible reason for such behaviour is that braking slows with age (Green, 2000). Interestingly, Summala and Koivisto (1990) demonstrated that old drivers compensate their degraded braking skills by driving slowly. We carried out the Wilcoxon rank sum test, which is a nonparametric test for equality of population medians of two independent samples (appropriate for unequal sample sizes), between the average speeds of young and old aged drivers in TE, and found a statistically significant difference at 90% confidence level between the two sets of average speeds (test statistic = 1878.5; p-value = 0.08). The average speed of old drivers is lower as compared to young drivers. This confirms the findings of Summala and Koivisto (1990). These findings also support the argument presented in previous studies that adaptation of a behaviour at the tactical level (adopting driving at slow-speed) occurs due to transient degradations in the operational performance (increased response time due to degraded braking skills) (van Winsum, 1998; van Winsum and Brouwer, 1997; van Winsum and Heino, 1996). Another possible reason for the old-aged drivers taking more time to respond is their driving experience. It is likely that old aged drivers have faced more safety-critical events as compared to the young and the middle-aged drivers, hence they can quickly judge dangerous situations and then intentionally delay their decision to brake if needed.

6.3 Impact of CE on the acceleration noise and the response time

As compared to TE, the acceleration noise decreases in CE, signifying that information assistance stabilises the braking behaviour of the follower in response to the leader's hard braking. Furthermore, the CE has a mixed effect on response time. For the majority of drivers, the response time increases when information assistance is provided. A possible reason for the increase in response time is the excess workload. In CE, drivers might have to allocate some time in comprehending the message in addition to perceiving and understanding the traffic in his/her field of view. However, for a few drivers, the response time decreases in CE. The deciding factors behind such different behaviours revealed by the decision tree analysis are speed section, spacing ratios, and speed ratios. At the low-speed section, the driver is likely to take more time to respond as compared to the high-speed section. A possible reason behind such behaviour is that the follower associates high risk with a hard braking at the high-speed car-following region due to less time to react relative to the hard braking at the low-speed car-following region. We can infer from our findings that by taking more time to respond in CE, the drivers are making better decisions that are reflected in lower acceleration noise (high stability) values in CE. Recent studies on CE have provided sufficient evidence underscoring that CE assists drivers in making better decisions when performing lane-changing and car-following tasks (Ali et al., 2019, 2018; Sharma et al., 2019a). Previous studies also report that collision avoidance warning decreased the accelerator release time and increased the time between initial brake press and maximum deceleration of undistracted drivers (Lee et al., 2002; McGehee et al., 2002).

The primary human factors behind an increase or decrease in the response time in CE as compared to TE are gender, university, and age. Particularly, for females, the response time decreases in CE as compared to TE, whereas, for males, the response time increases in CE against TE. Because females and males evaluate risk differently (Harris et al., 2006; Oltedal and Rundmo, 2006; Rhodes and Pivik, 2011), the gender difference in braking behaviour could be indicative of females having greater trust on the received message thereby recognising the potential risk of collision earlier than males. An early response is also an indicator of females giving themselves a larger margin of safety. Another possible reason could be that both males and females trust the CE, but males deliberately delay their decision to brake if there is no urgency or due to higher risk-taking tendency. Next, the results demonstrate that the young drivers take less time to react in CE as compared to the middle and the old aged drivers. Furthermore, in CE, young drivers demonstrate a low acceleration noise as compared to the middle and the old aged drivers. A possible reason for such behaviour is young drivers' more acceptance and trust, thereby more behavioural adaptation, towards CE relative to the middle and the old aged drivers. Green (2001) presented a review on the effects of age on driver performance when using telematics and reported that the young drivers require less visual demand and shorter time to read maps and enter destinations. This finding also implies the young drivers' higher ability to perceive, trust, and use tailored guidance. Previous studies also report that the old aged drivers are slower in response to a collision avoidance system (Kramer et al., 2007).

To conclude, in CE, the response time increases for the majority of drivers and acceleration noise decreases for every driver. These results indicate that, in CE, drivers are cautious, they have more time, and thus they make better decisions in safety-critical events during car-following. This conclusion underscores the findings of Sharma et al. (2019a) who conjuctured that the driving behaviour in CE is similar to the 'conscientious driving behaviour' (one of the five factors in the Big Five Factor model of personality (John and Srivastava, 1999)).

