Analysing and modelling of discretionary lane change duration considering driver heterogeneity

Gen Li, Zhen Yang, Yiyong Pan and Jianxiao Ma

College of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing, People’s Republic of China

ABSTRACT
This paper aims to investigate the characteristics of discretionary lane change (LC) duration on freeways based on an enriched dataset that contains the LC vehicle trajectories of 2905 passenger cars and 433 heavy vehicles. The LC duration is comprehensively analysed, and four stochastic LC duration models are established according to vehicle type and LC direction. LC duration varies with vehicle type and LC direction. Considering driver heterogeneity, accelerated failure time (AFT) models with fixed parameters, latent classes, and random parameters were established in this paper. The results show that drivers of heavy vehicles display greater heterogeneity and that vehicle types and LC directions have significant influence on the LC duration. The results of this study are helpful for understanding the mechanism of the LC process and the influence of LC on traffic flow and improving the safety of lane-changing behaviours of connected and autonomous vehicles.

ARTICLE HISTORY
Received 28 July 2021
Accepted 14 April 2022

KEYWORDS
Lane change duration; driver heterogeneity; Mann–Whitney U test; accelerated failure time model; heavy vehicle

1. Introduction
Lane change (LC) frequently occurs on freeways when drivers need to follow their routes (mandatory lane changes, MLC) or improve their driving conditions (discretionary lane changes, DLC). It is generally believed that the LC manoeuvre has a negative impact on traffic safety and traffic operation (Zheng 2014; Moridpour, Rose, and Sarvi 2010; Kusuma, Liu, and Choudhury 2020) and in some cases is an important cause of traffic breakdown. However, the LC has not received sufficient attention until the last two decades (Moridpour, Rose, and Sarvi 2010; Wang, Li, and Li 2014; Yang, Wang, and Quddus 2019; Ng et al. 2020). It is not easy to perform an LC manoeuvre because controlling the vehicle in both the longitudinal and lateral directions significantly increases the driver’s workload and stress, making driving more error-prone and dangerous (Zheng, Ahn, and Monsere 2010; van Winsum, De Waard, and Brookhuis 1999; Pande and Abdel-Aty 2006). Thus, building an accurate LC model can help establish more accurate microscopic traffic simulation (MTS) and help develop safer and more comfortable advanced driver assistance systems (ADAS).

For quite a long time, LC research was mainly focused on decision modelling. In many studies and MTSs (Finnegan and Green 1990; Knoop et al. 2018; Li 2018), the LC event was considered as an instant event. Since Gipps (1986) established a comprehensive LC framework based on gap acceptance theory, LCs have attracted increasing attention. Gipps’ framework has been widely used in several MTSs, including VISSIM, PARAMICS, and CORSIM (Bloomberg and Dale 2000; PTV-Vision 2011; SiAS 2005).
In these MTSs, various definitions of critical gaps were used to capture LC decision behaviour from different perspectives. Subsequently, some studies reported the inconsistency of gap acceptance theory, namely, when only the lead or the lag gap is greater than the critical gap (or even when neither gap is greater), vehicles may also take LC actions (Sun and Elefteriadou 2012a; Choudhury et al. 2007; Zheng 2014; Chu 2014; Daamen, Loot, and Hoogendoorn 2010; Li and Cheng 2019; Li 2018). Thus, discrete choice models such as logit and probit models were proposed for modelling LC decisions (Marczak, Daamen, and Buisson 2013; Weng et al. 2015; Kita 1993, 1999; Park et al. 2015). To make better use of the available data, data mining techniques such as CART, Bayesian networks and fuzzy logic models were applied to LC decision models (Meng and Weng 2012; Balal, Cheu, and Sarkodie-Gyan 2016; Hou, Edara, and Sun 2014; Gao et al. 2022; Zhang et al. 2021).

The LC process generally lasts for several seconds, during which the vehicle interacts with several surrounding vehicles, making LC execution one of the most dangerous driving behaviours. The literature shows that the LC execution process has a significant impact on traffic flow (Li et al. 2020; Moridpour, Sarvi, and Rose 2010b, 2010a; Cao et al. 2016; Zong et al. 2022). Some studies have focused on acceleration/deceleration behaviours during the lane-changing process and have modified typical car-following models to describe LC execution behaviour (Toledo, Koutsopoulos, and Ben-Akiva 2007; Sarvi and Kuwahara 2007; Moridpour, Sarvi, and Rose 2010b; Wan et al. 2014; Li et al. 2020). These studies regard the LC as a sequential decision process consisting of two steps: (i) decision-making and (ii) LC execution (Wan et al. 2017; Ahmed 1999; Choudhury, Ramanujam, and Ben-Akiva 2009; Sun and Elefteriadou 2012a; Li et al. 2020). It was pointed out by Jin (2010) that the lane-changing vehicle occupies two lanes during the LC process, thereby increasing the traffic density and leading to capacity reduction. As a consequence, LC duration becomes the most important factor that influences the extent of traffic capacity reduction. In this context, it is beneficial to investigate LC duration to understand its influence on traffic flow and to improve the accuracy of MTSs. Different ranges of LC duration have been observed. In general, LC duration ranges from 1 s to more than 10 s, as reported in many studies (Moridpour, Sarvi, and Rose 2010b; Toledo and Zohar 2007; Cao, Young, and Sarvi 2013; Hanowski et al. 2000; Yang, Wang, and Qudus 2019). It was found in these studies that LC duration is affected by many factors, including speed, traffic conditions, and surrounding vehicles. However, none of these studies succeeded in building an LC duration model or in determining how individual factors affect LC duration. Toledo and Zohar (2007) first established an LC duration model based on a log-normal model using the NGSIM dataset and found that the more challenges LC vehicles face, the longer the LC duration is. Wang, Li, and Li (2014) analysed the durations of discretionary LCs and found that speed significantly affected LC duration. Cao, Young, and Sarvi (2013) used a linear regression model based on 192 observational data collected in Australia. However, the NGSIM dataset was collected under conditions of congestion, whereas the data used by Cao, Young, and Sarvi (2013) were collected on an arterial road, and this affects the generality of the findings. Using the NGSIM dataset, Li and Cheng (2019) found that the duration of merging behaviour is differentially distributed under different traffic conditions. Limited by the dataset, it is still not clear whether a log-normal distribution is the best one to describe the duration of lane change at different places. Although the commonly used NGSIM dataset has provided many insightful observations about LCs, that dataset was collected at weaving sections under congested conditions, limiting the generalization of these studies.

Driving behaviour varies with traffic conditions and drivers. Due to the potential heterogeneous effects associated with the missing explanatory factors, it is challenging to accurately model the driving behaviour that accounts for unobserved effects. The heterogeneity of drivers’ lane change manoeuvres was not widely considered until the last decade (Sun and Elefteriadou 2011, 2012b; Keyvan-Ekbatani, Knoop, and Daamen 2016; Daamen, Loot, and Hoogendoorn 2010; Li 2018; Li and Cheng 2019; Li et al. 2019; Li and Sun 2018; Pang et al. 2020). These studies show that drivers adopt various strategies when performing lane-changing manoeuvres. However, it is still not clear how unobserved effects (such as the effects of individual heterogeneity) affect LC duration.

This paper explores the LC duration of vehicles of different types (passenger cars and heavy vehicles) and different LC directions using a correlated grouped random parameter hazard-based duration
model. Hazard-based duration is useful for determining factors that influence the amount of time that elapses between the beginning of an event and the time at which the event ends (i.e. its duration) (Washington et al. 2020) and is very suitable for modelling LC duration. Hazard-based methods have been widely used in transportation activity episode duration analysis (Tilahun and Levinson 2017; Anastasopoulos et al. 2012; Jordan et al. 2019; Hamdar and Mahmassani 2009). With respect to driving behaviours, hazard-based analysis has been used to model the overtaking duration of vehicles and motorcycles (Vlahogianni 2013; Bella and Gulisano 2020). Ali et al. (2019) used the hazard-based duration method to model the minimum gap time during lane-changing in a connected environment. The hazard-based duration model was also used to model the time-to-collision associated with discretionary lane changes in a connected environment (Ali et al. 2020). Recently, Ali, Zheng, and Haque (2021) compared the LC duration in the connected environment with the environment without driving aids based on a grouped random parameters with heterogeneity-in-means hazard-based model and found that the LC duration significantly increased in the connected environment, which can improve the traffic safety. To account for unobserved heterogeneity, latent class and random parameter hazard-based models were employed in this study (Anastasopoulos et al. 2012; Weng, Gan, and Du 2019).

