Delft University of Technology

A Bayesian correlated grouped random parameters duration model with heterogeneity in the means for understanding braking behaviour in a connected environment

Ali, Yasir; Haque, Md Mazharul; Zheng, Zuduo; Afghari, Amir Pooyan

DOI
10.1016/j.amar.2022.100221

Publication date
2022

Document Version
Final published version

Published in
Analytic Methods in Accident Research

Citation (APA)
Ali, Y., Haque, M. M., Zheng, Z., & Afghari, A. P. (2022). A Bayesian correlated grouped random parameters duration model with heterogeneity in the means for understanding braking behaviour in a connected environment. *Analytic Methods in Accident Research*, *35*, Article 100221. https://doi.org/10.1016/j.amar.2022.100221

Important note
To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright
Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy
Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

This work is downloaded from Delft University of Technology.
For technical reasons the number of authors shown on this cover page is limited to a maximum of 10.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' - Taverne project

https://www.openaccess.nl/en/you-share-we-take-care

Otherwise as indicated in the copyright section: the publisher is the copyright holder of this work and the author uses the Dutch legislation to make this work public.

Analytic Methods in Accident Research 35 (2022) 100221

Contents lists available at ScienceDirect

Analytic Methods in Accident Research

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

A Bayesian correlated grouped random parameters duration model with heterogeneity in the means for understanding braking behaviour in a connected environment

Yasir Ali [a], Md. Mazharul Haque [a,*], Zuduo Zheng [b], Amir Pooyan Afghari [c]

[a] Queensland University of Technology, School of Civil & Environment Engineering, Faculty of Engineering, Brisbane, Australia
[b] The University of Queensland, School of Civil Engineering, Faculty of Engineering, Architecture, and Information Technology, Brisbane, Australia
[c] Delft University of Technology, Safety and Security Science Section, Faculty of Technology, Policy and Management, Delft, Netherlands

ARTICLE INFO

Article history:
Received 29 January 2022
Received in revised form 10 April 2022
Accepted 10 April 2022
Available online 18 April 2022

Keywords:
Connected environment
Braking
Driving behaviour
Hierarchical
Random parameters
Decision tree

ABSTRACT

Driver's response to a pedestrian crossing requires braking, whereby both excess and inadequate braking is directly associated with crash risk. The highly anticipated connected environment aims to increase drivers' situational awareness by providing advanced information and assisting them during critical driving tasks such as braking. Focussing on this crucial behaviour and combined with the promise of a connected environment, the objective of this study is to examine the braking behaviour of drivers in response to a pedestrian at a zebra crossing in a connected environment. Seventy-eight participants from diverse backgrounds performed this driving task in the CARRS-Q Advanced Driving Simulator in two randomised driving scenarios: a baseline scenario (without driving aids) and a connected environment (with driving aids) scenario. A Weibull accelerated failure time duration modelling approach is adopted to model the braking behaviour of drivers. In particular, this duration model is specified to capture the panel nature of the data and unobserved heterogeneity through correlated grouped random parameters with heterogeneity-in-the-means in the Bayesian framework. Results indicate that, for most drivers in the connected environment, it takes longer to reduce their speed with less speed variation and a larger safety margin. In addition, a decision tree analysis for the braking time suggests that for older drivers, when the distance to the zebra crossing is larger in the connected environment than that in the baseline scenario, braking time is likely to increase. The model also reveals that the braking time of female drivers is longer in the connected environment compared to that of male drivers. Overall, the connected environment is associated with increased braking time by providing advanced information, giving drivers additional time to smoothly reduce their speed in response to a pedestrian at a zebra crossing, and ultimately making the vehicle–pedestrian interaction safer.

© 2022 Elsevier Ltd. All rights reserved.

* Corresponding author.
E-mail addresses: y2.ali@qut.edu.au (Y. Ali), m1.haque@qut.edu.au (M.M. Haque), zuduo.zheng@uq.edu.au (Z. Zheng), a.p.afghari-1@tudelft.nl (A.P. Afghari).

https://doi.org/10.1016/j.amar.2022.100221
2213-6657/© 2022 Elsevier Ltd. All rights reserved.

1. Introduction

With the technological advancements in communication and sensing technologies, connected and automated vehicles are soon to become a reality, as their deployment appears to be just on the horizon. Thus, unsurprisingly, recent research related to these vehicles has received significant attention from researchers. In particular, the information provided by a connected environment using vehicle-to-vehicle communication and vehicle-to-infrastructure communication has shown promise in solving various transport issues, such as improving safety, suppressing congestion, and minimising environmental impact. This study contributes to understanding the safety of a novel connected environment.

A connected environment is expected to provide event-based and advanced driving aids for assisting drivers in various driving tasks. Surrounding traffic information in a connected environment can help in car-following (Sharma et al., 2020) and lane-changing (Ali et al., 2020a) manoeuvres. Similarly, advanced information through a connected environment is expected to improve situational awareness among drivers and generate stimulus well before an event occurs, which can enhance safety. For instance, Ali et al. (2021b) reported that drivers in a connected environment make safer decisions at the onset of yellow light by deciding to stop before the stop line. Research on the effects of a connected environment on safety has been the focus of a large body of literature in recent years (see more details in the next section). Succinctly, much of the literature has focussed on connected environment's impact using numerical simulations, which lack the human factor that is considered critical in the safety evaluation of a connected environment (Sharma et al., 2017). Also, these simulation-based studies showed aggregated or macroscopic benefits of a connected environment, whereas studies demonstrating how an individual driver is affected by a connected environment are missing, which is critical for the success of this novel environment. As such, this study aims to investigate the effects of a connected environment on driving behaviour at a microscopic (or an individual driver) level using actual trajectory data, which contain human factor information.

Driving behaviour on urban streets differs significantly from that on motorways because of complex traffic interactions in an urban road traffic environment. This study focusses on microscopic driving behaviour on urban streets. While encountering traffic events on urban streets, two important aspects of driving behaviour are reaction (or response) time and braking behaviour. The former aspect has received significant attention in traffic safety literature. For instance, the effects of distraction (specifically caused by mobile phones) are frequently measured using reaction time (Hancock et al., 2003, Törnros and Bolling, 2006, Caird et al., 2008, Just et al., 2008, Ishigami and Klein, 2009, Haque and Washington, 2014). Similarly, the performance of different designs of intersections with dynamic use of exit lanes for the left turn has been evaluated using reaction time (Zhao et al., 2015). It has also been used to evaluate the effects of auditory alerts from in-vehicle information systems (Wiese and Lee, 2004) as well as in a connected environment (Sharma et al., 2019, Ali et al., 2020b). Comparatively, less attention has been paid to braking behaviour despite its importance in characterising driving behaviour. In general, braking behaviour is considered crucial because of its direct relationship with crash risk, as improper and abrupt braking is often associated with an increased likelihood of engaging in rear-end collisions. Thus, this study analyses braking behaviour in a connected environment.

More specifically, a detailed synthesis of the relevant literature (see Section 2) revealed a number of noteworthy observations along this research direction. First, although analysing braking behaviour has remained the objective of some studies (e.g., analysing the effects of distraction and warning signs), it has been rarely studied in the context of stopping in response to a routine traffic event, e.g., pedestrians at a zebra crossing in a connected environment. Second, for the impact of a connected environment, in general, and in-vehicle information, in particular, disagreements on braking behaviour have been noted in the literature (see the next section for more details). Third, our understanding remains elusive on whether drivers brake homogeneously in a connected environment when they interact with a pedestrian at a zebra crossing. For instance, the advanced information provided by a connected environment could trigger an early response, leading to smooth braking behaviour. In contrast, drivers may use this information to apply brake late and abruptly. An in-depth understanding of braking behaviour is not only critical for minimising rear-end collisions but also for maximising the impact of a connected environment for improving driving behaviour on urban streets. Finally, an important research question is how braking behaviour in a connected environment varies across driver demographics. Answers to this research question may trigger the need to tailor a connected environment design for different driver groups.

By focussing on these research gaps, the objective of this study is to examine the braking behaviour of drivers in a connected environment. The braking behaviour is studied and modelled when drivers receive advanced information about a pedestrian at a zebra crossing. A Bayesian random parameters duration modelling approach is applied to model the braking behaviour in a connected environment.

The contribution of this study is threefold. First, as one of the first studies that is focussed on the braking behaviour in a connected environment, this study presents a Bayesian random parameters duration modelling approach that provides an in-depth understanding of drivers braking behaviour in a routine driving task, i.e., interacting with a pedestrian at a zebra crossing. The findings of this study can help in identifying a group of risky drivers and suggesting suitable countermeasures for them. Second, by leveraging the capabilities of advanced econometric modelling, unobserved heterogeneity in braking behaviour is captured. Third, a decision tree algorithm is employed to further trace the source of underlying heterogeneity, revealing more insights about the differential braking behaviour in a connected environment corresponding to different driver demographics and driving behaviour.

The rest of the paper is organised as follows. Section 3 explains the design of the experimental setup, including an interaction with a pedestrian, design of driving aids, and data pre-processing. Section 4 describes the Bayesian random parameters duration model development process. While Section 5 presents results including a descriptive analysis of braking profiles, Bayesian random parameters model, and decision tree analysis, Section 6 discusses these results with respect to driver demographics. Finally, Section 7 summarises the main findings and outlines some future research directions.

2. Literature review

This section is divided into two parts, whereby the first part describes studies related to a connected environment and the second part deals with a review of braking behaviour studies.

2.1. Connected environment's impact on safety

Using a numerical simulation framework, Olia et al. (2016) reported that relative to a traditional environment, a connected environment enhances safety and traffic flow efficiency with reduced gas emissions at a network level. Similarly, Lee and Park (2012) found that basic safety messages obtained from the safety pilot model deployment project showed improved driving behaviour and enhanced intersection safety. Similar conclusions have been reported by several other studies (e.g., Park et al. (2011), McGurrin et al. (2012), Zeng et al. (2012), Rahman and Abdel-Aty (2018)). For instance, a study on connected vehicle platooning reported that a connected vehicular environment increases safety measured in terms of safety surrogates (Rahman and Abdel-Aty, 2018). In another study on a connected environment where drivers received assistance for merging, it was found that drivers safely merged to a freeway when they were assisted (Ahmed et al., 2017). Similar findings were also reported by Hayat et al. (2014).

Using the data from the Connected Vehicle Safety Pilot Model Deployment Program of the University of Michigan, Ghanipoor Machiani et al. (2017) developed a logistic model, which can be used to activate smart curve speed warnings in a connected environment. The same dataset has been used in another study that evaluated a real-time collision warning system based on time-to-collision and reported safety benefits of a connected environment (Zhang et al., 2017). In another study, data from the in-depth crash investigations by the Centre for Automotive Safety Research in South Australia were used, and crashes were reported to reduce significantly in a connected environment (Doecke et al., 2015).