6.4 Practical implications of the research findings

This is the first comprehensive study to investigate CE's impact on the driver decision making and task performance in the car-following scenarios. The results of this study can be applied to evaluate the connected vehicle technology for different driver populations and traffic conditions. The results can provide ways of tailoring connected vehicle technology to the needs of different drivers. Dedicated short range communication based advanced information assistance system can be placed which can directly benefit the drivers with long response time, e.g., old aged drivers (Tang and Yip, 2010). As revealed in this study, CE will benefit the young drivers more as compared to the middle and the old aged drivers and, hence, strategies and policies shall be framed to maximise the positive effect of CE on the young drivers and to make this new technology more beneficial to the middle and the old drivers. As a first step, an introduction to connected vehicle technology can be integrated with the framework of graduated driver licensing program (Shope, 2007; Williams et al., 2012) to provide the necessary exposure of this technology to novice drivers.

6.5 Future research and limitations of the study

The present study is limited to understanding the impact of CE on driving behaviour in a safety-critical event during car-following. Future research can focus on understanding the impact of CE on driving behaviour in other aspects of car-following, e.g., following the leader, follower's response to the leader accelerating from standstill, etc. Other than acceleration noise, future studies can also consider car-following variables such as average spacing, average time headway, and standard deviation of spacing as operation variables and measures of task performance in safety-critical events. Moreover, future studies can analyse the impact of communication delay and communication loss on driving behaviour in the car-following scenario as this study is limited to only uninterrupted communication. Furthermore, it is important to understand the gender and the age differences in the workload experienced in CE, and it is important to validate the conclusions drawn from this study using field experiments. Such efforts are ongoing.

Acknowledgements

The authors would like to thank Mr Andrew Haines for programming the simulator experiment and Mr Yasir Ali and Dr Mohammad Saifuzzaman for their help in the experiment design and data collection. This research was partially funded by the Australian Research Council (ARC) through Dr Zuduo Zheng's Discovery Early Career Researcher Award (DECRA; DE160100449).

References

Adell, E., Várhelyi, A., dalla Fontana, M., 2011. The effects of a driver assistance system for safe speed and safe distance–a real-life field study. Transportation research part C: emerging technologies 19, 145–155.

Ali, Y., Haque, Md.M., Zheng, Z., Washington, S., Yildirimoglu, M., 2019. A hazard-based duration model to quantify the impact of connected driving environment on safety during mandatory lane-changing. Transportation Research Part C: Emerging Technologies 106, 113–131.

Ali, Y., Zheng, Z., Haque, Md.M., 2018. Connectivity's impact on mandatory lane-changing behaviour: Evidences from a driving simulator study. Transportation Research Part C: Emerging Technologies 93, 292–309.

Anastasopoulos Panagiotis Ch., Mannering Fred L., 2015. Analysis of Pavement Overlay and Replacement Performance Using Random Parameters Hazard-Based Duration Models. Journal of Infrastructure Systems 21, 04014024.

Anastasopoulos, P.Ch., Mannering, F.L., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accident Analysis & Prevention 41, 153–159.

Babu, Y.S., Pattnaik, S.B., 1997. Acceleration noise and level of service of urban roads - A case study. Journal of Advanced Transportation 31, 325–342.

Barrett, J.K., Farewell, V.T., Siannis, F., Tierney, J., Higgins, J.P.T., 2012. Two-stage meta-analysis of survival data from individual participants using percentile ratios. Statistics in Medicine 31, 4296–4308.

Behnood, A., Mannering, F., 2017. The effect of passengers on driver-injury severities in single-vehicle crashes: A random parameters heterogeneity-in-means approach. Analytic Methods in Accident Research 14, 41–53.

Bella, F., Silvestri, M., 2017. Effects of directional auditory and visual warnings at intersections on reaction times and speed reduction times. Transportation Research Part F: Traffic Psychology and Behaviour 51, 88–102.

Belz, N.P., Aultman-Hall, L., 2011. Analyzing the Effect of Driver Age on Operating Speed and Acceleration Noise: On-Board Second-by-Second Driving Data. Transportation Research Record 2265, 184–191.

Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transportation Research Part B: Methodological 37, 837–855.

Breiman, L., 2017. Classification and Regression Trees. Routledge.

Broen, N.L., Chiang, D.P., 1996. Braking Response Times for 100 Drivers in the Avoidance of an Unexpected Obstacle as Measured in a Driving Simulator. Proceedings of the Human Factors and Ergonomics Society Annual Meeting 40, 900–904.