To understand the variation in LC duration under different traffic conditions, we conducted an in-depth analysis of lane change duration in this paper based on an enriched naturalistic vehicle trajectory dataset collected on German highways by unmanned aerial vehicles (UAV). First, we studied the LC durations of vehicles of different types and for different LC directions and establish corresponding models. Second, we applied latent class and random parameter duration models to analyse the explanatory factors that affect lane-changing duration.

2. Data preparation

2.1. The highD dataset

To analyse the LC behaviour of heavy vehicles, the highway drone (highD) dataset is selected in this study. The data were recorded across German highways using drones. A drone to which a high-resolution camera was attached hovered at a fixed position above the highway to collect naturalistic driving data. As shown in Figure 1 (Li, Ma, and Yang 2021), the camera covered a section of the highway approximately 420 metres in length. The highD dataset contains a total of 60 recordings (average length of 17 min) made at 6 different locations near Cologne, Germany. The data were collected between 2017 and 2018 and covered various numbers of lanes and speed limit conditions (e.g. 2 and 3 lanes). Compared with the commonly used NGSIM dataset, the highD dataset provides much more data, especially on heavy trucks. A total of 20,000 trucks were observed in the highD dataset, whereas only 278 trucks were observed in the NGSIM dataset. In addition, compared with the manually created labels, the resulting mean positional errors of the vehicle midpoints in the longitudinal and lateral directions are lower, being less than 3 cm (Krajewski et al. 2018). Thus, this dataset provides more information about the driving behaviour of heavy vehicles. The highD dataset was also used in car-following analysis by Kurtc (2020) with promising results. The speed analysis by Kurtc (2020) showed that although most of the highD dataset is obtained under free flow conditions, due to its large sample size it also contains a considerable number of datasets that record vehicle behaviour during conditions of impeded traffic or even jams with stop-and-go waves. For detailed information on the highD dataset, please refer to Krajewski et al. (2018). The highD dataset can be downloaded from https://www.highd-dataset.com/.

Among the locations at which the data in the highD dataset were collected, the data collected at Location 1 are used in this study for three reasons:

(1) Simplicity of the section: This section is a basic segment of the freeway in which there are 3 lanes in each direction (shown in Figure 1). It is approximately 2 km from the upstream on-ramp and
2 km from the downstream off-ramp. Therefore, there are more LCs of heavy vehicles at Location 1, and these can be generally considered discretionary lane changes (DLCs) due to their location.

(2) Diverse traffic conditions: 37 of the 60 recordings were collected at Location 1, and they cover both free flow and congestion conditions; thus, they provide sufficient data collected under different traffic conditions for analysis.

(3) Popular speed limitations: The speed limit at Location 1 is 120 km/h, a speed limit that is very common in other countries; thus, the data provide insightful and generalizable observations.

2.2. Trajectory extraction

The highD dataset contains details of all vehicles on the highways, including their lateral and longitudinal locations, velocities, accelerations, and surrounding vehicles, making the extraction easier. The steps are as follows:

(1) Extract the trajectories of vehicles that change lanes only once;
(2) For each discretionary LC trajectory, extract the trajectories of the vehicles surrounding the lane-changing vehicle;
(3) Filter out the trajectories that contain incomplete LC processes (some vehicles had initiated lane changing when they first entered the segment or had not finished lane changing when they left the segment).

When these three steps had been completed, the trajectories of 2905 passenger cars and 433 heavy vehicles were extracted from the dataset. Among them, 1350 passenger cars and 213 heavy vehicles change lanes to the left, and 1550 passenger cars and 220 heavy vehicles change lanes to the right. The distribution of different DLCs for passenger cars and heavy vehicles is shown in Table 1. In the table, ‘LCLL’ and ‘LCRL’ represent lane changes to the left lane and to the right lane, respectively, and lanes # 1, 2, and 3 denote the rightmost, middle, and leftmost lanes, respectively. It is found that most heavy vehicles made lane changes between the middle and rightmost lanes, while most passenger cars made lane changes between the middle and leftmost lanes.
Table 1. Numbers of DLC samples.

| Vehicle type     | LC direction | Original lane # | Target lane # | Number | Total |
|------------------|--------------|-----------------|---------------|--------|-------|
| Passenger Car    | LCLL         | 1               | 2             | 312    | 1355  |
|                  |              | 2               | 3             | 1043   |       |
|                  | LCRL         | 2               | 1             | 452    | 1550  |
|                  |              | 3               | 2             | 1098   |       |
| Heavy Vehicle    | LCLL         | 1               | 2             | 198    | 213   |
|                  |              | 2               | 3             | 15     |       |
|                  | LCRL         | 2               | 1             | 205    | 220   |
|                  |              | 3               | 2             | 15     |       |

Note: LCLL = Lane changes to the left lane. LCRL = Lane changes to the right lane.

Figure 2. Definition of LC duration.

The methods used to extract LC duration are based on previous studies (Toledo and Zohar 2007; Yang, Wang, and Quddus 2019). The start and end timepoints of an LC are defined as the times at which the lateral movement of the subject LC vehicle starts and ends, respectively (Toledo, Koutsopoulos, and Ben-Akiva 2007). The duration of the LC is defined as the time from the start of the LC to its endpoints, as shown in Figure 2. For each lane-changing vehicle, the start time is taken as the first time at which the lane-changing vehicle’s lateral velocity is greater than 0.1 m/s; accordingly, the end time is taken as the first time at which the lane-changing vehicle’s lateral velocity is less than 0.1 m/s after it crosses the lane line. Similar methods were adopted in previous studies (Wang, Li, and Li 2014; Li and Cheng 2019; Li and Sun 2018) in which the threshold of the lateral velocity was set as 0.2 m/s. However, considering the large size and the manoeuvrability of heavy vehicles, we use 0.1 m/s in this study. In fact, both 0.1 and 0.2 m/s were tested, and no significant difference was found.

3. Analysis of LC duration

3.1. Descriptive analysis of duration

Table 2 shows a basic statistical analysis of the observed LC duration for passenger cars and heavy vehicles. Figure 3 shows the cumulative distributions of LC duration for passenger cars and heavy vehicles. One can find that LC duration for heavy vehicles is generally longer than LC duration for passenger

| Vehicle type     | LC direction | Number of observations | Mean (s) | Median (s) | Standard Deviation (s) | Minimum (s) | Maximum (s) |
|------------------|--------------|------------------------|----------|------------|------------------------|-------------|-------------|
| Passenger Cars   | Left         | 1355                   | 7.568    | 7.400      | 1.600                  | 3.560       | 15.120      |
|                  | Right        | 1550                   | 7.785    | 7.560      | 1.629                  | 4.320       | 21.600      |
| Heavy Vehicles   | Left         | 213                    | 8.339    | 8.040      | 1.923                  | 4.160       | 15.120      |
|                  | Right        | 220                    | 8.452    | 8.120      | 1.761                  | 5.120       | 14.080      |
Figure 3. Cumulative distributions for LC durations of (a) Passenger cars and (b) Heavy vehicles.

Table 3. Results of Mann–Whitney U tests for different vehicle types and LC directions.

| Hypothesis                              | $P$ value | Results   |
|-----------------------------------------|-----------|-----------|
| Left LC vs. Right LC for Passenger Cars | 0.001     | Reject $H_0$ |
| Left LC vs. Right LC for Heavy Vehicles | 0.363     | Accept $H_0$ |
| Left LC for Passenger Cars vs. Heavy Vehicles | 0.000 | Reject $H_0$ |
| Right LC for Passenger Cars vs. Heavy Vehicles | 0.000 | Reject $H_0$ |

Note: $H_0$: The median values of the two samples are equal.