The aforementioned studies were conducted either using numerical simulation or field testbed. Although studies on numerical simulations confirmed the positive effects of a connected environment on safety at a macro level (or network level), these findings are preliminary and lack an important component, i.e., human factor, which is considered crucial for analysing safety at a microscopic level (Sharma et al., 2017). This issue is somewhat addressed in studies that use real testbed data. However, none of these studies focusses on examining and understanding driving behaviour in a connected environment at a microscopic (or an individual driver) level using actual trajectory data in an urban environment. This research gap motivates the present study.

2.2. A review of braking behaviour studies

Braking behaviour is often characterised by brake response time and the amount of braking, which has been studied for different driving tasks and conditions such as approaching a signalised intersection (Zöller et al., 2019, Ali et al., 2021b), distracted drivers approaching a pedestrian crossing (Haque and Washington, 2015), examining the effects of alcohol content on driving behaviour (Yadav and Velaga, 2019), automated emergency braking (Suzuki et al., 2019), and brake assistance for intelligent vehicles (McCall and Trivedi, 2007). Some studies used braking behaviour as an indicator of increased crash risk in distracted driving (e.g., Consiglio et al. (2003), Al-Darrab et al. (2009), Hancock et al. (2003), Harbluk et al. (2007)). Similarly, braking behaviour is also measured to analyse the effects of warnings provided by advanced driving assistance systems. Lerner et al. (2011), for example, observed a faster brake reaction when drivers received both audio and visual warnings compared to driving without warnings. Contrasting findings were reported by Bella and Silvestri (2017) as their study observed smooth braking when drivers received directional auditory and visual warnings. Similar findings have been reported in another study (Wan et al., 2016), where drivers were found to perform gradual braking when they received advance information of a traffic event. Complementing these studies, a recent study (Ali et al., 2020a) on a connected environment found that drivers have lower deceleration rates during mandatory lane-changing manoeuvres compared to discretionary lane-changing manoeuvres. In another study, a smooth braking behaviour is observed in a connected environment when drivers faced a failed lane-changing attempts on motorways (Ali et al., 2021a).

To summarise, braking behaviour has been studied in several other contexts, e.g., driver distraction, failed lane-changing attempts, etc. However, our understanding remains elusive on how driver braking behaviour will be changed when interacting with a pedestrian in a connected environment. Further, although a recent study has analysed braking behaviour in a connected environment (Ali et al., 2021a), their study focussed on failed lane-changing attempts on motorways, which is different from vehicle-pedestrian interactions in an urban driving condition. As such, our study aims to fill this research gap.

3. Design of experiment and data collection

A driving simulator experiment was designed to collect high-quality vehicle trajectory data and examine driving behaviour. To this end, an advanced driving simulator at the Centre for Accident Research and Road Safety-Queensland (CARRS-Q) of the Queensland University of Technology was utilised, and the data were collected in a controlled driving environment. Participants drove the simulator in two randomised driving scenarios: baseline driving (without advance information; the same as a traditional driving environment) and connected environment (with advanced information about a pedestrian walking from a sidewalk to a zebra crossing). This study considers the baseline driving as the '*default*' driving scenario with which the driving performance with advanced information is compared. Note that the experiment design is a within-subject experiment.

3.1. Specifications of the CARRS-Q advanced driving simulator

The CARRS-Q Advanced Driving Simulator (Fig. 1(a)) consisted of a fully working Holden Commodore car, which is fitted with three large size projectors providing a 180° field of view. In addition, the rear and wing mirrors of the simulator car were replaced by liquid crystal display (LCD) screens that provide a high-quality photorealistic view of surrounding traffic. The simulator car is rested on a flexible rotating base, providing six-degrees-of-freedom, and mimicking real driving features such as acceleration, deceleration, braking cornering, and road surface friction. The simulator was also capable of producing simulated engine noises, road interaction noises, and sounds of other traffic interactions. The simulator used SCANeR^{TM} studio software that connected eight computers for controlling the dynamics of the simulator car, simulated environment, and recorded basic driving parameters (speeds, accelerations, positions, etc.) at a frequency of 20 Hz.

3.2. Participants

Recognising the importance of randomly sampled participants, we advertised our experiment at various local public places and social media platforms, ensuring the diversity and representativeness of the general public. As a result, this study recruited 78 participants and their descriptive statistics are presented in Table 1. The mean age of the participants was 30.8 years (standard deviation [SD] 11.70 years), with approximately two-thirds of them being male. The mean ages for male and female participants were respectively 34.1 (SD 12.6) years and 24.9 (SD 6.7) years. The mean driving experience of the participants was 12.2 (SD 11.5) years, with more than two-thirds of them possessing an open driving licence (non-restricted). Note that in Queensland, Australia, newly licenced drivers receive a provisional licence for a period of 3 years before they obtain an open licence. Majority of the participants (61.6%) possessed a university degree, while 23.1% passed Grade 12. 24.4% and 23.1% of the participants reported that they usually drive between 5001–10,000 km and 15,001–20,000 km in a typical year, respectively. Eight (out 78) participants reported their involvement in a crash in the last year. About 42% of the participants responded that they have prior information or heard about connected vehicles. As a token of appreciation for volunteering in the experiment, each participant received AU$ 75 after completing the experiment. Note that the entire experiment design consists of several driving tasks, including car-following, lane-changing, interacting with traffic signals, and interacting with a pedestrian walking on a zebra crossing. A detailed discussion on these driving tasks is beyond the scope of this study and can be found in the full experiment design paper (Ali et al., 2020c).

3.3. Design of the vehicle–pedestrian interaction

The Brisbane Central Business District area and its surrounding environment were created in the simulated environment for the driving simulator experiment of ths study. A high degree of accuracy in replicating the real environment was ensured by generating a high-quality photorealistic environment as well as keeping traffic signs and road markings complying with Australian road design standards. Note that the posted speed limit was 40 km/h. The vehicle–pedestrian interaction was judiciously placed on two straight stretches along a city route (see Fig. 1(b) for more illustration). Prior to approaching a pedestrian crossing, drivers drove in the city to familiarise themselves with the driving environment. While approaching a zebra crossing, a driver was required to brake and completely stop their vehicle to yield to the pedestrian. In the experiment, a driver interacted with two zebra crossings in each drive, whereby the pedestrian crossed in one of the randomly selected zebra crossings.

The vehicle–pedestrian interaction was designed in such a way that the pedestrian started to walk from the sidewalk to the zebra crossing when the time taken by the subject vehicle to reach the zebra crossing was less than 6 s (see Fig. 1(c)). This time was calculated based on the speed of the subject driver, implying that the pedestrian would start to walk when the subject vehicle is about 70 m away from the zebra crossing. Drivers, therefore, had sufficient distance available to completely stop their vehicle on the road with the speed limit of 40 km/h that would require a braking distance of only 9 m. Although this time could be varied, e.g., 10 s as used by Haque and Washington (2015), we used a fixed time period to avoid confounding factors, as otherwise, it would be difficult to understand whether the difference in braking behaviour is caused by different time periods or due to a connected environment. Further, the simulated environment and interaction remained the

Fig. 1. Experiment design: (a) Advanced Driving Simulator; (b) Schematic of Brisbane Central Business District; (c) Designed vehicle–pedestrian interaction; (d) Design of advance information driving aid.

Table 1
Characteristics of the participants recruited for the driving simulator experiment.

Driver characteristics	Mean	SD	Count	Percentage
Driver's age (years)	30.8	11.7	–	–
Young drivers	22.11	2.44	38	48.72
Middle-aged drivers	35.34	3.36	32	41.03
Older drivers	58	4.08	8	10.26
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				
0–5000 km	–	–	10	12.8
5001–10,000 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

same for both the baseline and connected environment driving scenarios with one exception, i.e., the provision of advanced pedestrian information in the connected environment.

3.4. Design of the connected environment

Mimicking the vehicle-to-infrastructure communications between roadside units and vehicles, advance information was provided to the participants in the connected environment driving scenario. We conducted a thorough search with a special focus on how major car manufacturers design their in-vehicle information systems. Driving aids in the simulator were provided in two forms: visually (a text message) and auditory (a beep sound). The text message, along with a beep sound, was displayed at the bottom centre of the windscreen resembling the heads-up display fitted in some recent car models. Fig. 1(d) displays a typical example of advance information showing the message "*Watch for pedestrians*" when the subject driver was 6 s away from the zebra crossing. This threshold was selected in accordance with the Austroads guidelines (AUSTROADS, 1993), suggesting that a driver requires 4 s to safely stop before a zebra crossing while travelling at 40 km/h. Allowing an additional time of 2 s for reading and interpreting the message, advance information was disseminated when a driver was 6 s away from the zebra crossing. Note that this information was presented in advance before the pedestrian started to walk towards the zebra crossing.

As part of the participant testing protocol, each participant was briefed about the general objective of the experiment to avoid bias in their driving behaviour. The driving route, driving tasks, and driving aids were explained in detail when they arrived at the CARRS-Q facility. Once they were confident about all the details, they were taken to the simulator room.

Prior to the actual experimental drives, each participant performed a practice drive to get familiar with the simulator car, driving environment, designed interactions, and driving aids. Participants were asked if they felt confident to proceed to the actual experimental drive, and only after their positive response, the actual experiment was started.

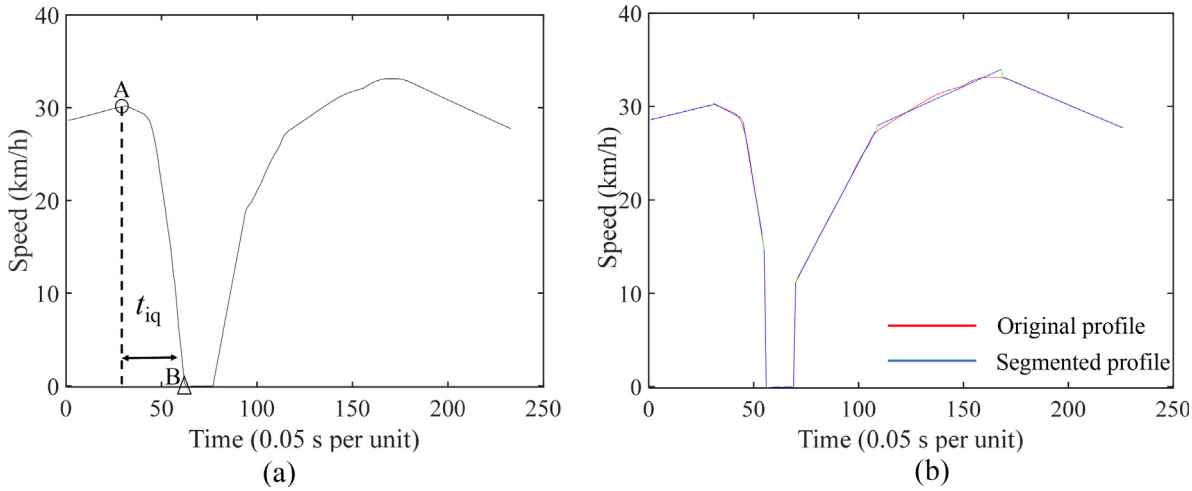

Fig. 2. A typical example of (a) original and (b) segmented speed profiles: Point A indicates when the pedestrian started walking; Point B shows the complete stop of a driver.

