Is the Car Following Behaviour of Human Drivers Affected when Following Autonomous Vehicles?

Ruixuan Zhang[1], Sara Masoud[2], and Neda Masoud[3]

[1] Graduate Research Assistant, Civil and Environmental Engineering

University of Michigan Ann Arbor

Address: 2350 Hayward St, Ann Arbor, MI 48109

Email: ruixuanz@umich.edu

[2] Assistant Professor, Industrial and Systems Engineering

Wayne State University

Address: 4815 4th St, Detroit, MI 48201

Email: saramasoud@wayne.edu

[3] Assistant Professor, Civil and Environmental Engineering

University of Michigan Ann Arbor

Address: 2350 Hayward St, Ann Arbor, MI 48109

Email: nmasoud@umich.edu

Abstract

The past few years have been witness to an increase in autonomous vehicle (AV) development and testing. However, even with a fully developed and comprehensively tested AV technology, AVs are anticipated to share the roadway network with human drivers for the unforeseeable future. In such a mixed environment, we use naturalistic driving data from the Next Generation Simulation (NGSIM) and Lyft Level 5 (Lyft L5) prediction datasets to investigate whether the existence of AVs influences the car following behavior of human drivers. We use time headway time series as a proxy to capture the car following behaviour of human drivers. We then develop a nested fixed model to find possible changes in behaviour when human drivers are following different types of vehicles (i.e., human-driven vehicles or AVs). The factors included in this model are the platoon structure (a legacy vehicle following a legacy vehicle, and a legacy vehicle following an autonomous vehicle), road type (freeway and urban), time period (morning and afternoon), lane (right, middle, and left), and the source of the data (NGSIM and Lyft L5). Results indicate a statistically significant difference between the car following behaviour of drivers when they follow a human-driven vehicle compared to an AV. This change in the car following behaviour of drivers is manifested in the form of a reduction in the mean and variance of time headways when human drivers follow an AV. These findings can bridge the gap between anticipated and real-world impacts of AVs on traffic streams and roadway stability and capacity, providing meaningful insights on the influence of AVs on the driving behavior of humans in a naturalistic driving environment.

Keywords: Autonomous vehicle-human driver interactions, Car following behaviour

1 Introduction

The past few years have been witness to an increase in autonomous vehicle (AV) development and testing, with many mobility-oriented companies as well as original equipment manufacturers (OEMS) attempting to either open AV divisions or partner with/acquire start-ups that focus on software or hardware development for AVs. This move toward a future autonomous transportation system is fueled by many anticipated benefits of AVs, such as increased safety and smoother traffic flow, which in turn leads to higher levels of fuel economy, less congestion, and curbing the environmental footprint of the transportation sector Stern et al. (2018). It might, however, take several decades for a fully autonomous transportation system to be deployed. Many experts argue that even with a fully developed and comprehensively tested AV technology, there will still be individuals who either have a distrust in the technology or do not wish to cease driving for other personal reasons. Therefore, it is safe to assume that AVs would have to share the roadway network with human drivers for the unforeseeable future.

Since the advent of personal automobiles traffic engineers have been interested in studying the car following behaviour of human drivers, with Bruce Greenshields being credited with the first recorded set of experiments to scientifically measure this car following behaviour Greenshields et al. (1934). The advent of AVs has given rise to an interesting research question: will the car following behaviour of human drivers be affected when they knowingly follow an autonomous vehicle? Few attempts have been made in the literature to answer this question. Rahmati et al. (2019) set up two sets of experiments with a platoon of size three, where the third vehicle in the platoon was a human-driven vehicle. In the first set of experiments the second vehicle was a human-driven vehicle, and in the second set of experiments it was an AV. They recorded the trajectory of the third vehicle, and used data-driven and model-based approaches to discern any changes in the car following behaviour of the third vehicle in reaction to its proceeding vehicle. They concluded that when following an AV, a human driver's car following behaviour is significantly different than following a human-driven vehicle.