Table 4. Numbers of samples within different speed ranges and results of Mann–Whitney U tests for heavy vehicles.

|         | [0,20) m/s | [20,25) m/s | [25,30) m/s | [30,35) m/s | Total |
|---------|------------|-------------|-------------|-------------|-------|
| Left    | 28         | 101         | 75          | 9           | 213   |
| Right   | 0          | 85          | 116         | 19          | 220   |
| Total   | 28         | 186         | 191         | 28          | 433   |
| Percentage | 6.47%  | 42.96%      | 44.11%      | 6.47%       | 100%  |
| Mann–Whitney U test ($P$ value) | -         | 0.891        | 0.104        | 0.539       | -     |

cars and that the difference in LC duration between the two directions is small. To investigate whether there are statistically significant differences in the median values of LC duration between different LC directions and different vehicle types, Mann–Whitney U tests were performed. Table 3 shows that the average LC duration is significantly longer for heavy vehicles than for passenger cars. The results in Table 3 also show that there is no significant difference at the 0.95 confidence level in LC duration for heavy vehicles travelling in different directions, whereas this is not the case for passenger cars. This result differs from the results of a previous study that used the NGSIM dataset (Wang, Li, and Li 2014); the authors of that study concluded that direction of travel did not have a significant influence on LC duration. There are three possible reasons for this difference. First, for heavy vehicles, the sample size is small, and that could lead to different results. Second, the traffic conditions under which the NGSIM data were collected were more congested than those under which the highD data were collected, and congestion restricts the movement of vehicles when drivers are changing lanes. Third, heavy vehicle drivers have been found to maintain much more stable speeds during the lane change process, while passenger car drivers adapt their speed more actively (Moridpour, Rose, and Sarvi 2010). In our study, the data were classified into 4 groups according to the subject vehicle’s speed, and Mann–Whitney U tests were performed individually on the data in these 4 groups, as shown in Tables 4 and 5.

Interestingly, The data show that LC direction only has a significant effect on LC duration within the speed ranges of [25, 30) m/s and [30, 35) m/s. This contrasts with the results obtained in a previous
study based on the NGSIM dataset (Wang, Li, and Li 2014). Considering that the NGSIM data were collected under conditions of congestion and transition flow, this finding is not unexpected. When traffic is congested or undergoing transition flow, most passenger cars pursue better driving conditions but are usually restricted by the presence of surrounding vehicles. However, when traffic is undergoing free flow, the motivations of passenger car drivers are more complicated, resulting in a significant difference in LC duration for vehicles travelling in different directions. The speed limit in this study was 120 km/h; thus, passenger cars travelling within the speed range of (35, 45) m/s (126–162 km/h) were exceeding the speed limit. This suggests that these drivers were more aggressive than other drivers and that they were more likely to change lanes to overtake surrounding vehicles. In this situation, LC direction does not significantly affect LC duration. To our knowledge, these findings have not been reported in previous studies.

### 3.2. Effect of speed on LC duration

It was reported by Wang, Li, and Li (2014) that vehicle speed significantly influences LC duration. LC duration tends to saturate as vehicle speed increases. To investigate whether this conclusion remains valid under free flow traffic conditions, the effect of vehicle speed on LC duration is investigated in this section. Considering that a significant influence of LC direction on LC duration was found for passenger cars, the LCs of passenger cars are classified according to LC directions and speed ranges, while the LCs of heavy vehicles are classified only according to speed range. Table 6 shows the statistics for LC duration for passenger cars and heavy vehicles within different speed ranges. Tables 7 and 8 give the results of two-sample Mann–Whitney U tests that were conducted to determine whether there is a significant difference in median LC duration for vehicles travelling at speeds within different speed ranges.

As seen from Table 6, the mean values, median values, and standard deviations of LC duration for passenger cars generally decrease with increasing speed. However, the difference in LC duration for passenger cars travelling in either direction within the speed ranges [0, 20) m/s and [20, 25) m/s is not significantly different. On the other hand, the LC duration for heavy vehicles is relatively consistent across different speed ranges.
Table 7. Results of Mann–Whitney U tests for LC duration for passenger cars within different velocity ranges.

| LC direction | Speed range | (20,25) m/s | (25,30) m/s | (30,35) m/s | (35,45) m/s |
|--------------|-------------|-------------|-------------|-------------|-------------|
| Left         | [0, 20) m/s | 0.754       | 0.174       | 0.023       | 0.029       |
|              | [20, 25) m/s| –           | 0.041       | 0.003       | 0.004       |
|              | [25, 30) m/s| –           | –           | 0.096       | 0.124       |
|              | [30, 35) m/s| –           | –           | –           | 0.677       |
| Right        | [0, 20) m/s | 0.878       | 0.162       | 0.004       | 0.000       |
|              | [20, 25) m/s| –           | 0.201       | 0.005       | 0.000       |
|              | [25, 30) m/s| –           | –           | 0.017       | 0.000       |
|              | [30, 35) m/s| –           | –           | –           | 0.000       |

Table 8. Results of Mann–Whitney U tests for LC duration for heavy vehicles within different velocity ranges.

| Speed range | (0,20) m/s | (20,25) m/s | (25,30) m/s | (30,35) m/s |
|-------------|-------------|-------------|-------------|-------------|
| [0, 20) m/s | –           | 0.049       | 0.020       | 0.000       |
| [20, 25) m/s| –           | –           | 0.778       | 0.008       |
| [25, 30) m/s| –           | –           | –           | 0.005       |

large. It can be observed from Table 7 that most of the $P$ values are below 0.05, indicating that LC duration varies significantly with vehicle speed. The LC duration within the speed range [0, 20) m/s is not significantly different from the LC duration within the speed ranges [20, 25) m/s and [25, 30) m/s. It is also interesting to find that LC duration within the speed range [35, 45) m/s is significantly different from LC duration within the other speed ranges for LCs to the right (LCsR). However, the $P$-value for LCs to the left (LCsL) for vehicles travelling between [35, 45) m/s and [30, 35) m/s is 0.677, indicating that the LC duration for LCsL has reached a saturation value. Determination of the saturation value of LCsR requires more data.

It is clear from Table 6 that the mean and median durations of discretionary LCs made by heavy vehicles generally decrease with increasing vehicle speed; this is confirmed by the results of the Mann–Whitney U tests, as most of the obtained $P$ values are below 0.05 (see Table 4). However, the results also show that there is no significant difference between the LC durations of vehicles travelling [20, 25) m/s and those travelling [25, 30) m/s; the $P$-value is 0.778. However, LC duration further decreases when the speed is faster than 30 m/s. This indicates that vehicle speed does have a significant influence on LC duration for heavy vehicles, and LC duration for heavy vehicles will reach a saturation value in nonfree traffic flow with increasing vehicle speed. However, LC duration will further decrease under ultrafree flow conditions, in contrast to the findings reported in a previous study (Wang, Li, and Li 2014).

3.3. Duration of different LC stages

To further investigate the difference in LC duration for vehicles travelling in different directions, the LC duration was divided into two stages. The first stage, denoted $T_1$, is defined as the time from the starting point of the LC execution process to the point when the subject vehicle just crosses the line that divides the traffic lanes. The second stage, denoted $T_2$, is defined as the time from the end of $T_1$ to the end of the LC execution process, as shown in Figure 2. Table 9 shows the statistics of $T_1$ and $T_2$ for passenger cars and heavy vehicles, and Table 10 shows the results of Mann–Whitney U tests comparing $T_1$ and $T_2$ for passenger cars and heavy vehicles.