3.5. Data collection

The braking profile of each driver in response to the pedestrian at the zebra crossing was extracted from the driving simulator. The braking profile was captured between the periods when the pedestrian started to move from the sidewalk to the zebra crossing, and the subject driver responded by decelerating and reaching their minimum speed. Graphically, the braking profile is the segment between Point A and Point B (see Fig. 2(a)). The Bottom-Up algorithm was adopted (Keogh and Pazzani, 1998) to determine when a driver responded after detecting the pedestrian movement (Fig. 2(b)). Once the driver responded (Point A in Fig. 2(a)), the minimum point on their braking profile was traced (Point B in Fig. 2(a)), and the time taken to reach the minimum speed was taken as braking time (see t_{iq} in Fig. 2(a)).

Table 2 presents the descriptive statistics of variables extracted from the driving simulator experiment. An advanced econometric model corresponding to the survival time of speed changes was developed as a function of driving scenarios (i.e., baseline and connected environment), vehicle dynamics, and driver demographics. Note that Table 2 only contains statistically significant variables found in the parsimonious model, whereas several other demographic variables were tested in the model, such as driving experience, education level, licence type, crash involvement etc. Models with these variables neither yielded better statistical fit nor statistically significant, possibly caused by small sample size that can be addressed by collecting data from a larger number of participants. Further, vehicle dynamics consisted of variables for initial speed, accel-

Table 2
Summary statistics of explanatory variables included in the duration model.

Variable	Description of variables	Count	Percentage	Mean (SD)
Driving scenario				
Baseline	Driving without driving aids (reference)	78	100	–
Connected environment	Driving with driving aids (dummy)	78	100	–
Vehicle dynamics				
Initial speed	Instantaneous speed of the subject vehicle before the driver starts braking in response to a pedestrian crossing (m/s)	–	–	9.68 (1.92)
Distance to the zebra crossing	Distance to the zebra crossing when the driver started to brake (m)	–	–	58.87 (14.88)
Acceleration noise	The standard deviation of acceleration/deceleration of a driver prior to the braking event (m/s^2)	–	–	2.11 (0.55)
Maximum deceleration	Maximum deceleration over the range of the initial speed to the minimum speed (m/s^2)	–	–	0.12 (0.56)
Demographic variables				
Age groups				
Young	Participant is 18–26 years old (dummy)	38	48.72	–
Middle-aged	Participant is 27–50 years old (reference)	32	41.03	–
Older	Participant is 51+ years old (dummy)	8	10.26	–
Gender				
Male	Participant is male (reference)	50	64.10	–
Female	Participant is female (dummy)	28	35.90	–

eration noise, maximum deceleration, and distance to the zebra crossing. Both initial speed and distance to the zebra crossing were measured at the instant just before a driver started braking in response to the pedestrian (i.e., Point A in Fig. 2(a)). The maximum deceleration was measured as the highest deceleration over the braking profile. Acceleration noise is measured as the standard deviation of acceleration/deceleration prior to the braking event. Note that 78 participants faced the vehicle–pedestrian interaction and stopped at the zebra crossing in two repeated drives, forming a panel dataset of 156 observations.

Descriptive analyses were conducted to compare the braking profiles between the two drives (i.e., baseline and connected environment) using repeated measures *t*-tests.

4. Modelling technique

A random parameters duration model was applied to model the braking behaviour of drivers in a connected environment. Succinctly, the braking time (or survival time of speed changes) was modelled using a hazard-based duration modelling approach. The random parameters specification took into account the unobserved heterogeneity associated with the braking behaviour in a connected environment. The developed model was estimated in the Bayesian framework. A detailed description of the model is presented in ensuing subsections.

4.1. Model structure

A hazard-based duration modelling approach is a probabilistic approach that has been frequently used in various transport applications, especially when analysing time-related phenomena. More specifically, this approach is suitable for studying the time-varying probabilities of an event or the duration of an event (Washington et al., 2020). Some applications of duration modelling in transportation are studying travel distance and times in urban environments (Anastasopoulos et al., 2012a, Anastasopoulos et al., 2012b), understanding household evacuation timing behaviour (Hasan et al., 2013), analysing pavement overlay and replacement performance (Anastasopoulos and Mannering, 2015), understanding travellers' habits of using new energy type mode for their transport (Anastasopoulos et al., 2017), identifying elderly travel time disparities (Jordan et al., 2019), modelling lane-changing execution behaviour in a connected environment (Ali et al., 2021c), and identifying environmental factors affecting accident occurrences during snow events (Pang et al., 2022). Similarly, this study applies a hazard-based modelling approach to model braking time or the survival time of speed changes, which is considered as the duration variable-the time taken by a driver to reduce their initial speed to the minimum speed. The parametric duration models, i.e., accelerated failure time, allow covariates to accelerate directly in a baseline survival function where all covariates are zero (Washington et al., 2020). In doing so, an acceleration factor is obtained that can capture the direct effects of exposure on survival time. This attribute also leads to a simpler and intuitive interpretation of modelling results since the estimated parameters quantify the effect of a covariate on the mean survival time (Haque and Washington, 2015). These characteristics, combined with the appropriateness of duration modelling for braking time data, motivated us to adopt a hazard-based duration modelling approach in this study.

Mathematically, the accelerated failure time model is an intrinsically linear form of the survival time, *T*, expressed as a function of covariates in a linear regression setting as.

$$\ln(t_{iq}) = \boldsymbol{\beta}'_i \boldsymbol{x}_{iq} + \boldsymbol{\gamma}' \boldsymbol{z}_i + \sigma \varepsilon_{iq}, \tag{1}$$

where t_{iq} is the survival time of speed changes for each driver $i \in \{1, \ldots, 78\}$ and driving scenario $q \in \{1, 2\}$ (baseline and connected environment), $\boldsymbol{\beta}$ indicates a (column) vector of unknown (and to be estimated) driver-specific parameters, $\mathbf{x}_{iq}$ denotes a (column) vector of driving scenario-specific explanatory variables defined in Table 2, $\boldsymbol{z}_i$ represents column vectors of corresponding driver-specific values of the sociodemographic variables, and $\boldsymbol{\gamma}$ is a column vector of coefficients for the sociodemographic variables to describe unobserved heterogeneity across drivers. ε_{iq} denotes random error term that is assumed to be independently and identically normally distributed with mean zero and standard deviation σ. Further, following Washington et al. (2020), the survival function for accelerated failure time model can be written as.

$$S(t|\boldsymbol{x}_{iq}, \boldsymbol{z}_i) = S_0\left(te^{\boldsymbol{\beta}'_i \boldsymbol{x}_{iq} + \boldsymbol{\gamma}' \boldsymbol{z}_i}\right), \tag{2}$$

which leads to the conditional hazard function as

$$h(t|\boldsymbol{x}_{iq}, \boldsymbol{z}_i) = h_0\left(te^{\boldsymbol{\beta}'_i \boldsymbol{x}_{iq} + \boldsymbol{\gamma}' \boldsymbol{z}_i}\right)\left(e^{\boldsymbol{\beta}'_i \boldsymbol{x}_{iq} + \boldsymbol{\gamma}' \boldsymbol{z}_i}\right), \tag{3}$$

where, S_0 and h_0 are baseline survival and hazard functions, respectively. Eqs. (2) and (3) suggest that there exists a direct relationship between the effects of covariates and the survival time of speed changes, implying that these effects may increase or decrease the braking time.

A precursor of parametric duration modelling is to specify the distribution of the duration variable. A wide range of distributions is used in the literature for estimating survival function, including Weibull, lognormal, exponential, gamma, log-logistic and Gompertz distribution. This study, however, used Weibull distribution, primarily because of two reasons. First, it is flexible that allows modelling of duration data with monotone hazard rates that increase or decrease exponentially with

time or remain constant over time. Second, statistically, Weibull distribution was found to better fit the braking time data compared to other distributions, as indicated by an Anderson Darling test (*test statistics* = 0.47; *p*-value = 0.25). Thus, this study applied Weibull distribution to model the survival time of speed changes (from the initial speed to the minimum), and its hazard function can be written as.

$$h(t) = (\lambda P)\,(\lambda t)^{P-1}, \tag{4}$$

and the survival function of the Weibull duration model is shown as.

$$S(t) = exp\,(-\lambda t^{P}), \tag{5}$$

where, λ and P are respectively the location and scale parameters.

Recent literature reported that different drivers might perceive and react differently to driving aids provided by a connected environment (Sharma et al., 2020, Ali et al., 2020b), which leads to preference heterogeneity and differential effects of the same driving aids. For instance, when drivers are provided with the information of available gaps in the adjacent lane, they tend to select different gap sizes in a connected environment during mandatory and discretionary lane-changing manoeuvres (Ali et al., 2018, Ali et al., 2020d). Similarly, when the same advanced information related to traffic light change was presented to drivers, it was found that some drivers stopped before the stop line while the others ran through the yellow light (Ali et al., 2021b). As such, driving behaviour in a connected environment is likely to be heterogeneous, which is mostly unobserved. To account for such unobserved heterogeneity, a random parameters modelling approach is adopted in this study, which allows the estimated parameters to vary across individual drivers. More specifically, by using random parameters in hazard-based duration modelling, we consider $\boldsymbol{\beta}_i$ to be driver-specific random parameters for the operational variables defined as.

$$\boldsymbol{\beta}_i = \boldsymbol{\mu} + \boldsymbol{\Psi z}_i + \boldsymbol{\Omega\varphi}, \tag{6}$$

where $\boldsymbol{\varphi}$ is a column vector of independent standard normally distributed random variables, and $\boldsymbol{\beta}_i$ is assumed to follow a multivariate normal distribution with mean $\boldsymbol{\mu} + \boldsymbol{\Psi z}_i$ and covariance matrix $\boldsymbol{\Omega\Omega}'$, where $\boldsymbol{\Omega}$ is a lower triangular Cholesky matrix containing information about (co)variances and accounting for possible correlations in the coefficients (Fountas et al., 2018, Greene, 2012). Assuming that $\boldsymbol{\sigma}$ denotes the non-zero elements in the Cholesky matrix, our study aims to estimate coefficients in $\boldsymbol{\sigma}$, in vectors $\boldsymbol{\mu}$, γ, and in matrix $\boldsymbol{\Psi}$ (describing unobserved heterogeneity across drivers with respect to the sensitivity towards traffic operational scenarios). Further, this study used an unrestrictive form of the Cholesky matrix that allows capturing correlation between two or more random parameters as well as accounting for the panel nature of data or group effect (see Fountas et al. (2018), Yu et al. (2015) and Eker et al. (2019)) for more details).