Conducting controlled field experiments allows for assessing the impact of a single factor at a time on the car following behaviour of human drivers, while keeping all other factors fixed. However, controlled field experiments have a number of downsides. First, a combinatorial number of experiments are required to capture the impact of multiple factors changing at once. This could easily render comprehensive controlled field experiments impractical, since a wide range of environmental factors as well as the presence of other agents (e.g., other AVs or legacy vehicles, pedestrians, bicycles, etc.) may play a role in the car following behaviour of drivers. As a result, the conclusions obtained from basic and contained field experiments, although insightful, may not be readily generalizable to a naturalistic driving environment. As such, in this paper we seek to investigate the car following behavior of human drivers who follow an AV in a naturalistic driving environment using a naturalistic and large dataset that allows for making statistically significant conclusions. To this end, we use the Lyft Level 5 (Lyft L5) Houston et al. (2020) data repository, in which a fleet of AVs travels on a fixed route in an urban environment, providing over 1,000 hours of AV trajectories, their surrounding agents, and the transportation network. The route encompasses a variety of transportation facility types, including intersections and corridors. This dataset is the first to enable analysis of the car following behaviour of a heterogeneous set of drivers who follow an AV in a naturalistic and dynamically changing driving environment.

Despite the benefits of using naturalistic driving data in analyzing the changes in the car following behaviour of human drivers when following an AV, it also poses a unique set of challenges. More specifically, the appearance of an AV is a key factor that can influence a human driver's car following behavior. For the presence of an AV to change the behaviour of human drivers, they should be able to discern that they are following an AV based on clear visual cues. Garnished by lidars and cameras, AVs generally have a distinctive look that human drivers are likely to discern. Additionally, a human driver's car following behaviour depends on their subjective opinion on how an AV operates and its risk-taking attitude Zhao et al. (2020). As such, to mitigate the risk of unwanted bias in data collection, data should be collected in an extended period of time from a diverse set of drivers.

The car following behaviour of a driver can be reflected using a number of parameters, e.g., velocity, acceleration, and time headway Wang et al. (2014). Here, we use time headway (THW)–defined as the time it takes for the following vehicle to reach its leading vehicle–to model car following behaviour. As such, we conduct change point analysis on THW of the following driver to identify the moment in time when the human driver has identified its leading AV.

The remainder of the paper is organized as follows. In section 2 we present the existing work and list the contributions of this paper. In section 3 we present our analytical approach in detail. In section 5 we lay out our analysis using Lyft L5 and NGSIM datasets and present our findings. We conclude the paper in section 6.

2 Literature review

In traffic modeling, the car-following behavior has been intensively studied to establish how a vehicle interacts with its leading vehicle. The main idea is to work with longitudinal dynamics of the vehicle pair, such as velocity, acceleration, time headway, and time-to-collision inverse, to uncover the behavior patterns of the following vehicle in various driving scenarios. There are two main components involved in the study of car-following behavior: modeling and analysis. These two components are discussed in the following.