For passenger cars, the data in Table 9 show that $T_1$ is much smaller than $T_2$ for LCsL. Although $T_1$ and $T_2$ of the LCsR are similar for the whole sample, it is interesting to find that $T_1$ and $T_2$ of the LCsR differ significantly for all speed ranges. Table 10 also shows that $T_1$ of LCsL differs significantly from $T_1$ of LCsR for passenger cars in all speed ranges. Unlike $T_1$, $T_2$ does not show a significant difference under conditions of congestion and transition flow (when the speed is below 30 m/s) for either direction of travel. This finding suggests that drivers may use different strategies to perform LCs in different
### Table 9. \( T_1 \) and \( T_2 \) statistics for passenger cars and heavy vehicles.

| Vehicle type  | Direction | Duration of stage | Mean(s) | Standard Deviation(s) | Minimum(s) | Median(s) | Maximum(s) |
|---------------|-----------|-------------------|---------|-----------------------|------------|-----------|------------|
| Passenger Cars| Left \( T_1 \) | 3.465             | 1.048   | 1.380                 | 3.220      | 3.220     | 9.540      |
|               | \( T_2 \) | 4.103             | 1.135   | 1.620                 | 3.940      | 3.940     | 8.420      |
|               | Right \( T_1 \) | 3.883             | 1.036   | 1.900                 | 3.660      | 3.660     | 17.660     |
|               | \( T_2 \) | 3.902             | 1.223   | 2.182                 | 3.660      | 4.020     | 9.220      |
| Heavy Vehicles| Left \( T_1 \) | 4.032             | 1.208   | 2.100                 | 3.780      | 4.060     | 9.380      |
|               | \( T_2 \) | 4.307             | 1.329   | 1.860                 | 4.020      | 4.060     | 9.220      |
|               | Right \( T_1 \) | 4.126             | 1.166   | 2.380                 | 3.860      | 4.060     | 9.220      |
|               | \( T_2 \) | 4.327             | 1.204   | 2.500                 | 4.060      | 4.060     | 9.220      |

### Table 10. Results of Mann–Whitney U tests for \( T_1 \) and \( T_2 \).

| Vehicle type  | Hypothesis test | All data | \([0,20)\) m/s | \([20,25)\) m/s | \([25,30)\) m/s | \([30,35)\) m/s | \([35,45)\) m/s |
|---------------|-----------------|----------|----------------|----------------|----------------|----------------|----------------|
| Passenger Cars| Left \( T_1 \) vs. Left \( T_2 \) | 0.000    | 0.000          | 0.000          | 0.000          | 0.000          | 0.000          |
|               | Right \( T_1 \) vs. Right \( T_2 \) | 0.348    | 0.004          | 0.001          | 0.012          | 0.033          | 0.000          |
|               | Left \( T_1 \) vs. Right \( T_1 \) | 0.000    | 0.012          | 0.132          | 0.000          | 0.000          | 0.000          |
|               | Left \( T_2 \) vs. Right \( T_2 \) | 0.000    | 0.84           | 0.583          | 0.815          | 0.000          | 0.000          |
| Heavy Vehicles| Left \( T_1 \) vs. Left \( T_2 \) | 0.018    | 0.010          | 0.083          | 0.910          | 0.156          | –              |
|               | Right \( T_1 \) vs. Right \( T_2 \) | 0.033    | –              | 0.346          | 0.099          | 0.148          | –              |
|               | Left \( T_1 \) vs. Right \( T_1 \) | 0.310    | –              | 0.121          | 0.713          | 1.000          | –              |
|               | Left \( T_2 \) vs. Right \( T_2 \) | 0.597    | –              | 0.335          | 0.108          | 0.884          | –              |

Note: \( H_0 \): The median values of the two samples are equal.

Directions, resulting in different LC durations; this would mean that individual LC models for the two LC directions are needed for passenger cars.

For heavy vehicles, \( T_1 \) is slightly smaller than \( T_2 \) in both directions for the whole sample. However, the results for different speed ranges differ. For LCsL, the difference between \( T_1 \) and \( T_2 \) becomes insignificant when the vehicle’s speed is greater than 25 m/s. In contrast, the \( P \)-value for LCsR within the speed range \([25,30)\) m/s is 0.099. This means that heavy vehicle drivers may adopt different strategies to perform LCs in different directions under different traffic conditions. Table 10 shows that \( T_1 \) and \( T_2 \) do not differ significantly for the two directions, indicating the consistency of heavy vehicles’ LCs in different directions. The results of Mann–Whitney U tests show that \( T_1 \) is significantly shorter than \( T_2 \) for both directions; this means that heavy vehicles generally need more time to adjust their position and speed after encroaching on the target lane.

### 4. Methodology

#### 4.1. Accelerated failure time models

Let \( T \) be a nonnegative random variable representing the duration of an activity (such as a lane change), and let it have a continuous probability distribution \( f(t) \) in which \( t \) is the realization of \( T \). The cumulative probability of \( t \) can be obtained by

\[
F(t) = \int_0^t f(t) dt = P(T \leq t)
\]

The survival function of \( T \), indicating that the duration of the activity is at least \( t \), should be

\[
S(t) = \text{Prob}(P \geq t) = 1 - F(t) = \int_t^\infty f(t) dt
\]
The hazard function, \( h(t) \), is defined as the conditional probability that an activity ends between \( t \) and \( t + \Delta t \). Given that the activity does not end before time \( t \), it can be expressed as

\[
h(t) = \lim_{\Delta t \to 0} \frac{P(t \leq T \leq t + \Delta t | T \geq t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{F(t + \Delta t) - F(t)}{\Delta t S(t)} = \frac{f(t)}{S(t)} = -\frac{d}{dt} \log S(t) \tag{3}
\]

\( h(t) \) can also be interpreted as the instantaneous ‘failure’ rate at time \( T = t \) upon survival to time \( t \). Two methods can be used to investigate the influence of covariates. One is the proportional hazards (PH) model, and the other is the accelerated failure time (AFT) model. PH models assume that the covariates act proportionally on the underlying hazard function and ignore the relationship between the covariates and the target variable. Unlike PH models, AFT models assume that the effect of covariates is to directly accelerate or decelerate the lifetime. Since the estimated parameters quantify the corresponding effect of a covariate on the mean survival time, the interpretation of the results is much simpler when an AFT model is used (Patel, Kay, and Rowell 2006; Haque and Washington 2014). Thus, the AFT model is applied in this study.

It is assumed in AFT models that the joint effects of the covariates are \( \theta = \exp(\beta X) \). Then, the hazard function can be written as

\[
h(t|\theta) = \theta h_0(t\theta) \tag{4}
\]

An AFT model can also be written as

\[
\ln T = \beta X + \varepsilon \tag{5}
\]

where \( \beta \) is an unknown coefficient vector associated with the covariates vector \( X \), and \( \varepsilon \) is an error variable. Different distributions of \( \varepsilon \) can produce different baseline distributions of the survival time. Weibull, log-normal, exponential, gamma, log-logistic, and Gompertz distributions are commonly used in AFT models (Washington et al. 2020; Li, Ma, and Yang 2021). Previous studies have shown that the log-normal distribution can fit lane-changing duration data well (Cao, Young, and Sarvi 2013; Toledo and Zohar 2007). However, Li and Cheng (2019) found that the optimal distribution varies with the traffic conditions. Therefore, different distributions (Weibull, log-normal, and log-logistic) were compared to find the best-fitting form in this study.

Fixed-parameter AFT models can be affected by unobserved heterogeneity such as heterogeneity of driver characteristics. To incorporate unobserved heterogeneity and better describe the stochastic characteristics of lane-changing duration, a random parameter model is applied in this paper. The observational specific parameter vectors of the explanatory variables are estimated in a random parameters model to capture the underlying variations. The specific parameter vector of observation \( i \) can be expressed as

\[
\beta_i = \beta + \omega_i \tag{6}
\]

where \( \beta \) is the mean value of the random parameter vector and the \( \omega_i \) random term follows the standard normal distribution (Fountas, Anastasopoulos, and Abdel-Aty 2018).