Caird, J.K., Willness, C.R., Steel, P., Scialfa, C., 2008. A meta-analysis of the effects of cell phones on driver performance. Accident Analysis & Prevention 40, 1282–1293.

Chang, C.Y., Wei, C.H., 2013. Driving simulation for analyzing the safety and fuel saving effects of a connected bus system on freeways, in: 2013 IEEE Intelligent Vehicles Symposium (IV). Presented at the 2013 IEEE Intelligent Vehicles Symposium (IV), pp. 618–623.

Choudhary, P., Velaga, N.R., 2017. Modelling driver distraction effects due to mobile phone use on reaction time. Transportation Research Part C: Emerging Technologies 77, 351–365.
Demuth, H., Beale, M., 1993. Neural Network Toolbox For Use with Matlab--User'S Guide Verion 3.0.
Drew, D.R., Dudek, C.L., Keese, C.J., 1967. Freeway level of service as described by an energy- acceleration noise model. Highway Research Record.
Duda, R.O., Hart, P.E., Stork, D.G., 2012. Pattern classification. John Wiley & Sons.
Edwards, C.J., Creaser, J.I., Caird, J.K., Lamsdale, A.M., Chisholm, S.L., 2003. Older and younger driver performance at complex intersections: implications for using perception-response time and driving simulation. Presented at the Driving Assessment 2003: The Second International Driving Symposium on Human Factors in Driver Assessment, Training and Vehicle Design.
Fairclough, S.H., May, A.J., Carter, C., 1997. The effect of time headway feedback on following behaviour. Accident Analysis & Prevention 29, 387–397.
Farah, H., Koutsopoulos, H.N., Saifuzzaman, M., Kölbl, R., Fuchs, S., Bankosegger, D., 2012. Evaluation of the effect of cooperative infrastructure-to-vehicle systems on driver behavior. Transportation Research Part C: Emerging Technologies 21, 42–56.
Fu, C., Zhang, Y., Bie, Y., Hu, L., 2016. Comparative analysis of driver's brake perception-reaction time at signalized intersections with and without countdown timer using parametric duration models. Accident Analysis & Prevention, Traffic Safety in China: Challenges and Countermeasures 95, 448–460.
Ghadiri, S., Prasetijo, J., Sadullah, A., Hoseinpour, M., Sahranavard, S., 2013. Intelligent speed adaptation: Preliminary results of on-road study in Penang, Malaysia. IATSS research 36, 106–114.
Green, M., 2000. "How Long Does It Take to Stop?" Methodological Analysis of Driver Perception-Brake Times. Transportation Human Factors 2, 195–216.
Green, P., 2001. Variations in task performance between younger and older drivers: UMTRI research on telematics. Presented at the association for the advancement of automotive medicine conference on aging and driving.
Haque, M.M., Ohlhauser, A.D., Washington, S., Boyle, L.N., 2016. Decisions and actions of distracted drivers at the onset of yellow lights. Accident Analysis & Prevention 96, 290–299.
Haque, M.M., Washington, S., 2014. A parametric duration model of the reaction times of drivers distracted by mobile phone conversations. Accident Analysis & Prevention 62, 42–53.
Harris, C.R., Jenkins, M., Glaser, D., 2006. Gender differences in risk assessment: Why do women take fewer risks than men? Judgment and Decision Making 1, 48–63.
Hasan, S., Mesa-Arango, R., Ukkusuri, S., 2013. A random-parameter hazard-based model to understand household evacuation timing behavior. Transportation Research Part C: Emerging Technologies, Selected papers from the Seventh Triennial Symposium on Transportation Analysis (TRISTAN VII) 27, 108–116.
Hearst, M.A., Dumais, S.T., Osuna, E., Platt, J., Scholkopf, B., 1998. Support vector machines. IEEE Intelligent Systems and their applications 13, 18–28.
Hensher, D.A., Rose, J.M., Greene, W.H., 2015. Applied Choice Analysis, 2nd ed. Cambridge University Press.
Herman, R., Montroll, E.W., Potts, R.B., Rothery, R.W., 1959. Traffic Dynamics: Analysis of Stability in Car Following. Operations Research 7, 86–106.
.idcommunity, 2016. Service age groups | Australia | Community profile. URL https://profile.id.com.au/australia/service-age-groups (accessed 4.24.18).
John, O.P., Srivastava, S., 1999. The Big Five Trait taxonomy: History, measurement, and theoretical perspectives. in: Handbook of Personality: Theory and Research, 2nd Ed. Guilford Press, New York, NY, US, pp. 102–138.
Jones, T.R., Potts, R.B., 1962. The Measurement of Acceleration Noise-A Traffic Parameter. Operations Research 10, 745–763.
Kim, J.-T., Courage, K., Washburn, S., Bonyani, G., 2003. Framework for Investigation of Level-of-Service Criteria and Thresholds on Rural Freeways. Transportation Research Record: Journal of the Transportation Research Board 1852, 239–245.
Ko, J., Guensler, R., Hunter, M., 2010. Analysis of effects of driver/vehicle characteristics on acceleration noise using GPS-equipped vehicles. Transportation Research Part F: Traffic Psychology and Behaviour 13, 21–31.
Ko, J., Guensler, R., Hunter, M., 2006. Variability in Traffic Flow Quality Experienced by Drivers: Evidence from Instrumented Vehicles. Transportation Research Record: Journal of the Transportation Research Board 1988, 1–9.