The developed model can be estimated using simulated maximum likelihood estimation. Alternatively, the same model can also be estimated in the Bayesian framework, which is frequently used for hierarchical models. Note that the model developed in this study could also be viewed as a hierarchical model where the mean of the random parameters (at level 1) is modelled as a function of covariates. Bayesian estimation offers a significant advantage over the maximum likelihood estimation in that complicated hierarchical likelihood functions (such as the one in this study) and posteriors can be considered in model estimation (Oviedo-Trespalacios et al., 2020).

4.2. Model inference

In Bayesian models, inferences are made on model parameters based on their posterior distributions built from the model's likelihood and prior distributions assigned to all estimable parameters. In other words, the posterior distributions of model parameters depend on the prior probabilities assigned to parameters. As prior information of model parameters was unavailable, we used uninformative priors, assuming that parameters follow a normal distribution with mean $\hat{\beta}_i$ and large variance, i.e.,$N(\hat{\beta}_i, 10^6)$. Note that $\hat{\beta}_i$ is the mean of parameters estimated for a model, where we also accounted for the panel nature of the data (i.e., each participant drives both in the baseline and connected environment scenarios). Instead of using zero mean, we opted $\hat{\beta}_i$ to accelerate the model convergence for the parameters.

In this study, the inference of posterior distributions is obtained using Markov Chain Monte Carlo simulation and Gibbs sampling (Spiegelhalter et al., 2002).

4.3. Model selection

The Deviance Information Criterion (DIC) is adopted in this study, which is frequently used in the Bayesian framework. The basic notion of DIC is to select the simplest model that can explain as much of the variation in the data as possible (Spiegelhalter et al., 2002). Mathematically, it can be obtained as.

$$DIC = \bar{D} + p_D \tag{7}$$

where $\bar{D}$ and p_D respectively denote the posterior mean deviation measuring the model fit and the effective number of model parameters reflecting the complexity of a given model. Similar to other statistical fit measures, a lower DIC value indicates a better model and vice versa.

To develop further insights into the effects of the explanatory variables on the survival time of speed changes, the exponent of each coefficient $(1 - \exp(\beta))$ was computed (Haque and Washington, 2015, Washington et al., 2020), indicating a percent change in survival time corresponding to a unit increase in the continuous variable or a change from zero to one for categorical variables. Further, the statistical significance of model parameters is assessed using the Bayesian credible intervals.

5. Results

5.1. Descriptive analysis of braking profile

Several driving indicators during the braking episode were measured for comparing the braking performance of drivers between the connected environment and the baseline scenario, and results are summarised in Table 3.

Acceleration noise-an indicator of reckless driving-was found to be statistically significant between the baseline and connected environment driving scenarios. Notably, drivers' acceleration noise was reduced by about 0.5 m/s^2 in the connected environment, reflecting their safer braking behaviour in the connected environment. Similarly, it was found that drivers in the connected environment reacted early to the pedestrian walking from the sidewalk to the zebra crossing, which could be attributed to the provision of advanced information. Finally, the difference in the initial speed (or approaching speed) was found to be statistically significant between the baseline and connected environment driving scenarios. A paired *t*-test indicated that initial speed was about 0.941 m/s lower when drivers were assisted with driving aids while approaching the zebra crossing. It can be inferred that a lower initial speed in the connected environment could have triggered an early response from drivers.

Braking time or the time taken by drivers to reduce their initial speed to the minimum was also found to be statistically significant. A paired *t*-test showed that drivers took about 0.4 s longer in the connected environment to reach their minimum speeds. By taking a long time in the connected environment, drivers appeared to smoothly reduce their speeds, which is also evident from speed variations reported in Fig. 3. Differences in the speed variations were also found to be statistically sig-

Table 3
Summary of descriptive analysis for the braking profile.

Indicator	Baseline (SD)	Connected environment (SD)	Significance by paired *t*-tests	Remark
Acceleration noise (m/s^2)	2.351 (0.526)	1.863 (0.483)	*t*-stat = 6.031; *p*-value < 0.001	Significant
Initial (or approaching) speed (m/s)	10.158 (1.886)	9.217 (1.858)	*t*-stat = 3.141; *p*-value = 0.002	Significant
Time to reduce the initial speed to the minimum (s)	1.585 (0.603)	1.992 (0.485)	*t*-stat = −4.646; *p*-value < 0.001	Significant

SD: standard deviation.

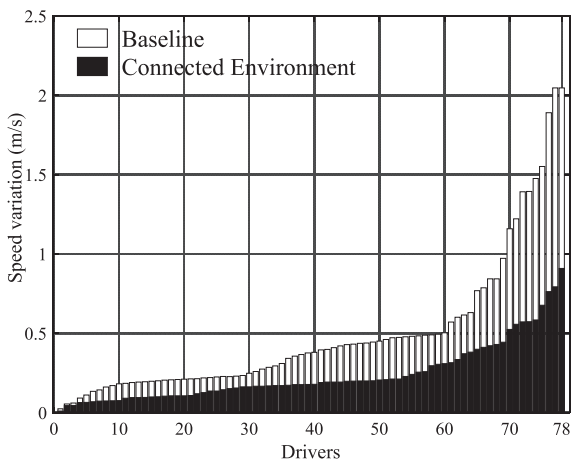

Fig. 3. Speed variations during the braking episode (sorted in ascending order).

Table 4
Estimation results of the Bayesian random parameters duration model.

Variable	*Parameter estimate*	*s. d.*	*MC error*	exp(β)	Bayesian credible intervals		
					2.50%	*Median*	*97.5%*
Non-random parameters							
Constant	−1.109	0.754	0.025	–	–	–	–
Initial speed	0.263	0.093	0.003	1.3	0.089	0.268	0.467
Acceleration noise	−0.385	0.245	0.007	0.68	−0.863	−0.381	−0.003
Maximum deceleration	−0.302	0.205	0.004	0.739	−0.703	−0.298	−0.005
Young drivers	−0.382	0.21	0.002	0.683	−0.794	−0.379	−0.003
Older drivers	0.487	0.293	0.004	1.627	0.086	0.492	1.066
Random parameters							
Connected environment (mean)	0.428	0.156	0.005	1.534	0.139	0.436	0.732
Distance to the zebra crossing (mean)	0.049	0.013	0.0004	1.051	0.024	0.049	0.075
Diagonal values in Cholesky matrix							
Connected environment (CE)	0.285	0.223	0.004	–	–	–	–
Distance to the zebra crossing (DZC)	0.056	0.094	0.003	–	–	–	–
Below diagonal values in Cholesky matrix							
Distance to the zebra crossing: connected env.	0.021	0.003	0.0009	–	–	–	–
Heterogeneity in the mean of connected environment							
Female	0.743	0.038	0.0007	–	–	–	–
$\bar{D}$ = 201.32; p_D= 39.589; DIC = 240.91; Scale = 4.416 (p-value <0.001); No. of observations = 156; No. of groups = 78; Group size = 2							

s. d.: standard deviation; MC: Monte Carlo.

nificant (p-value <0.001), with speed variations being approximately 50% higher in the baseline scenario compared to the connected environment, suggesting that driving aids help drivers to maintain a smooth speed profile.

5.2. Bayesian random parameters duration model for braking

The Bayesian random parameters duration model estimates for the survival time of speed changes are presented in Table 4. Note that this model was estimated in WinBUGS software. Two separate Markov chains were used for each parameter with different initial values, and the the Markov chain Monte Carlo (MCMC) was performed for 110,000 iterations. The first 10,000 iterations were discarded as burn-in samples. The model converged after 70,000 iterations, which was meticulously examined by (i) calculating the Gelman-Rubin statistics of two chains and (ii) visually inspecting the trace plots of parameters chains (Spiegelhalter et al., 2002), which confirmed the model convergence. Further, the simulation was continued for additional 30,000 iterations to obtain the posterior distributions of each parameter.

As mentioned in Section 3.1, this model captures correlations between random parameters and accounts for the panel nature of data. Two variants of the Bayesian random parameters duration model were compared, namely a correlated random parameters model and an uncorrelated random parameters model. The comparative analysis results suggested that the correlated random parameters model outperformed its counterpart, with a smaller Deviance Information Criterion (DIC) value. The selected model had a scale parameter of 4.416, which is significantly greater than 1 (p-value <0.001), implying a positive duration dependence and that an event follows a monotone hazard function. In other words, the survival time of speed changes decreased with an increase in time. For instance, the speed survival probability after 4 s was, on average, about 12 times (i.e., $(4/2)^{4.416-1}$) lower than that of 2 s. This decreasing survival probability trend reflected the scenario of drivers' stopping when approaching the zebra crossing, thereby ensuring the appropriateness of the duration modelling approach.

The estimated model identified two random parameters, namely a dummy variable for the connected environment and distance to the zebra crossing. Several distributions such as normal, log-normal, uniform, and triangular distributions were tested for specifying the distribution of these parameters, and the normal distribution was found to outperform others in terms of goodness-of-fit. The connected environment variable included gender as the heterogeneity in the mean of the random parameter. Fixed parameters in the model were initial speed, acceleration noise, maximum deceleration, and driver age. The braking time function (Equation (1)) can be rewritten as.

$$\begin{aligned} mean\,(t_{iq}) = \exp(&-1.109 \\ &+\beta_{CE} \times CE + \beta_{DZC} \times DZC \\ &+0.428 \times initial\,speed - 0.385 \times acc.\ noise - 0.302 \times \max.\ acc. \\ &-0.382 \times YoungDriver + 0.487 \times OlderDriver), \end{aligned} \tag{8}$$

where the first line contains the constant, the second line indicates random parameters (CE: connected environment, DZC: distance to the zebra crossing), and the third and fourth lines represent non-random parameters, including vehicle dynamics and driver demographics, respectively, where.

$$\begin{pmatrix} \beta_{CE} \\ \beta_{DZC} \end{pmatrix} = \begin{pmatrix} 0.428 \\ 0.049 \end{pmatrix} + \begin{pmatrix} 0.743 \\ 0 \end{pmatrix} \times \text{FemaleDriver} + \begin{pmatrix} 0.258 & 0 \\ 0.018 & 0.056 \end{pmatrix} \begin{pmatrix} S_1 \\ S_2 \end{pmatrix} \tag{9}$$

is the specified correlation structure between random parameters with S_1 and S_2 be the independent standard normally distributed random variables.