Another way to account for driver heterogeneity is the latent class model, which segments the population into several subpopulations or components (McLachlan and Peel 2000; Fraley and Raftery 2002; Greene 2003; Greene and Hensher 2003; Li et al. 2019; Li, Ma, and Yang 2021). A \( K \)-class AFT has the following form:

\[
\ln T = \sum_{k=1}^{K} \pi_k (\beta_k X + \varepsilon_k) \sum_{k=1}^{K} \pi_k = 1 \tag{7}
\]

where \( \pi_k \) is the proportion of latent class \( k \), \( \beta_k \) is the component-specific parameter vector, and \( \varepsilon_k \) is the error term for class \( k \). We can find that the driver heterogeneity is assumed to be distributed discretely, which is not the case in the random-parameter model.
4.2. Estimation methods

Unlike fixed-parameter AFT models, random parameter and latent class AFT models cannot be directly estimated using the maximum likelihood estimation (MLE) method. However, the random-parameter AFT model can be estimated by a simulation-based MLE technique. Halton sequence draws are used in this study because this technique has been proved to outperform random draws and to produce a more efficient simulation process (Fountas, Anastasopoulos, and Abdel-Aty 2018; Fountas et al. 2018; Anastasopoulos et al. 2012; Jordan et al. 2019; Greene 2016). Previous studies have shown that 200 Halton draws are sufficient (Anastasopoulos et al. 2012; Weng, Gan, and Du 2019); thus, we use 200 Halton draws in this study. It should be noted that not all parameters are considered random parameters. A parameter is considered random only when both its mean and its variance are significant at the 90% confidence level and when inclusion of that parameter reduces the Akaike information criterion (AIC) value compared to the model with fixed parameters. The AIC is defined as

\[
AIC = 2N - 2 \ln L
\]  

where \(N\) is the number of parameters. Please refer to Greene (2003) for a detailed description of the process.

The latent class model can be efficiently estimated by the EM algorithm (Dempster, Laird, and Rubin 1977), which alternates between the expectation and maximization steps until the likelihood of improvement falls below a prespecified threshold or a maximum number of iterations is reached. In this paper, latent class models with numbers of classes ranging from 1 to 4 will be estimated, and the model with the lowest Bayesian information criterion (BIC) will be chosen as the best (Greene and Hensher 2003) because the estimation of the latent class AFT model was a nonconvex optimization. The BIC can be calculated as follows:

\[
BIC = -2LL + \gamma \log(N)
\]

where \(LL\) is the log-likelihood value, \(\gamma\) is the number of free parameters to be estimated, and \(N\) is the number of observations in the dataset.

5. Model estimation results

5.1. Parameter estimation

As summarized in previous studies, LC duration can be influenced by several factors, as shown in Table 11 (Cao, Young, and Sarvi 2013; Toledo and Zohar 2007; Ali et al. 2019).

The log-normal, log-logistic, Weibull, and Gamma distributions were tested in fixed-parameter, random parameter, and latent class AFT models. The estimation results showed that the log-normal

| Candidate variable | Description |
|--------------------|-------------|
| \(V_s\) (m/s)      | Speed of the subject LC vehicle. |
| \(DV_l\) (m/s)     | The speed difference between the subject LC vehicle and its leading vehicle in the original lane. |
| \(DV_f\) (m/s)     | The speed difference between the subject LC vehicle and its following vehicle in the original lane. |
| \(DV_{PL}\) (m/s)  | The speed difference between the subject LC vehicle and its putative leading (PL) vehicle in the target lane. |
| \(DV_{PF}\) (m/s)  | The speed difference between the subject LC vehicle and its putative following (PF) vehicle in the target lane. |
| \(Gap\) (m)        | The space gap between the PL and PF vehicles in the target lane. |
| \(Gap_l\) (m)      | The space gap between the leading vehicle and the subjective vehicle in the original lane. |
| \(T_L\) (m/s)      | The dummy variable of the leading vehicle in the original lane (0 for passenger car and 1 for heavy vehicle). |
| \(T_F\) (m/s)      | The dummy variable of the following vehicle in the original lane (0 for passenger car and 1 for heavy vehicle). |
| \(T_{PL}\) (m/s)   | The dummy variable of the PL vehicle in the target lane (0 for passenger car and 1 for heavy vehicle). |
| \(T_{PF}\) (m/s)   | The dummy variable of the PF vehicle in the target lane (0 for passenger car and 1 for heavy vehicle). |
distribution outperformed the other distributions in all three models, consistent with previous studies (Toledo and Zohar 2007; Wang, Li, and Li 2014). Thus, we only present the estimation results of log-normal AFT models in the remainder of this paper.

Figure 4 shows the BIC values of latent class models with numbers of classes ranging from 1 to 8. From the BIC values of the four models, it can be seen that the latent class AFT models do not better fit the lane change durations than do the fixed-parameter AFT models. This means that latent class AFT models are not suitable for describing the lane change durations; that is, there is no obvious discrete heterogeneity among lane change durations. Thus, latent class AFT models will not be discussed in the remainder of this paper.

Tables 12 and 13 present the model estimation results of the fixed-parameter and random-parameter AFT models for passenger cars and heavy vehicles, respectively. All competitive models were estimated using the same explanatory variables for comparison purpose. In Table 13, according to the log-likelihood and AIC values, the random-parameter AFT models of heavy vehicles are superior to the fixed-parameter models. For passenger cars, however, they are not. Both the log-likelihood and AIC values show that the random-parameter AFT models cannot outperform the fixed-parameter models for passenger cars, especially when the cars are moving to the right lanes. This means that the heterogeneity of LC duration is greater for heavy vehicles than for passenger cars.

5.2. Discussion

Tables 12 and 13 show that the parameters of the random-parameter AFT models are similar to those of the fixed-parameter models. However, the random-parameter AFT models can provide more information than the fixed-parameter models. Thus, this section will discuss the results of the random-parameter AFT models.

For passenger cars (please see Table 12), both $Gap_L$ and the constant term have random parameters in the LCSL model, and only the constant term has a random parameter in the LCSR model, simplifying the random-parameter model to the random effect model. In addition, the standard deviation of the constant term of LCSR is much smaller than that of LCSL. These results indicate that for passenger cars LCSR is more consistent than LCSL. This conclusion is consistent with the analysis presented in Section 4; that is, the subject vehicle’s speed has a negative impact on LC duration in both directions, a finding that is also consistent with several previous studies (Wang, Li, and Li 2014; Toledo and Zohar 2007). The LC duration in both directions decreases with increasing $DV_L$. $Gap_L$ has a positive effect on LC duration in both directions, indicating that with a larger $Gap_L$, the subject vehicles have more space and time in which to change lanes, resulting in longer LC durations. In addition, from Table 12, we can also find
Table 12. Estimation results of the models for LC duration of passenger cars.

| LC to the Left | LC to the Right |
|----------------|-----------------|
| **Random parameters AFT model** | **Random-parameter AFT model** |
| Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat | Coeff. | t-stat |
|---|---|---|---|---|---|---|---|
| **Constant** | 2.1557 | 75.17 | 2.1889 | 1876.39 | 2.2753 | 106.20 | 2.2745 | 95.13 |
| **SD of parameter** | – | – | 0.2023 | 1082.28 | – | – | 0.0534 | 9.05 |
| **VS** | –0.0053 | –5.17 | –0.0066 | –163.69 | –0.0095 | –10.79 | –0.0093 | –9.22 |
| **DVPL** | 0.0073 | 3.53 | 0.0093 | 133.72 | –0.0002 | –1.81 | –0.0026 | –1.65 |
| **GapL** | 0.0017 | 7.32 | 0.0021 | 231.67 | 0.0005 | 4.29 | 0.0004 | 2.78 |
| **SV** | –0.0053 | –5.17 | –0.0066 | –163.69 | –0.0095 | –10.79 | –0.0093 | –9.22 |
| **DVPL** | 0.0073 | 3.53 | 0.0093 | 133.72 | –0.0002 | –1.81 | –0.0026 | –1.65 |
| **GapL** | 0.0017 | 7.32 | 0.0021 | 231.67 | 0.0005 | 4.29 | 0.0004 | 2.78 |

that **GapPL** has positive effects on the duration of LCsR. However, it is interesting that the presence of a leading heavy vehicle in the current lane will increase the LC duration, while a leading heavy vehicle in the target lane will cause the LC vehicle to complete the LC manoeuvre earlier for LCsL. Another interesting finding is that the estimates of **DVPL** for different LC directions are opposite. This may be due to different motivations for LCs and to differences in traffic conditions, as in many studies, the left lane is generally considered to be a faster lane (Moridpour, Mazloumi, and Mesbah 2015; Moridpour, Rose, and Sarvi 2010; Moridpour, Sarvi, and Rose 2010b).