Koppa, R.J., 2000. Revised Monograph on Traffic Flow Theory: Chapter 3 Human Factors (Technical Report), Transportation Research Board (TRB) Special Report 165, Traffic Flow Theory. US Department of Transportation Federal Highway Administration.
Kotsiantis, S.B., Zaharakis, I., Pintelas, P., 2007. Supervised machine learning: A review of classification techniques. Emerging artificial intelligence applications in computer engineering 160, 3–24.
Kotsiantis, S.B., Zaharakis, I.D., Pintelas, P.E., 2006. Machine learning: a review of classification and combining techniques. Artif Intell Rev 26, 159–190.
Kramer, A.F., Cassavaugh, N., Horrey, W.J., Becic, E., Mayhugh, J.L., 2007. Influence of age and proximity warning devices on collision avoidance in simulated driving. Hum Factors 49, 935–949.
Lee, J.D., McGehee, D.V., Brown, T.L., Reyes, M.L., 2002. Collision warning timing, driver distraction, and driver response to imminent rear-end collisions in a high-fidelity driving simulator. Human Factors: The Journal of the Human Factors and Ergonomics Society 44, 314–334.
Lu, Q., Xin, C., 2018. Pavement Rehabilitation Policy for Reduced Life-Cycle Cost and Environmental Impact Based on Multiple Pavement Performance Measures (Final Report No. 69A3551747119 (Phase II)). U.S. Department of Transportation.
Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R news 2, 18–22.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Analytic Methods in Accident Research 11, 1–16.
Martinez, W.L., Martinez, A.R., 2007. Computational statistics handbook with MATLAB. Chapman and Hall/CRC.
MathWorks, 2017. Train Classification Models in Classification Learner App - MATLAB & Simulink - MathWorks Australia. URL https://au.mathworks.com/help/stats/train-classification-models-in-classification-learner-app.html (accessed 1.4.19).
Mathworks, 2017. Estimates of predictor importance - MATLAB - MathWorks Australia. URL https://au.mathworks.com/help/stats/compactclassificationtree.predictorimportance.html (accessed 1.4.19).
May, A., Carter, C., Smith, F., Fairclough, S., 1995. An evaluation of an in-vehicle headway feedback system with a visual and auditory interface. Presented at the Design of the Driver Interface, IEE Colloquium on, IET, pp. 5–1.
McGehee, D., Brown, T., Lee, J., Wilson, T., 2002. Effect of Warning Timing on Collision Avoidance Behavior in a Stationary Lead Vehicle Scenario. Transportation Research Record: Journal of the Transportation Research Board 1803, 1–6.
McLachlan, G., 2004. Discriminant analysis and statistical pattern recognition. John Wiley & Sons.
Mehmood, A., Easa, S.M., 2009. Modeling reaction time in car-following behaviour based on human factors. International Journal of Applied Science, Engineering and Technology 5, 93–101.
Michiels, S., Piedbois, P., Burdett, S., Syz, N., Stewart, L., Pignon, J.-P., 2005. Meta-analysis when only the median survival times are known: A comparison with individual patient data results. International Journal of Technology Assessment in Health Care 21, 119–125.
Mirchandani, P., Head, L., 2001. A real-time traffic signal control system: architecture, algorithms, and analysis. Transportation Research Part C: Emerging Technologies 9, 415–432.
Moon, S., Moon, I., Yi, K., 2009. Design, tuning, and evaluation of a full-range adaptive cruise control system with collision avoidance. Control Engineering Practice 17, 442–455.
Ni, D., Li, J., Andrews, S., Wang, H., 2011. A Methodology to Estimate Capacity Impact due to Connected Vehicle Technology. International Journal of Vehicular Technology 2012, e502432.
Olia, A., Abdelgawad, H., Abdulhai, B., Razavi, S.N., 2016. Assessing the potential impacts of connected vehicles: mobility, environmental, and safety perspectives. Journal of Intelligent Transportation Systems 20, 229–243.
Oltedal, S., Rundmo, T., 2006. The effects of personality and gender on risky driving behaviour and accident involvement. Safety Science 44, 621–628.
Ossen, S., Hoogendoorn, S., 2007. Driver heterogeneity in car following and its impact on modeling traffic dynamics. Transportation Research Record, 95–103.
Ossen, S., Hoogendoorn, S.P., 2011. Heterogeneity in car-following behavior: Theory and empirics. Transportation Research Part C: Emerging Technologies, Emerging theories in traffic and transportation and methods for transportation planning and operations 19, 182–195.
Rakha, H., El-Shawarby, I., Setti, J.R., 2007. Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger. IEEE Transactions on Intelligent Transportation Systems 8, 630–640.
Rhodes, N., Pivik, K., 2011. Age and gender differences in risky driving: The roles of positive affect and risk perception. Accident Analysis & Prevention 43, 923–931.