Table 4 also presents the diagonal and below diagonal elements of the Cholesky matrix for each random parameter. Using these elements, the standard deviation of each random parameter can be calculated as the square root of the variance (i.e., elements of the variance–covariance matrix, which can be computed as $\mathbf{\Omega\Omega}'$). For example, the standard deviations for the connected environment and distance to the zebra crossing parameters are computed as $\sqrt{0.066} = 0.258$ and $\sqrt{0.003} = 0.059$, respectively. Note that several other random parameters were also tested in the Bayesian model, but they did not improve the model fit, and hence left out of the parsimonious model.

A driver's *initial speed* at the approach to the zebra crossing was significant and positively associated with braking time. The model suggests that a 1 m/s increase in the initial speed was associated with a 30% increase in the time required to reach the minimum speed (Table 4). With a higher speed, drivers will require more time to stop before the zebra crossing safely.

Unlike initial speed, *acceleration noise* had a significant and negative effect on the time to reach the minimum speed. More specifically, with a 1 m/s^2 increase in acceleration noise, the time required to reach the minimum speed decreased by 32% (Table 4). This finding suggests that drivers with higher acceleration noise are likely to reach their minimum speed earlier, perhaps because of hard braking. Similarly, the *maximum deceleration* parameter was found to be significant and negatively associated with braking time (Table 3).

Dummy variables for *young and older drivers* were negatively and positively associated with speed survival time, respectively. Compared to the middle-aged drivers, young (older) drivers took about 31% (63%) shorter (longer) in reducing their speeds to the minimum speed or zero (Table 4).

Table 4 also illustrates that the mean and standard deviation of the connected environment dummy variable were significant. The presence of randomness in the variable connected environment indicates significant heterogeneity in braking time for this parameter, as shown in Table 5. The braking time was found to increase for the majority of drivers (90%) in the connected environment compared to the baseline scenario, but there was a group of drivers who took shorter time in the connected environment to reduce their speeds. This finding underscored the existence of preference heterogeneity in the connected environment, reflecting that not every driver perceived the advanced information in the same way. Some drivers used this information to brake slowly, exhibiting their safer behaviour, while other drivers used it in a counterproductive manner by delaying their response and then perhaps brake faster to compensate for the increased crash risk.

From Table 4, it can be observed that the heterogeneity in the connected environment is a function of driver gender. Specifically, we use a simulation approach to calculate the braking times for both male and female drivers. Results from simulations indicate that the braking time of female drivers in the connected environment was twice more than that of male drivers. We further elaborate on this finding in the next section.

The mean and standard deviation of the second random parameter, i.e., distance to the zebra crossing when a driver started to brake, was found to be significant and positively associated with braking time, revealing a significant heterogeneity in braking time corresponding to the distance to the zebra crossing. As the heterogeneity is confirmed in Table 5, more than three-quarters of drivers took longer to stop in the connected environment, whereas the remaining drivers were found to brake faster. The mean of this parameter was positive, implying that when this distance was large, drivers were more likely to take longer time to stop, which is intuitive because there was no urgency to stop, and they could slowly reduce their speed. On the other hand, braking time was found to decrease with an increase in the distance for some drivers, reflecting the cautious behaviour of drivers who were hesitant to delay their response and then compensate by hard braking.

The unrestricted form of the Cholesky matrix allowed the estimation of correlation between random parameters, which offers more insights into the braking behaviour of drivers in response to the pedestrian at the zebra crossing. Results indicated that the random parameters for connected environment and distance to the zebra crossing were statistically correlated at a 95% confidence level (t-stats = 4.07; p-value <0.001), with a covariance of 0.01 and a correlation coefficient of 0.31. Of note, t-stats and correlation coefficients were calculated following the post estimation technique presented in Fountas et al. (2018), and interested readers are referred to this study for mathematical formulations. Correlation between random parameters indicates the presence of interactions of unobserved heterogeneity associated with random parameters. A positive correlation between the random parameters (i.e., distance to the zebra crossing and the connected environment) implies that an

Table 5
Distributional effects of the random parameters.

Random parameter	Above zero	Below zero
Connected environment	90%	10%
Distance to the zebra crossing	79%	21%

increase in the effect of the distance to the zebra crossing (indicated by β_{DZC}) in the connected environment was likely to increase braking time because of unobserved heterogeneity associated with these random parameters. This finding underscores the notion that when drivers are informed about pedestrian movements on the zebra crossing in the connected environment, they make early decisions and smoothly reduce their speed by taking more time.

5.3. Decision tree analysis

As reported in the previous section, the Bayesian random parameters duration model identifies two classes of braking time in a connected environment: drivers with increased braking times and drivers with decreased braking times. To further develop insights into what causes this differential behaviour in a connected environment given that the same information is presented to all the drivers, we adopted a decision tree analysis because such information cannot be obtained from the Bayesian random parameters model. Several classification algorithms are available in the literature, namely neural networks, support vector machine, decision tree, and ensemble algorithms. However, the decision tree algorithm selection was governed by a combination of better accuracy and interpretability, which corroborates with a previous study (Sharma et al., 2020).

To employ a decision tree algorithm, the same variables as reported in Table 1 were used. However, vehicle dynamics variables were used to calculate their respective ratio. For instance, the initial speed ratio was defined as the initial speed in the connected environment divided by the initial speed in the baseline environment. Similarly, all other ratios were calculated. Note that the outcome variable for the decision tree was a binary variable capturing an increase or decrease in braking time in the connected environment.

Fig. 4 presents the decision tree for classifying the braking time, which is estimated using Python *CHAID* library. Considering the right branch of the tree, it can be observed that for older drivers, when the distance to the zebra crossing was larger in the connected environment than that in the baseline scenario, braking time was likely to increase (i.e., outcome *A* in Fig. 4). Older drivers being aware of the zebra crossing in the connected environment took more time to gradually reduce their speed without any safety concerns. Similarly, the left branch of the tree suggests that when the initial speed of young and middle-aged drivers in the connected environment was higher than that in the baseline scenario and distance to the zebra crossing in the connected environment was smaller compared to the baseline scenario, drivers took less time to reduce their speed (outcome *B* in Fig. 4). Young drivers in the connected environment with higher speed and short distance from the stop line were found to brake hard. The rest of the decision tree can be interpreted in a similar manner.

Two noteworthy conclusions from the decision tree analysis can be made. First, while there is significant heterogeneity in braking time when drivers receive advance information from the connected environment, braking time increases for most drivers. This finding is consistent with our model estimation results in the previous section. Secondly, a necessary scenario

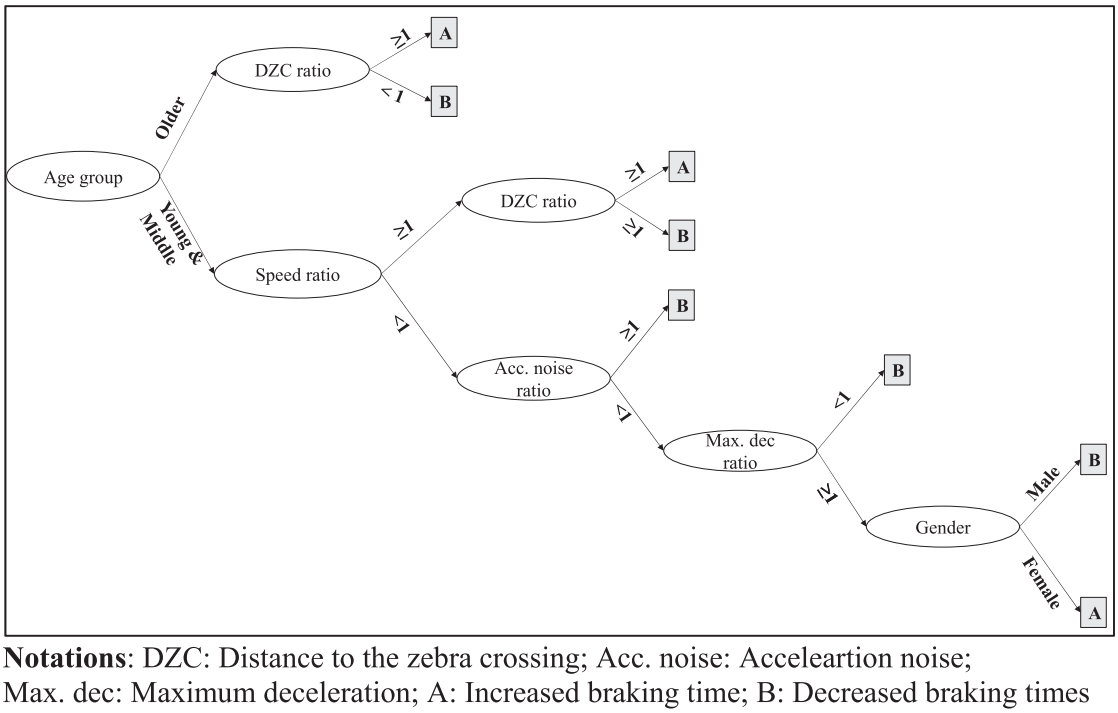

Fig. 4. The estimated decision tree for classifying differential braking time.

for an increase in braking time was the early response of drivers (measured as the large distance to the zebra crossing) in the connected environment than that of the baseline scenario. This finding is intuitive and has been explained previously.