2.1 Modeling

As the most commonly encountered driving maneuver in the real world, car following behavior has been extensively studied in investigating many specific driving situations. To properly describe the interaction between the leading and following vehicles, several measures are proposed. Time-to-collision (TTC) reflects human drivers' perception of their safety for potential collision and it is strongly related to longitudinal acceleration/deceleration (Jin et al., 2011). (Vogel, 2003) compares time headway and TTC with real-world traffic data and concludes that time headway and TTC are independent but suitable for different usages. They also argue that time headway directly reflects potential danger and thus prevents risky TTC, while TTC should be used for actual danger, i.e., on-road obstacle or collision. (Boer, 1999) also mentioned that time headway characterizes the safety margin in the situation where the preceding vehicle decelerates while TTC means the time left for drivers to intervene to avoid for a crash. Headway is not considered here as it can not include velocity-related information which is necessary to learn car following behavior. As we are interested in human drivers' reaction to on-road stimuli (the preceding AV) without evaluating an actual collision, time headway works better. Several car-following behavior models are formulated using ordinary differential equations (ODE) that take positions and velocities of vehicles as inputs. The intelligent driving model (IDM) Treiber et al. (2000) and optimal velocity model (OVM) Sugiyama (1999) are two extensively-applied ODE-based models capable of modeling nonlinear dynamics. Additionally, a linearized model can be further derived from ODEs via Taylor expansion. The full velocity difference model (FVDM) Jiang et al. (2001) was developed based on OVM and the generalized force model (GFM) Helbing and Tilch (1998) by taking both positive and negative velocity differences into account. It could obtain more precise predictions of vehicle motion in traffic jam density. Wiedemann 74 (W-74) model and Wiedemann 99 (W-99) model Durrani et al. (2016) are two car-following models developed by Rainer Wiedemann, where the drivers change their behaviors at discrete time steps only when certain thresholds (predefined for headway, speed, or relative speed) are reached. However, the values of parameters in W-99 are empirical, and no literature exists to indicate how ranges for these parameter should be established, which prompted many related works Durrani et al. (2016); Mathew and Radhakrishnan (2010); Gallelli et al. (2017) in calibrating the W-99 model. Newell's car-following model Newell (2002) applied a similar concept to W-99, assuming that a vehicle will maintain a minimum space and time gap between itself and its preceding vehicle. Some studies which pursue a more general way of modeling the car-following behavior are discussed in Ro et al. (2017) and Koutsopoulos and Farah (2012), where not only the car-following dynamics is considered, but also random human factors and different driving scenarios (such as following and emergency braking) were accounted for. Other car-following models such as adaptive cruise control (ACC) and cooperative adaptive cruise control (CACC) were designed for commercial vehicles, applying automated longitudinal control by adjusting acceleration with a linear function to maintain preset velocity and headway values.

All of the aforementioned car-following models are based on mathematical formulations with longitudinal dynamics, taking advantage of traditional control theory. On the other hand, predictive techniques enable a data-driven approach and can directly learn the car-following behavior using real-world data. Zhang et al. (2008) utilized time headway and time-to-collision inverse data and a back-propagation neural network to reproduce longitudinal accelerations. A long short-term memory (LSTM) neural network in Zhang et al. (2019) used the position information of surrounding vehicles to predict the car-following behavior with low longitudinal trajectory error. A deep deterministic policy gradient reinforcement learning car-following model was developed in Zhu et al. (2018), where a mapping from speed, relative speed, and headway to acceleration regime of the following vehicle were obtained to deliver a human-like car-following model. A Gaussian mixture model (GMM) was developed in Angkititrakul et al. (2011) to anticipate the future car-following behavior based on velocity and headway. Such learning-based methods require a large amount of training data, and the quality of data significantly influences model performance. Neural network-based designs also require careful tuning when learning the longitudinal dynamics of vehicles Da Lio et al. (2020).

From the literature, it can be noticed that multiple longitudinal dynamics impact the car-following behaviors of both the following vehicle and the proceeding vehicle, among which relative distance and velocity are the two most essential factors. To leverage this finding and reduce the complexity of the model, we select time headway as the main feature for modeling car-following behavior as it accounts for both relative distance and velocity (Chen et al. (2015) and Vogel (2002)).

2.2 Analysis

Car-following behavior is of interest to transportation researchers as it can provide insights into the best ways to approach flow throughput control, on-road safety, and energy consumption, etc. There are two directions followed in the current literature to analyze car-following behavior of drivers: one studies the stability (string stability and plant stability) of traffic flow, while the other quantifies the car-following behavior using statistical tools such as mean and variance. As this work focuses on patterns of interactions between human-driven vehicles and AVs, the analysis of string stability and plant stability is out of the scope this study.

Car-following behavior may be affected by many factors such as road condition, weather, and vehicle type. When dealing with data relevant to multiple factors, Analysis of Variance (ANOVA) is a powerful tool to investigate the influence level of each factor. In Liu et al. (2019), two one-way ANOVA tests were conducted, indicating that different speed limits have a significant influence on the time headway and headway, and the mean of time headway is closely centered around a fixed value. A factorial ANOVA analysis was conducted in Hjelkrem (2015) to determine the interactions between area type, number of lanes, and vehicle type influencing the car-following behavior. Road condition is suggested to be a critical factor in influencing both headway and time headway by Wang et al. (2015) and Houchin (2015). Significant influence from vehicle type (2-door car v.s. 4-door vehicles, sedans v.s. trucks, vehicles v.s. motorcycles) is also observed in Evans and Wasielewski (1983), Houchin (2015), and Amini et al. (2019).