The estimation results of the models for heavy vehicles are shown in Table 13. The last section shows that the distributions of the heavy vehicles’ LC durations for different directions are not significantly different, but the estimation results show that the influencing variables are significantly different.
These results are not contradictory. The results mean that the lane change durations of heavy vehicles are not significantly influenced by the direction of the lane change, but the influences of the variables on different individuals are heterogeneous. Similar to the findings for passenger cars, \( V_f \) has a negative influence on LC duration in both directions, and the parameters in both directions are random. \( T_F \) is another parameter that affects LC duration in both directions. For LCsL, LC duration will decrease when the following vehicle is a heavy vehicle. However, we find that although the parameter \( T_F \) for LCsR is 0.0121 (positive), the standard deviation of the parameter is approximately 3 times the magnitude of the parameter, indicating that the effect of \( T_F \) on the duration of LCsR is significantly heterogeneous. However, \( DVL \) and \( Gap_L \) only affect the duration of LCsL. The parameter \( DVL \) is a fixed negative value, indicating that LC duration will be reduced if the LC vehicle moves faster than the leading vehicle. The positive sign of the parameter \( Gap_L \) indicates that as \( Gap_L \) increases, the LC durations of heavy vehicles will also be increased. \( Gap \) and \( T_{PI} \) are significant only in the LCsR model. The parameters \( Gap \) and \( T_{PI} \) are both positive and random, indicating that with an increase in \( Gap \) and when the leading vehicle in the target lane is a heavy vehicle, the subject heavy vehicle will increase its LC duration. Similar to the parameter of \( T_F \), the standard deviations of \( Gap \) and \( T_{PI} \) are also larger than the parameters themselves, indicating that LC duration is more heterogeneous and complicated when heavy vehicles change lanes to the right.

Comparing the estimation results of the models for passenger cars and heavy vehicles, it is clear that the factors that influence the LC durations of heavy vehicles are more heterogeneous than those that influence the LC durations of passenger cars. This may be due to the difference in mobility of these two types of vehicles. Previous studies have shown that heavy vehicles cannot adjust their speeds as actively as passenger cars can (Moridpour, Rose, and Sarvi 2010; Moridpour, Mazloumi, and Mesbah 2015), resulting in different interactions between heavy vehicles and surrounding vehicles when changing lanes.

6. Concluding remarks

To explore the characteristics of the durations of discretionary LCs, an LC vehicle trajectory dataset containing LC vehicle trajectories of 2905 passenger cars and 433 heavy vehicles was used in this study. First, LC duration was comprehensively analysed based on vehicle type, LC direction, and vehicle speed. It was found that more time is required for heavy vehicles to complete the LC manoeuvre. LC direction significantly influences LC duration for passenger cars, while it has almost no influence on LC duration for heavy vehicles in this dataset. Vehicle speed has an important effect on LC duration. However, this influence varies with vehicle type and LC direction. In general, the LC duration of heavy vehicles is more easily affected by vehicle speed than is the LC duration of passenger cars. For passenger cars, the duration of LCsR is more easily affected by vehicle speed than is the duration of LCsL. Further analysis of LC durations at different stages shows that when changing lanes to the left, passenger cars tend to use less time to encroach on the target lane, while heavy vehicles take more time to adjust their speed and position after encroaching on the target lane. However, similar amounts of time were used when passenger car drivers changed lanes to the right. This means that drivers may use different strategies to perform LCs when they change lanes in different directions. Unlike drivers of passenger cars, drivers of heavy vehicles used less time to encroach on the target lane in both directions.

Second, in our study, fixed-parameter AFT models, latent class AFT models, and random-parameter AFT models were used to fit LC duration for different vehicle types and LC directions. Interestingly, the fixed-parameter AFT models can better simulate the lane change duration of passenger cars, while the random-parameter AFT models outperform other models for heavy vehicles. This indicates that heavy vehicles show more heterogeneity when changing lanes than do passenger cars, a fact that is mainly due to the poor manoeuvrability of heavy vehicles. The results also revealed that LC duration for vehicles of the same type depends on different covariates in different LC directions. In addition, the
estimated standard deviations of the random parameters show that the heterogeneity of LC duration usually increases when heavy vehicles move to the right lane.

Finally, the results of this study confirmed that LC duration varies with vehicle type and LC direction. In addition, the results of this study are helpful in allowing us to better understand the mechanism of LC and its impact on traffic flow. The proposed LC duration model indicates that LCs need to be carefully handled, especially in the case of heavy vehicles. Therefore, heterogeneity cannot be ignored when developing LC assistance systems for heavy vehicles. The results of this paper can benefit the modelling of lane-changing behaviours of connected and autonomous vehicles as the LC duration is an important parameter that may affect traffic safety.

However, this paper still has some limitations. First, the data were collected at only one location at which there are 3 lanes and did not consider the effect of the road configurations, such as the number of lanes in this study. In addition, the driving rules, laws, and behaviour in Germany are different from those in the USA, where the NGSIM dataset was collected. Additional data will be collected in the future to verify the generality of the conclusions and the model. In addition, more insight is needed into how the lane change direction affects the lane change durations of heavy vehicles by collecting more data. We will also investigate LC durations of intelligent connected vehicles in the future.

Data availability

Some or all data, models, or code generated or used during the study are available in a repository or online in accordance with funder data retention policies. The data used to support the findings of this study have been deposited at https://www.highd-dataset.com/.

Acknowledgements

The authors would like to thank Robert Krajewski and Julian Bock (RWTH Aachen University, Aachen, Germany) for providing the highD trajectory data for this study.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the Natural Science Fund for Colleges and Universities in Jiangsu Province under [grant number 21KJB580014], the Science and Technology Innovation Fund for Youth Scientists of Nanjing Forestry University under [grant number CX2019021], and Basic Research Program of Science and Technology Commission Foundation of Jiangsu Province under [grant number BK20170932].

ORCID

Gen Li http://orcid.org/0000-0001-5535-6467
Zhen Yang http://orcid.org/0000-0001-6391-7260
Yiyong Pan http://orcid.org/0000-0002-2435-970X
Jianxiao Ma http://orcid.org/0000-0002-4324-8175

References

Ahmed, Kazi Iftekhar. 1999. “Modeling Drivers’ Acceleration and Lane Changing Behavior.” Massachusetts Institute of Technology.

Ali, Yasir, Md Mazharul Haque, Zuduo Zheng, Simon Washington, and Mehmet Yildirimoglu. 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, Yasir, Zuduo Zheng, and Md Mazharul Haque. 2021. “Modelling Lane-Changing Execution Behaviour in a Connected Environment: A Grouped Random Parameters with Heterogeneity-in-Means Approach.” Communications in Transportation Research 1: 100009.
Ali, Yasir, Zuduog Zheng, Md Mazharul Haque, Mehmet Yildirimoglu, and Simon Washington. 2020. “Understanding the Discretionary Lane-Changing Behaviour in the Connected Environment.” Accident Analysis & Prevention 137: 105463.

Anastasopoulos, Panagiots Ch, Mouyid Bin Islam, Dionysia Perperidou, and Matthew G. Karlaftis. 2012. “Hazard-Based Analysis of Travel Distance in Urban Environments: Longitudinal Data Approach.” Journal of Urban Planning Development 138 (1): 53–61.

Balal, Esmaeil, Ruey Long Cheu, and Thompson Sarkodie-Gyan. 2016. “A Binary Decision Model for Discretionary Lane Changing Move Based on Fuzzy Inference System.” Transportation Research Part C: Emerging Technologies 67: 47–61.