6. Discussion

6.1. Braking behaviour in the connected environment

The Bayesian random parameters duration model (Table 3) facilitated examining the impact of the connected environment on various combinations of explanatory variables after controlling for vehicle dynamics and other exogenous factors. The developed model allows us to plot survival curves that assist in examining braking behaviours in response to a pedestrian at the zebra crossing. More specifically, speed survival probabilities were calculated using Equation (10) and the parameter estimates in Table 3. For instance, the survival probability at time t under average driving scenarios in the connected environment (see Table 1 for the average values of continuous variables and reference category for categorical variables) can be calculated according to Eq. (10). Baseline survival probabilities can be computed in a similar way. The corresponding survival curves are plotted in Fig. 5.

$$S(t)_{CE} = \exp\left(-\exp\left(-4.416(-1.109 + 0.263 \times 9.68 - 0.382 \times 2.11 - 0.302 \times 0.12 + (0.428 \times 1 + 0.743 \times 0))\right)t^{4.416}\right). \tag{10}$$

Fig. 5 shows that the speed survival probabilities decreased with elapsed time. Drivers in the baseline scenario (without any driving aids) appeared to reduce their initial speed earlier compared to the connected environment scenario. The speed survival probability, for instance, at 2 s in the baseline scenario was 21%, whereas the corresponding probability in the connected environment scenario was 79%, suggesting a 58% difference in the latter driving scenario. Using the survival curves, we found that drivers in the connected environment scenario took about 4.3 s to reduce their initial speed to the minimum, while speed in the baseline scenario, on average, survived 1.4 s less, reflecting more aggressive braking to reduce the speed in the baseline scenario. This finding implies that when drivers are not assisted with driving aids, they tend to decelerate more sharply in response to a pedestrian at the zebra crossing.

A connected environment increases situational awareness, thereby providing the benefit of additional time to react during different driving tasks (Sharma et al., 2019). For instance, when drivers received advanced information about hard braking events during car-following, they were found to take more time to respond to the event reflecting their proactive driving behaviour (Sharma et al., 2020). Similarly, Ali et al. (2020a) reported that hard braking events in the connected environment were reduced by more than 50% during lane-changing manoeuvres, which can be attributed to an early response when advance information of traffic events was provided. These studies suggest a direct relation of a driver's early response to a situation with the awareness of the surrounding traffic environment created by a connected environment, thereby increasing safety margin. On the other hand, a delayed response is often associated with increased crash risk unless drivers perform

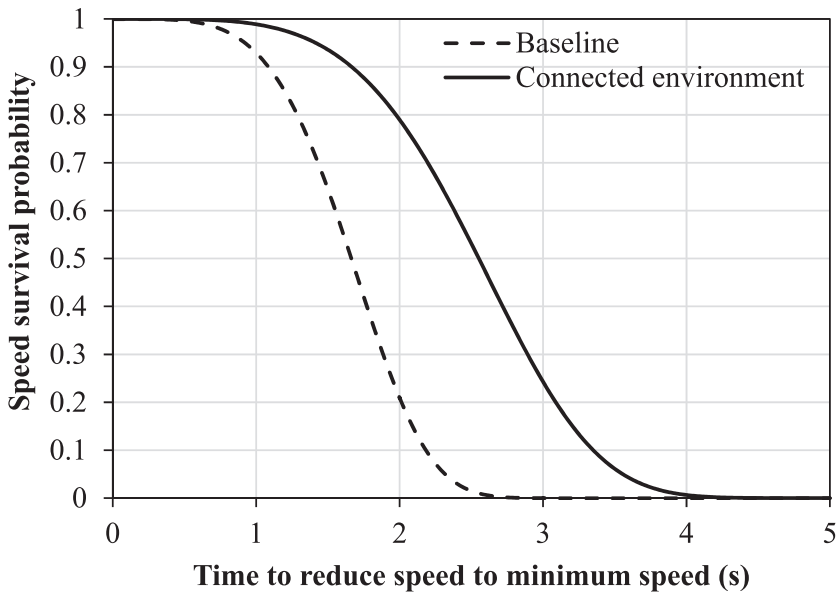

Fig. 5. Speed survival graphs for different driving scenarios.

a hard (or rapid) deceleration to avoid engaging in safety–critical events (Harbluk et al., 2007). A study reported that drivers, who were slow in responding to traffic light changes, brake harder as compensation for slow response (Hancock et al., 2003). In summary, shorter braking time in the baseline scenario can be associated with hard braking, which is directly linked to a delay in responding, perhaps because of an inaccurate perception of the presence of the pedestrian (Haque and Washington, 2015). However, such risk is minimised in the connected environment when drivers receive advanced information about the presence of a pedestrian at the zebra crossing, and thereby drivers are prepared to brake early and smoothly.

Although most drivers took a longer time to brake, the model also suggested that some drivers took a shorter time to brake. Drivers with shorter braking time may have used advanced information of the connected environment to respond earlier to minimise risk, reflecting their cautious, proactive, and safer braking behaviour. Ali et al. (2020b) found that drivers in the connected environment took shorter time to respond to a lane-changing request and reduced their speed gradually, as illustrated by smaller jerks and speed variations.

To summarise, drivers were found to drive safely when the connected environment provided advanced information as they avoid hard braking, which may lead to engaging in safety–critical events during their interaction with the pedestrian.

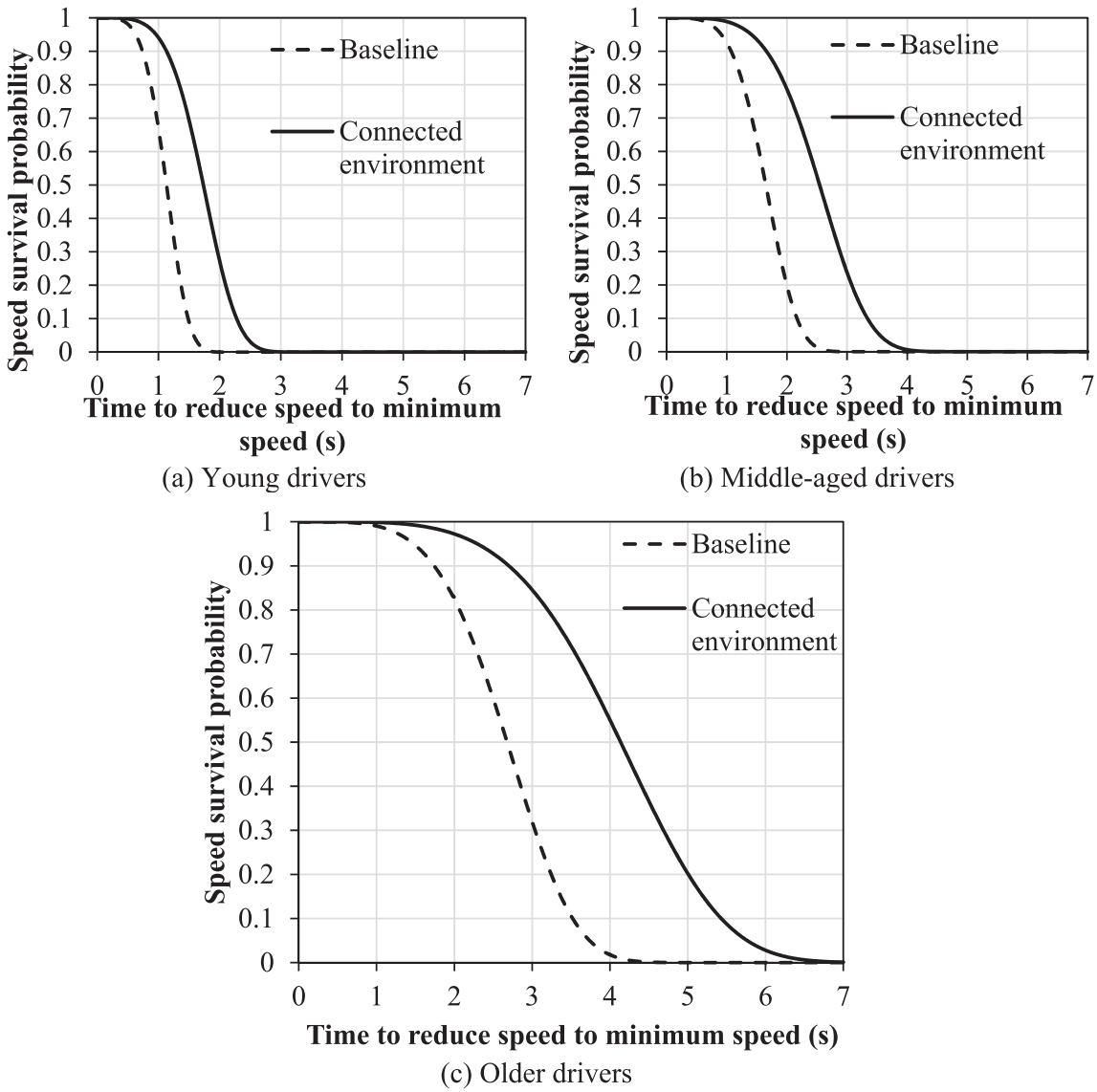

Fig. 6. Effects of the connected environment on braking behaviour of different age groups.

6.2. Effects of drivers' characteristics on braking behaviour

6.2.1. Driver age

Fig. 6 displays the braking behaviour of young, middle-aged, and older drivers in the baseline and connected environment driving scenarios. The speed survival probabilities in the connected environment for young, middle-aged, and older drivers at 2 s were 28%, 78%, 97%, respectively, while the corresponding probabilities in the baseline scenarios were respectively <1%, 21%, and 83%.

The difference between the areas under the curve for the speed survival probabilities of the baseline and connected environment scenarios was calculated for comparing braking behaviours of different age groups, whereby a large difference in the area reflects a higher safety margin in the connected environment. The difference in the areas under the curve for young and middle-aged drivers was 0.60 (Fig. 6(a)) and 0.88 (Fig. 6(b)), respectively, indicating that compared to young drivers, middle-aged drivers in the connected environment took more time, indicating more gradual and safe braking behaviour. Complementing this finding, the model further showed that relative to the baseline scenario, middle-aged drivers took about 0.94 s longer in the connected environment to reduce their initial speeds to their minimum speed compared to young drivers. The decision tree analysis also confirmed that young drivers took longer to reduce their initial speeds in the connected environment. Young drivers, who are often reported to be risky, aggressive, and tend to brake harder, took a shorter time to come to a complete stop in response to a pedestrian at the zebra crossing. Simons-Morton et al. (2009) reported that young drivers exhibited more hard braking events, which could be attributed to their risk-taking behaviour. However, such behaviour was significantly reduced in the connected environment when young drivers were assisted with driving aids. On the other hand, Simons-Morton et al. (2013) found that middle-aged drivers braked slowly when they received feedback from a safety monitoring system, complementing the behaviour of middle-aged drivers observed in our study. Overall, the connected environment appeared to provide safety benefits to both age groups, with higher benefits for middle-aged drivers.