Rusli, R., Haque, Md.M., Afghari, A.P., King, M., 2018. Applying a random parameters Negative Binomial Lindley model to examine multi-vehicle crashes along rural mountainous highways in Malaysia. Accident Analysis & Prevention 119, 80–90.
Saad, M., Abdel-Aty, M., Lee, J., 2018. Analysis of driving behavior at expressway toll plazas. Transportation Research Part F: Traffic Psychology and Behaviour.
Saifuzzaman, M., Haque, Md. M., Zheng, Z., Washington, S., 2015. Impact of mobile phone use on car-following behaviour of young drivers. Accident Analysis & Prevention 82, 10–19.
Saifuzzaman, M., Zheng, Z., 2014. Incorporating human-factors in car-following models: A review of recent developments and research needs. Transportation Research Part C: Emerging Technologies 48, 379–403.
Schweitzer, N., Apter, Y., Ben-david, G., Liebermann, D.G., Parush, A., 1995. A field study on braking responses during driving. II. Minimum driver braking times. Ergonomics 38, 1903–1910.
Scott, J.J., Gray, R., 2008. A comparison of tactile, visual, and auditory warnings for rear-end collision prevention in simulated driving. Hum Factors 50, 264–275.
Sharma, A., Ali, Y., Saifuzzaman, M., Zheng, Z., Haque, M.M., 2017. Human Factors in Modelling Mixed Traffic of Traditional, Connected, and Automated Vehicles, in: Advances in Human Factors in Simulation and Modeling, Advances in Intelligent Systems and Computing. Springer, Cham, pp. 262–273.
Sharma, A., Zheng, Z., Bhaskar, A., 2018a. A pattern recognition algorithm for assessing trajectory completeness. Transportation Research Part C: Emerging Technologies 96, 432–457.
Sharma, A., Zheng, Z., Bhaskar, A., 2018b. Is more always better? The impact of vehicular trajectory completeness on car-following model calibration and validation. Transportation Research Part B: Methodological 120C, 49–75.
Sharma, A., Zheng, Z., Bhaskar, A., Haque, Md.M., 2019a. Modelling car-following behaviour of connected vehicles with a focus on driver compliance. Transportation Research Part B: Methodological 126, 256–279.
Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, Md.M., 2019b. Estimating and comparing response times in traditional and connected environments. Transportation Research Record 0361198119837964.
Shinar, D., Schechtman, E., Compton, R., 2001. Self-reports of safe driving behaviors in relationship to sex, age, education and income in the US adult driving population. Accid Anal Prev 33, 111–116.
Shope, J.T., 2007. Graduated driver licensing: Review of evaluation results since 2002. Journal of Safety Research, Novice Teen Driving: GDL and Beyond – Research Foundations for Policy and Practice Symposium 38, 165–175.
Son, H. "Daniel," Kweon, Y.-J., Park, B. "Brian," 2011. Development of crash prediction models with individual vehicular data. Transportation Research Part C: Emerging Technologies 19, 1353–1363.