Bella, Francesco, and Federico Gulisano. 2020. “A Hazard-Based Model of the Motorcyclists’ Overtaking Duration.” Accident Analysis Prevention 141: 105522.

Bloomberg, Loren, and Jim Dale. 2000. “A Comparison of the VISSIM and CORSIM Traffic Simulation Models.” Paper presented at the Institute of Transportation engineers Annual meeting. Washington, DC, USA.

Cao, X., W. Young, and M. Sarvi. 2013. “Exploring Duration of Lane Change Execution.” Paper presented at the australasian Transport Research forum, Brisbane, Queensland, Australia.

Cao, Xiaoying, William Young, Majid Sarvi, and Inhi Kim. 2016. “Study of Mandatory Lane Change Execution Behavior Model for Heavy Vehicles and Passenger Cars.” Transportation Research Record 2561 (1): 73–80.

Choudhury, Charisma F, Moshe E Ben-Akiva, Tomer Toledo, Gunwoo Lee, and Anita Rao. 2007. “Modeling Cooperative Lane Changing and Forced Merging Behavior.” Paper presented at the 86th Annual Meeting of the Transportation Research board, Washington, DC, USA.

Choudhury, Charisma, Varun Ramanujam, and Moshe Ben-Akiva. 2009. “Modeling Acceleration Decisions for Freeway Merges.” Transportation Research: Journal of the Transportation Research Board 2124: 45–57.

Chu, Tien Dung. 2014. “A Study on Merging Behavior at Urban Expressway Merging Sections.” Nagoya University.

Daamen, Winnie, Martijn Loot, and Serge P Hoogendoorn. 2010. “Empirical Analysis of Merging Behavior at Freeway on-Ramp.” Transportation Research Record 2188 (1): 108–118.

Dempster, Arthur P, Nan M Laird, and Donald B Rubin. 1977. “Maximum Likelihood from Incomplete Data via the EM Algorithm.” Journal of the Royal Statistical Society. Series B (Methodological) 39 (1): 1–22.

Finnegan, Paula, and Paul Green. 1990. “The Time to Change Lanes: A Literature Review.” In.: University of Michigan, Transportation Research Institute.

Fountas, Grigoris, Panagiots Ch Anastasopoulos, and Mohamed Abdel-Aty. 2018a. “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.

Fountas, Grigoris, Md Tawfiqu Sarwar, Panagiots Ch Anastasopoulos, Alan Blatt, and Kevin Majka. 2018. “Analysis of Stationary and Dynamic Factors Affecting Highway Accident Occurrence: A Dynamic Correlated Grouped Random Parameters Binary Logit Approach.” Accident Analysis Prevention 113: 330–340.

Fraleys, Chris, and Adrian E Raftery. 2002. “Model-based Clustering, Discriminant Analysis, and Density Estimation.” Journal of the American Statistical Association 97 (458): 611–631.

Gao, Jun, Yi Lu Murphey, Jiangang Yi, and Honghui Zhu. 2022. “A Data-Driven Lane-Changing Behavior Detection System Based on Sequence Learning.” Transportmetrica B: Transport Dynamics 10 (1): 1–18.

Gipps, Peter G. 1986. “A Model for the Structure of Lane-Changing Decisions.” Transportation Research Part B: Methodological 20 (5): 403–414.

Greene, William H. 2003. Econometric Analysis. New Jersey: Pearson Education Inc.

Greene, W. 2016. LIMDEP Ver11 Reference Guide. Plainview, NY: Econometric.

Greene, William H, and David A Hensher. 2003. “A Latent Class Model for Discrete Choice Analysis: Contrasts with Mixed Logit.” Transportation Research Part B: Methodological 37 (8): 681–698.

Hamdar, Samer H, and Hani S. Mahmassani. 2009. “Life in the Fast Lane:Duration-Based Investigation of Driver Behavior Differences Across Freeway Lanes.” Transportation Research Record 2124 (1): 89–102. doi:10.3141/2124-09.

Hanowski, Richard J, Walter W Wierwille, Sheila A Garness, Thomas A Dingus, Ronald R Knipling, and Robert J Carroll. 2000. “A Field Evaluation of Safety Issues in Local/Short Haul Trucking.” Paper presented at the proceedings of the human factors and ergonomics society annual meeting.

Haque, Md Mazharul, and Simon Washington. 2014. “A Parametric Duration Model of the Reaction Times of Drivers Distracted by Mobile Phone Conversations.” Accident Analysis Prevention 62: 42–53.

Hou, Yi, Praveen Edara, and Carlos Sun. 2014. “Modeling Mandatory Lane Changing Using Bayes Classifier and Decision Trees.” IEEE Transactions on Intelligent Transportation Systems 15 (2): 647–655.

Jin, Wen-Long. 2010. “A Kinematic Wave Theory of Lane-Changing Traffic Flow.” Transportation Research Part B: Methodological 44 (8-9): 1001–1021.

Jordan, Gary A, Panagiots Ch Anastasopoulos, Srinivas Peeta, Sekhar Somenahalli, and Peter A Rogerson. 2019. “Identifying Elderly Travel Time Disparities Using a Correlated Grouped Random Parameters Hazard-Based Duration Approach,” Research in Transportation Business Management 30: 100369.

Keyvan-Ekbatani, Mehdi, Victor L Knoop, and Winnie Daamen. 2016. “Categorization of the Lane Change Decision Process on Freeways.” Transportation Research Part C: Emerging Technologies 69: 515–526.

Kita, Hideyuki. 1993. “Effects of Merging Lane Length on the Merging Behavior at Expressway on-Ramps.” Proceedings of the 12th International Symposium on the Theory of Traffic Flow and Transportation, 37–51.
Kita, Hideyuki. 1999. "A Merging–Giveaway Interaction Model of Cars in a Merging Section: A Game Theoretic Analysis." Transportation Research Part A: Policy and Practice 33 (3-4): 305–312.

Knoop, V. L., M. Keyvan-Ekbatani, Marco de Baat, H. Taale, and S. P. Hoogendoorn. 2018. "Lane Change Behavior on Freeways: An Online Survey Using Video Clips." Journal of Advanced Transportation 2018.

Krajewski, Robert, Julian Bock, Laurent Kloeker, and Lutz Eckstein. 2018. "The Highd Dataset: A Drone Dataset of Naturalistic Vehicle Trajectories on German Highways for Validation of Highly Automated Driving Systems." Paper presented at the 2018 21st international conference on Intelligent Transportation Systems (ITSC).

Kurtc, Valentina. 2020. "Studying Car-Following Dynamics on the Basis of the HighD Dataset." Transportation Research Record On Line. doi:10.1177/0361198120925063.

Kusuma, Andyka, Ronghui Liu, and Charisma Choudhury. 2020. "Modelling Lane-Changing Mechanisms on Motorway Weaving Sections." Transportmetrica B: Transport Dynamics 8 (1): 1–21.

Li, Gen. 2018. "Application of Finite Mixture of Logistic Regression for Heterogeneous Merging Behavior Analysis." Journal of Advanced Transportation 2018.

Li, Gen, and Juan Cheng. 2019. "Exploring the Effects of Traffic Density on Merging Behavior." IEEE Access 7: 51608–51619.

Li, Gen, Song Fang, Jianxiao Ma, and Juan Cheng. 2020. "Modeling Merging Acceleration and Deceleration Behavior Based on Gradient-Boosting Decision Tree." Journal of Transportation Engineering, Part A: Systems 146 (7): 05020005.

Li, Gen, Jianxiao Ma, and Zhen Yang. 2021a. "Characteristics of Heavy Vehicle Discretionary Lane Changing Based on Trajectory Data." Transportation Research Record. doi:10.1177/03611981211051337.

Li, Gen, Yiyong Pan, Zhen Yang, and Jianxiao Ma. 2019. "Modeling Vehicle Merging Position Selection Behaviors Based on a Finite Mixture of Linear Regression Models." IEEE Access 7: 158445–58.