Compared to the baseline driving scenario, it was found that older drivers in the connected environment took about 2.12 s more to brake in response to the zebra crossing (Fig. 6(c)) compared to young drivers. From the decision tree analysis, it was observed that older drivers with a large distance to the zebra crossing took longer to reduce their initial speeds. Past studies have reported slower sensory motors and processing power of older drivers, which is often reported as one of the causes for increased crash risk of older drivers (Karthaus and Falkenstein, 2016, Salvia et al., 2016, Strayer and Drew, 2004). However, the connected environment had been shown to reduce such crash risk as older drivers were found to utilise this information by taking more time to gradually brake in response to the zebra crossing.

A similar comparison of middle-aged and older drivers revealed that older drivers took about 1.20 s more time than middle-aged drivers, suggesting that older drivers benefitted more from the advanced information, as they slowly reduced their speed by taking longer time. Salvia et al. (2016) reported shorter braking time of middle-aged drivers compared to older drivers while approaching a traffic light, suggesting the cautious behaviour of older drivers. Nevertheless, the connected environment was found to assist in reducing speed for both driver groups when interacting with the zebra crossing. Aligned with this finding, Wan et al. (2016) found that the braking times of middle-aged and older drivers decreased in a connected environment when they received warning messages in a driving task.

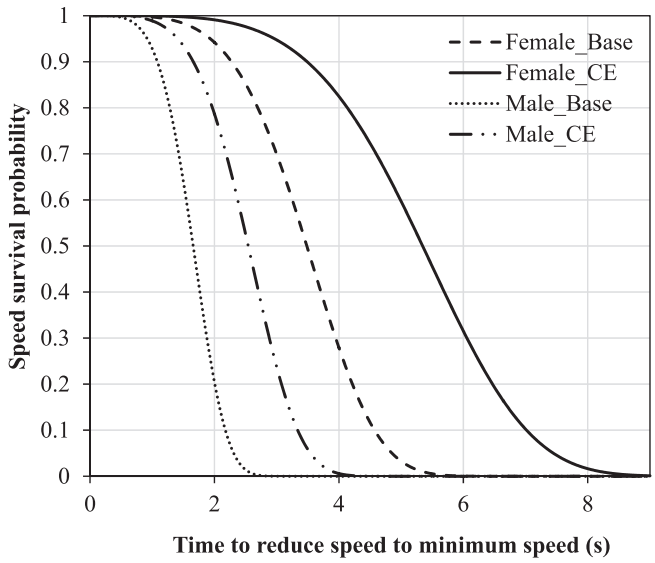

Fig. 7. Gender difference in braking behaviour in the connected environment (CE); Base: baseline.

6.2.2. Driver gender

Fig. 7 shows the braking behaviour of male and female drivers in the baseline and connected environment driving scenarios when they approached the zebra crossing. The speed survival probability for female drivers in the connected environment at time 2 s was 99%, while the corresponding probability for male drivers in the same driving scenario was 79%, suggesting a 20% increase in braking time for female drivers. The decision tree analysis also indicated that female drivers took longer to reduce their initial speed when initial speed and maximum deceleration ratio were lower in the connected environment. This finding underscores that female drivers reduced their speeds more slowly, thereby exhibiting safe braking behaviour. A recent study reported smoother braking response of female drivers in a connected environment, reflecting their safer driving behaviour during a car-following task (Chang et al., 2019).

As male drivers are frequently reported to be risky and more likely to be engaged in safety–critical events (Montgomery et al., 2014, Iversen and Rundmo, 2004), their braking times appeared to increase in the connected environment by 1.58 s than that in the baseline scenario. This suggests that male drivers also benefitted from the connected environment and exhibited safer behaviour. The corresponding increase in the braking time of female drivers in the connected environment was 3.06 s, suggesting that females took more benefits from the connected environment compared to male drivers. Ali et al. (2020a) also reported a similar finding that female drivers exhibited 10% lower hard-braking events than male drivers during lane-changing manoeuvres in a connected environment. Similar safety benefits for female drivers in other contexts, such as the effects of fog and warning systems, have also been reported (e.g., Li et al. (2016) and Li et al. (2015)).

7. Conclusions

The objective of this study was to examine the braking behaviour of drivers when they interacted with a pedestrian walking from a sidewalk to a zebra crossing. A diverse group of drivers aged between 18 years and 65 years was recruited for a driving simulator experiment with two randomised driving scenarios (i.e., baseline and connected environment). A Bayesian random parameters duration modelling approach was applied for investigating the braking behaviour in the connected environment. In particular, the time taken by a driver to reduce their speed to the minimum was modelled in the Bayesian framework using a correlated grouped random parameters accelerated failure time Weibull duration model with heterogeneity-in-the-means. The developed model contained random (a dummy variable for the connected environment with heterogeneity in the mean explained by gender and distance to the zebra crossing) and non-random parameters (initial speed, acceleration noise, maximum deceleration rate, and dummy variables for driver demographics).

Overall, the model suggests that drivers in the connected environment were more likely to take a longer time to reduce their initial speeds in response to the zebra crossing. The improved situational awareness by the advanced information of the connected environment might have helped drivers to swiftly reduce their speed without aggressive braking. The random parameter of the connected environment revealed that most drivers in the connected environment had longer braking times when they interacted with the pedestrian, which implies that most drivers decelerated more gradually and smoothly in response to the pedestrian crossing. Moreover, to understand the factors linked to such differential braking behaviour, a decision tree analysis was performed that revealed that a driver's braking time generally increased in the connected environment but may also decrease if the distance to the zebra crossing was smaller in the connected environment compared to the baseline scenario. Age-related differential braking behaviours were also observed in the connected environment, whereby older drivers benefitted the most from the connected environment and took the longest time to reduce their speeds, while young drivers were found to quickly reduce their speeds. Similarly, the braking time of female drivers was longer in the connected environment compared to male drivers, suggesting that female drivers better utilise the advanced information for their safer braking behaviour.

As it has been repeatedly mentioned in the literature that both excess and inadequate braking is associated with rear-end collisions, the findings of this study provide an in-depth understanding of how such risks can be minimised, if not completely eliminated, in a connected environment where drivers are assisted with advanced information. By understanding the differential braking behaviour, the design of the connected environment could be further improved and tailored to the needs of a specific age/gender group. For instance, findings of this study reveal that young and men drivers tend to improve their braking behaviour in the connected environment, but this improvement is not as much as other groups of drivers. Considering that both young and male drivers are generally risky groups of drivers, a more appealing interface of a connected environment should be designed that can fascinate young and men drivers who are mostly tech-savvy and can be lured by an attractive driving environment.

Recognising the difficulty in controlling for multiple exogenous factors (e.g., driver heterogeneity, different forms of information display, and messages), this study only analysed one type of information display in the connected environment. However, to fully understand the effects of a connected environment and proper deployment of this environment, it is imperative to evaluate various modes of design so that car manufacturers can consider an optimal design.

This study analysed braking behaviour in a fully functioning connected environment, whereas recent studies have shown that impairment in a connected environment could deteriorate driving behaviour, thus merits a separate investigation. Further, in this study, the time when the pedestrian started to walk was kept constant to avoid any confounding factors. Future studies can vary this time and analyse the effectiveness of a connected environment. Finally, since there was no traffic in the direction of the travel, we could not analyse the effects of a driver position in the traffic stream and link it to braking behaviour in a connected environment.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This research is partly funded by the Australian Research Council Discovery Grants, Project ID DP210102970.