Spruance, S.L., Reid, J.E., Grace, M., Samore, M., 2004. Hazard Ratio in Clinical Trials. Antimicrobial Agents and Chemotherapy 48, 2787–2792.
Summala, H., Koivisto, I., 1990. Unalerted drivers' brake reaction times: older drivers compensate their slower reactions by driving more slowly. Driving behaviour in a social context 680–683.
Sun, Jie, Zheng, Z., Sun, Jian, 2018. Stability analysis methods and their applicability to car-following models in conventional and connected environments. Transportation Research Part B: Methodological 109, 212–237.
Talebpour, A., Mahmassani, H.S., 2016. Influence of connected and autonomous vehicles on traffic flow stability and throughput. Transportation Research Part C: Emerging Technologies 71, 143–163.
Talebpour, A., Mahmassani, H.S., Bustamante, F.E., 2016. Modeling driver behavior in a connected environment: integrated microscopic simulation of traffic and mobile wireless telecommunication systems. Presented at the Transportation Research Board 95th Annual Meeting.
Talebpour, A., Mahmassani, H.S., Hamdar, S.H., 2015. Modeling Lane-Changing Behavior in a Connected Environment: A Game Theory Approach. Transportation Research Procedia, 21st International Symposium on Transportation and Traffic Theory Kobe, Japan, 5-7 August, 2015 7, 420–440.
Tang, A., Yip, A., 2010. Collision avoidance timing analysis of DSRC-based vehicles. Accident Analysis & Prevention 42, 182–195.
Tavassoli Hojati, A., Ferreira, L., Washington, S., Charles, P., 2013. Hazard based models for freeway traffic incident duration. Accident Analysis & Prevention 52, 171–181.
Taylor, M.A.P., Woolley, J.E., Zito, R., 2000. Integration of the global positioning system and geographical information systems for traffic congestion studies. Transportation Research Part C: Emerging Technologies 8, 257–285.

Treiber, M., Kesting, A., Helbing, D., 2007. Influence of Reaction Times and Anticipation on Stability of Vehicular Traffic Flow. Transportation Research Record 1999, 23–29.

Uno, N., Iida, Y., Itsubo, S., Yasuhara, S., 2002. A microscopic analysis of traffic conflict caused by lane-changing vehicle at weaving section. Presented at the Proc. 9th meeting of Euro Working Group on Transportation (EWGT), Bari, Italy.

van Winsum, W., 1998. Preferred time headway in car-following and individual differences in perceptual-motor skills. Percept Mot Skills 87, 863–873.

van Winsum, W., Brouwer, W., 1997. Time headway in car following and operational performance during unexpected braking. Percept Mot Skills 84, 1247–1257.

van Winsum, W., Heino, A., 1996. Choice of time-headway in car-following and the role of time-to-collision information in braking. Ergonomics 39, 579–592.

Vogel, K., 2003. A comparison of headway and time to collision as safety indicators. Accident Analysis & Prevention 35, 427–433.

Warshawsky-Livne, L., Shinar, D., 2002. Effects of uncertainty, transmission type, driver age and gender on brake reaction and movement time. Journal of Safety Research 33, 117–128.

Washington, S.P., Karlaftis, M.G., Mannering, F., 2010. Statistical and econometric methods for transportation data analysis. Chapman and Hall/CRC.

Williams, A.F., Tefft, B.C., Grabowski, J.G., 2012. Graduated Driver Licensing Research, 2010-Present. Journal of Safety Research 43, 195–203.

Wu, Y., Abdel-Aty, M., Park, J., Zhu, J., 2018. Effects of crash warning systems on rear-end crash avoidance behavior under fog conditions. Transportation Research Part C: Emerging Technologies 95, 481–492.

Zheng, Z., 2014. Recent developments and research needs in modeling lane changing. Transportation Research Part B: Methodological 60, 16–32.

Zheng, Z., Washington, S., Hyland, P., Sloan, K., Liu, Y., 2016. Preference heterogeneity in mode choice based on a nationwide survey with a focus on urban rail. Transportation Research Part A: Policy and Practice 91, 178–194.