Li, Gen, and Lu Sun. 2018. "Characterizing Heterogeneity in Drivers’ Merging Maneuvers Using Two-Step Cluster Analysis." Journal of Advanced Transportation 2018.

Marczak, Florian, Winnie Daamen, and Christine Buisson. 2013. "Merging Behaviour: Empirical Comparison Between two Sites and new Theory Development." Transportation Research Part C: Emerging Technologies 36: 530–546.

McClellan, Geoffrey, and David Peeli. 2000. Finite Mixture Models, Willey Series in Probability and Statistics. New York: John Wiley & Sons.

Meng, Qiang, and Jinxian Weng. 2012. "Classification and Regression Tree Approach for Predicting Drivers’ Merging Behavior in Short-Term Work Zone Merging Areas." Journal of Transportation Engineering 138 (8): 1062–1070.

Moridpour, Sara, Ehsan Mazloumi, and Mahmoud Mesbah. 2015. "Impact of Heavy Vehicles on Surrounding Traffic Characteristics." Journal of Advanced Transportation 49 (4): 535–552.

Moridpour, Sara, Geoff Rose, and Majid Sarvi. 2010. "Effect of Surrounding Traffic Characteristics on Lane Changing Behavior." Journal of Transportation Engineering 136 (11): 973–985.

Moridpour, Sara, Majid Sarvi, and Geoff Rose. 2010a. "Lane Changing Models: A Critical Review." Transportation Letters 2 (3): 157–173.

Moridpour, Sara, Majid Sarvi, and Geoff Rose. 2010b. "Modeling the Lane-Changing Execution of Multiclass Vehicles Under Heavy Traffic Conditions." Transportation Research Record: Journal of the Transportation Research Board 2161: 11–19.

Ng, Christina, Susilawati Susilawati, Md Abdus Samad Kamal, and Irene Mei Leng Chew. 2020. "Development of a Binary Logistic Lane Change Model and its Validation Using Empirical Freeway Data." Transportmetrica B: Transport Dynamics 8 (1): 49–71.

Pande, Anurag, and Mohamed Abdel-Aty. 2006. "Assessment of Freeway Traffic Parameters Leading to Lane-Change Related Collisions." Accident Analysis & Prevention 38 (5): 936–948.

Pang, Meng-Yuan, Bin Jia, Dong-Fan Xie, and Xin-Gang Li. 2020. "A Probability Lane-Changing Model Considering Memory Effect and Driver Heterogeneity." Transportmetrica B: Transport Dynamics 8 (1): 72–89.

Park, Minju, Kitae Jang, Jinwoo Lee, and Hwasoo Yeo. 2015. "Logistic Regression Model for Discretionary Lane Changing Under Congested Traffic." Transportmetrica A: Transport Science 11 (4): 333–344.

Patel, Katie, Richard Kay, and Lucy Rowell. 2006. "Comparing Proportional Hazards and Accelerated Failure Time Models: An Application in Influenza." Pharmaceutical Statistics: The Journal of Applied Statistics in the Pharmaceutical Industry 5 (3): 213–224.

PTV-Vision. 2011. VISSIM 5.30.05 User Manual. Karlsruhe.

Sarvi, Majid, and Masao Kuwahara. 2007. "Microsimulation of Freeway Ramp Merging Processes Under Congested Traffic Conditions." IEEE Transactions on Intelligent Transportation Systems 8 (3): 470–479.

Rafael, S. 2005. S-Paramics 2005 Reference Manual. Edinburgh, UK: SIAS Ltd.

Sun, Daniel Jian, and Lily Elefteriadou. 2011. "Lane-changing Behavior on Urban Streets: A Focus Group-Based Study." Applied Ergonomics 42 (5): 682–691.

Sun, Daniel, and Lily Elefteriadou. 2012a. "A Driver Behavior-Based Lane-Changing Model for Urban Arterial Streets." Transportation Science 48 (2): 184–205.

Sun, Daniel, and Lily Elefteriadou. 2012b. "Lane-Changing Behavior on Urban Streets: An “in-Vehicle” Field Experiment-Based Study." Computer-Aided Civil Infrastructure Engineering 27 (7): 525–542.

Tilahun, Nebiyou, and David Society Levinson. 2017. "Contacts and Meetings: Location, Duration and Distance Traveled." Travel Behaviour and Society 6: 64–74.
Toledo, Tomer, Haris N Koutsopoulos, and Moshe Ben-Akiva. 2007. “Integrated Driving Behavior Modeling.” *Transportation Research Part C: Emerging Technologies* 15 (2): 96–112.

Toledo, Tomer, and David Zohar. 2007. “Modeling Duration of Lane Changes.” *Transportation Research Record* 1999 (1): 71–78.

van Winsum, Wim, Dick De Waard, and Karel A Brookhuis. 1999. “Lane Change Manoeuvres and Safety Margins.” *Transportation Research Part F: Traffic Psychology and Behaviour* 2 (3): 139–149.

Vlahogianni, Eleni I. 2013. “Modeling Duration of Overtaking in two Lane Highways.” *Transportation Research Part F: Traffic Psychology Behaviour* 20: 135–146.

Wan, Xia, Peter J. Jin, Haiyan Gu, Xiaoxuan Chen, and Bin Ran. 2017. “Modeling Freeway Merging in a Weaving Section as a Sequential Decision-Making Process.” *Journal of Transportation Engineering, Part A: Systems* 143 (5), doi:10.1061/jtepbs.0000048.

Wan, Xia, Peter Jin, Fan Yang, Jian Zhang, and Bin Ran. 2014. “Modeling Vehicle Interactions During Merge in Congested Weaving Section of Freeway Ramp.” *Transportation Research Record: Journal of the Transportation Research Board* 2421: 82–92.

Wang, Qi, Zhiheng Li, and Li Li. 2014. “Investigation of Discretionary Lane-Change Characteristics Using Next-Generation Simulation Data Sets.” *Journal of Intelligent Transportation Systems* 18 (3): 246–253.

Washington, Simon, Matthew Karlaftis, Fred Mannering, and Panagiotis Anastasopoulos. 2020. *Statistical and Econometric Methods for Transportation Data Analysis*. 3rd ed. Boca Raton: Chapman and Hall/CRC.

Washington, S, Karlaftis M., Mannering F, and Anastasopoulos P. 2020. *Statistical and Econometric Methods for Transportation Data Analysis*. 3rd ed. New York: CRC Press, Taylor and Francis Group.

Weng, Jinxian, Xiafan Gan, and Gang Du. 2019. “Random Coefficient Models for Work Zone Headway Distribution.”

Weng, Jinxian, Shan Yue, Ying Yang, Xuedong Yan, and Xiaobo Qu. 2015. “In-depth Analysis of Drivers’ Merging Behavior and Rear-end Crash Risks in Work Zone Merging Areas.” *Accident Analysis & Prevention* 77: 51–61.

Yang, Minming, Xuesong Wang, and Mohammed Quddus. 2019. “Examining Lane Change gap Acceptance, Duration and Impact Using Naturalistic Driving Data.” *Transportation Research Part C: Emerging Technologies* 104: 317–331.

Zhang, Yue, Yajie Zou, Jinjun Tang, and Jian Liang. 2021. “Long-term Prediction for High-Resolution Lane-Changing Data Using Temporal Convolution Network.” *Transportmetrica B: Transport Dynamics*, 1–15.

Zheng, Zuduo. 2014. “Recent Developments and Research Needs in Modeling Lane Changing.” *Transportation Research Part B: Methodological* 60: 16–32.

Zheng, Zuduo, Soyoung Ahn, and Christopher M Monsere. 2010. “Impact of Traffic Oscillations on Freeway Crash Occurrences.” *Accident Analysis & Prevention* 42 (2): 626–636.

Zhong, Fang, Zhengbing He, Meng Zeng, and Yixuan Liu. 2022. “Dynamic Lane Changing Trajectory Planning for CAV: A Multi-Agent Model with Path Preplanning.” *Transportmetrica B: Transport Dynamics* 10 (1): 266–292.