References

Ahmed, M., Hoque, M., Rios-Torres, J., Khattak, A., 2017. Freeway merge assistance system using dsrc. In: Proceedings of the 2nd ACM International Workshop on Smart, Autonomous, and Connected Vehicular Systems and Services, pp. 83–84.
Al-Darrab, I., Khan, Z., Ishrat, S., 2009. An experimental study on the effect of mobile phone conversation on drivers' reaction time in braking response. Journal of Safety Research 40 (3), 185–189.
Ali, Y., Bliemer, M., Haque, M., Zheng, Z., 2021a. Examining braking behaviour during failed lane-changing attempts in a simulated connected environment with driving aids. Transportation Research Part C 136, 103531.
Ali, Y., Bliemer, M., Zheng, Z., Haque, M., 2020a. Comparing the usefulness of real-time driving aids in a connected environment during mandatory and discretionary lane-changing manoeuvres. Transportation Research Part C 121, 102871.
Ali, Y., Bliemer, M., Zheng, Z., Haque, M., 2020b. Cooperate or not? Exploring drivers' interactions and response times to a lane-changing request in a connected environment. Transportation Research Part C 120, 102816.
Ali, Y., Haque, M., Zheng, Z., Bliemer, M., 2021b. Stop or go decisions at the onset of yellow light in a connected environment: A hybrid approach of decision tree and panel mixed logit model. Analytic Methods in Accident Research 31, 100165.
Ali, Y., Sharma, A., Haque, M.M., Zheng, Z., Saifuzzaman, M., 2020c. The impact of the connected environment on driving behavior and safety: A driving simulator study. Accident Analysis & Prevention 144, 105643.
Ali, Y., Zheng, Z., Haque, M., 2018. Connectivity's impact on mandatory lane-changing behaviour: evidences from a driving simulator study. Transportation Research Part C 93, 292–309.
Ali, Y., Zheng, Z., Haque, M., Yildirimoglu, M., Washington, S., 2020d. Understanding the discretionary lane-changing behaviour in the connected environment. Accident Analysis and Prevention 137, 105463.
Ali, Y., Zheng, Z., Haque, M.M., 2021c. Modelling lane-changing execution behaviour in a connected environment: A grouped random parameters with heterogeneity-in-means approach. Communications in Transportation Research 1, 100009.
Anastasopoulos, P.C., Fountas, G., Sarwar, M.T., Karlaftis, M.G., Sadek, A.W., 2017. Transport habits of travelers using new energy type modes: A random parameters hazard-based approach of travel distance. Transportation Research Part C 77, 516–528.
Anastasopoulos, P.C., Haddock, J.E., Karlaftis, M.G., Mannering, F.L., 2012a. Analysis of urban travel times: Hazard-based approach to random parameters. Transportation Research Record 2302 (1), 121–129.
Anastasopoulos, P.C., Islam, M.B., Perperidou, D., Karlaftis, M.G., 2012b. Hazard-based analysis of travel distance in urban environments: longitudinal data approach. Journal of Urban Planning and Development 138 (1), 53–61.
Anastasopoulos, P.C., Mannering, F.L., 2015. Analysis of pavement overlay and replacement performance using random parameters hazard-based duration models. Journal of Infrastructure Systems 21 (1), 04014024.
Austroads, 1993. Rural Road Design. Guide to the Geometric Design of Rural Roads. AUSTROADS Publication AP-1/89, Sydney, Australia.
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 51, 88–102.
Caird, J., Chisholm, S., Lockhart, J., 2008. Do in-vehicle advanced signs enhance older and younger drivers' intersection performance? Driving simulation and eye movement results. International Journal of Human-Computer Studies 66 (3), 132–144.
Chang, X., Li, H., Rong, J., Huang, Z., Chen, X., Zhang, Y., 2019. Effects of on-board unit on driving behavior in connected vehicle traffic flow. *Journal of Advanced Transportation* 2019.
Consiglio, W., Driscoll, P., Witte, M., Berg, W., 2003. Effect of cellular telephone conversations and other potential interference on reaction time in a braking response. Accident Analysis and Prevention 35 (4), 495–500.
Doecke, S., Grant, A., Anderson, R., 2015. The real-world safety potential of connected vehicle technology. Traffic Injury Prevention 16, S31–S35.
Eker, U., Ahmed, S.S., Fountas, G., Anastasopoulos, P.C., 2019. An exploratory investigation of public perceptions towards safety and security from the future use of flying cars in the United States. Analytic Methods in Accident Research 23, 100103.
Fountas, G., Anastasopoulos, P., Abdel-Aty, M., 2018. Analysis of accident injury-severities using a correlated random parameters ordered probit approach with time variant covariates. Analytic Methods in Accident Research 18, 57–68.
Ghanipoor Machiani, S., Jahangiri, A., Balali, V., Belt, C., 2017. Predicting driver risky behavior for curve speed warning systems using real field connected vehicle data. 96th Annual Meeting of Transportation Research Board Washington DC, United States. United States.
Greene, W., 2012. LIMDEP Version 10/NLOGIT Version 5. Econometric Modeling Guide.
Hancock, P., Lesch, M., Simmons, L., 2003. The distraction effects of phone use during a crucial driving maneuver. Accident Analysis and Prevention 35 (4), 501–514.
Haque, M., Washington, S., 2014. A parametric duration model of the reaction times of drivers distracted by mobile phone conversations. Accident Analysis and Prevention 62, 42–53.
Haque, M., Washington, S., 2015. The impact of mobile phone distraction on the braking behaviour of young drivers: A hazard-based duration model. Transportation Research Part C 50, 13–27.
Harbluk, J., Noy, Y., Trbovich, P., Eizenman, M., 2007. An on-road assessment of cognitive distraction: Impacts on drivers' visual behavior and braking performance. Accident Analysis and Prevention 39 (2), 372–379.
Hasan, S., Mesa-Arango, R., Ukkusuri, S., 2013. A random-parameter hazard-based model to understand household evacuation timing behavior. Transportation Research Part C 27, 108–116.
Hayat, M., Park, H., Smith, B., 2014. Connected Vehicle Enabled Freeway Merge Assistance system-field test: Preliminary results of driver compliance to advisory. In: IEEE Intelligent Vehicles Symposium Proceedings. IEEE, pp. 1017–1022.
Ishigami, Y., Klein, R., 2009. Is a hands-free phone safer than a handheld phone? Journal of Safety Research 40 (2), 157–164.
Iversen, H., Rundmo, T., 2004. Attitudes towards traffic safety, driving behaviour and accident involvement among the Norwegian public. Ergonomics 47 (5), 555–572.
Jordan, G.A., Anastasopoulos, P.C., Peeta, S., Somenahalli, S., Rogerson, P.A., 2019. Identifying elderly travel time disparities using a correlated grouped random parameters hazard-based duration approach. Research in Transportation Business & Management 30, 100369.
Just, M., Keller, T., Cynkar, J., 2008. A decrease in brain activation associated with driving when listening to someone speak. Brain Research 1205, 70–80.
Karthaus, M., Falkenstein, M., 2016. Functional changes and driving performance in older drivers: assessment and interventions. Geriatrics 1 (2), 12.

Keogh, E., Pazzani, M., 1998. An enhanced representation of time series which allows fast and accurate classification, clustering and relevance feedback. KDD, 239–243.
Lee, J., Park, B., 2012. Development and evaluation of a cooperative vehicle intersection control algorithm under the connected vehicles environment. IEEE Transactions on Intelligent Transportation Systems 13 (1), 81–90.
Lerner, N., Jenness, J., Robinson, E., Brown, T., Baldwin, C. and Llaneras, R. 2011. Crash Warning Interface Metrics Final Report, DOT HS 812 207, National Highway Traffic Safety, Administration, Washington, DC.
Li, X., Yan, X., Wong, S., 2015. Effects of fog, driver experience and gender on driving behavior on S-curved road segments. Accident Analysis and Prevention 77, 91–104.
Li, X., Yan, X., Wu, J., Radwan, E., Zhang, Y., 2016. A rear-end collision risk assessment model based on drivers' collision avoidance process under influences of cell phone use and gender—A driving simulator based study. Accident Analysis and Prevention 97, 1–18.
Mccall, J., Trivedi, M., 2007. Driver behavior and situation aware brake assistance for intelligent vehicles. Proceedings of the IEEE 95 (2), 374–387.
Mcgurrin, M., Vasudevan, M., Tarnoff, P., 2012. Benefits of Dynamic Mobility Applications: Preliminary Estimates from the Literature. Intelligent Transportation Systems Joint Program Office.
Montgomery, J., Kusano, K., Gabler, H., 2014. Age and gender differences in time to collision at braking from the 100-car naturalistic driving study. Traffic Injury Prevention 15, S15–S20.
Olia, A., Abdelgawad, H., Abdulhai, B., Razavi, S., 2016. Assessing the potential impacts of connected vehicles: mobility, environmental, and safety perspectives. Journal of Intelligent Transportation Systems 20 (3), 229–243.
Oviedo-Trespalacios, O., Afghari, A., Haque, M., 2020. A hierarchical Bayesian multivariate ordered model of distracted drivers' decision to initiate risk-compensating behaviour. Analytic Methods in Accident Research 26, 100121.
Pang, J., Krathaus, A., Benedyk, I., Ahmed, S.S., Anastasopoulos, P.C., 2022. A temporal instability analysis of environmental factors affecting accident occurrences during snow events: The random parameters hazard-based duration model with means and variances heterogeneity. Analytic Methods in Accident Research 34, 100215.
Park, H., Bhamidipati, C., Smith, B., 2011. Development and evaluation of enhanced intellidrive-enabled lane changing advisory algorithm to address freeway merge conflict. Transportation Research Record 2243 (1), 146–157.
Rahman, M., Abdel-Aty, M., 2018. Longitudinal safety evaluation of connected vehicles' platooning on expressways. Accident Analysis and Prevention 117, 381–391.
Salvia, E., Petit, C., Champely, S., Chomette, R., Di Rienzo, F., Collet, C., 2016. Effects of age and task load on drivers' response accuracy and reaction time when responding to traffic lights. Frontiers in Aging Neuroscience 8, 169.
Sharma, A., Ali, Y., Saifuzzaman, M., Zheng, Z., Haque, M., 2017. Human factors in modelling mixed traffic of traditional, connected, and automated vehicles. In: International Conference on Applied Human Factors and Ergonomics. Springer, pp. 262–273.
Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, M., 2019. Estimating and comparing response times in traditional and connected environments. Transportation Research Record 2673 (4), 674–684.
Sharma, A., Zheng, Z., Kim, J., Bhaskar, A., Haque, M., 2020. Is an informed driver a better decision maker? A grouped random parameters with heterogeneity-in-means approach to investigate the impact of the connected environment on driving behaviour in safety-critical situations. Analytic Methods in Accident Research 27, 100127.
Simons-Morton, B., Bingham, C., Ouimet, M., Pradhan, A., Chen, R., Barretto, A., Shope, J., 2013. The effect on teenage risky driving of feedback from a safety monitoring system: A randomized controlled trial. Journal of Adolescent Health 53 (1), 21–26.
Simons-Morton, B., Ouimet, M., Wang, J., Klauer, S., Lee, S., Dingus, T., 2009. Hard braking events among novice teenage drivers by passenger characteristics. In: Proceedings of the... International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design. NIH Public Access, p. 236.
Spiegelhalter, D., Best, N., Carlin, B., Van Der Linde, A., 2002. Bayesian measures of model complexity and fit. Journal of the Royal Statistical Society 64 (4), 583–639.
Strayer, D., Drew, F., 2004. Profiles in driver distraction: Effects of cell phone conversations on younger and older drivers. Human Factors 46 (4), 640–649.
Suzuki, K., Kanbe, A., & Oguri, T. (2019). Method to Set the Control Start Timing of Pedestrian AEB System for Preventing Excessive Dependence. International Advance Journal of Engineering Research., 2 (6), 18–28.
Törnros, J., Bolling, A., 2006. Mobile phone use–effects of conversation on mental workload and driving speed in rural and urban environments. Transportation Research Part F 9 (4), 298–306.
Wan, J., Wu, C., Zhang, Y., 2016. Effects of lead time of verbal collision warning messages on driving behavior in connected vehicle settings. Journal of Safety Research 58, 89–98.
Washington, S., Karlaftis, M., Mannering, F., Anastasopoulos, P., 2020. Statistical and Econometric Methods for Transportation Data Analysis. CRC Press.
Wiese, E., Lee, J., 2004. Auditory alerts for in-vehicle information systems: The effects of temporal conflict and sound parameters on driver attitudes and performance. Ergonomics 47 (9), 965–986.
Yadav, A., Velaga, N., 2019. Modelling the relationship between different Blood Alcohol Concentrations and reaction time of young and mature drivers. Transportation Research Part F 64, 227–245.
Yu, R., Xiong, Y., Abdel-Aty, M., 2015. A correlated random parameter approach to investigate the effects of weather conditions on crash risk for a mountainous freeway. Transportation Research Part C 50, 68–77.
Zeng, X., Balke, K., Songchitruksa, P., 2012. Potential connected vehicle applications to enhance mobility, safety, and environmental security. SWUTC/12/161103-1 (ed.).
Zhang, R., Cao, L., Bao, S., Tan, J., 2017. A method for connected vehicle trajectory prediction and collision warning algorithm based on V2V communication. International Journal of Crashworthiness 22 (1), 15–25.
Zhao, J., Yun, M., Zhang, H., Yang, X., 2015. Driving simulator evaluation of drivers' response to intersections with dynamic use of exit-lanes for left-turn. Accident Analysis and Prevention 81, 107–119.
Zöller, I., Abendroth, B., Bruder, R., 2019. Driver behaviour validity in driving simulators–Analysis of the moment of initiation of braking at urban intersections. Transportation Research Part F 61, 120–130.