The literature on the analysis of car-following behavior mainly focuses on human-driven vehicles, and AV-involved scenarios are rarely studied. Human-AV interactions at the microscopic level were first studied in Rahmati et al. (2019), where a field experiment was conducted though setting up two two-vehicle platoon structures of human-following-human and human-following-AV. Rahmati et al. (2019) showed that a shorter headway is selected when human drivers follow an AV. Other field experiments conducted by Zhao et al. (2020) suggest that a driver's subjective attitude toward to AV technology dominates the actual AV's driving behavior in the speed-headway relationship. Observations from these two field experiments indicate that the limited data collected from field experiments degrades the robustness of the intersection effect(s). Recently, Li et al. (2021) leveraged the Lyft L5 dataset as the data source for operational safety analysis in human-AV interactions in car-following scenarios. In this study we utilize the Lyft L5 and NGSIM datasets to provide a comprehensive and robust evaluation of the car following behaviour of humans, accounting for multiple factors that may affect the car-following behaviour of human drivers. This naturalistic study serves as a necessary complement to the existing field experiments.

2.3 Contributions

The objective of this paper is to provide insights on the potential influence of AVs on the car-following behavior of human drivers. The contributions of this paper are two-fold: (*i*) we apply statistical analysis on time headway data from Lyft L5, using NGSIM datasets (US101, I-80, Lankershim Blvd) as the control group, to find the influence of leading AVs on the car-following behaviour of following drivers; (*ii*) This naturalistic study provides evidence that human drivers are regulated as a result of introducing AVs, as evidenced by the statistically significant reduction in the mean value and variance of their time headways.

3 Methods

The objective of this study is to investigate whether, and the extent to which, the existence of AVs in the traffic stream influences the car following behaviour of human drivers. To answer this question, we propose a comprehensive framework demonstrated in Figure 1. Data used in this study is obtained from two public datasets: Lyft L5 Houston et al. (2020) and NGSIM NGS (2021). We use time headway time series in our analysis as a proxy to quantify the car-following behaviour of vehicles. Time headway between two vehicles is defined as the travel time from the centroid of the following vehicle to the centroid of the preceding/leading vehicle based on the following vehicle's speed. In the rest of this paper, we denote a legacy vehicle following an autonomous vehicle as LFA, and a legacy vehicle following a legacy vehicle as LFL. We refer to LFA and LFL as platoon structures.

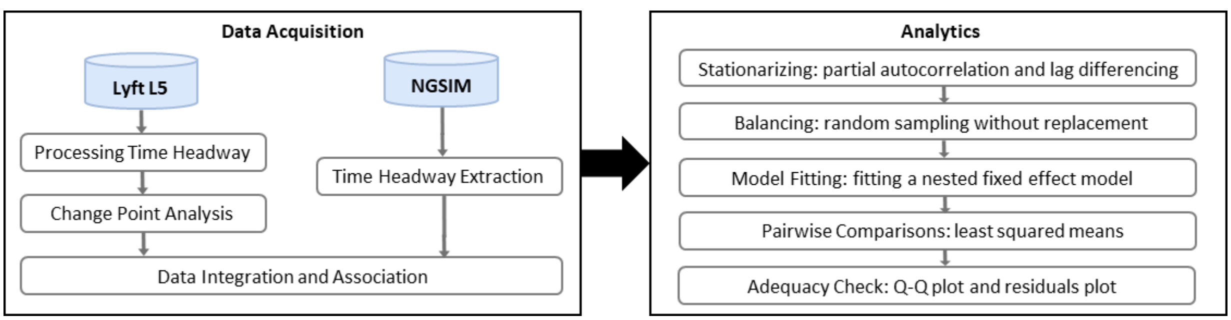


Figure 1: The proposed framework to study the car-following behavior of driversing in LFL and LFA platoon structures.

As displayed in Figure 1, the proposed framework consists of two main phases, namely, data acquisition and data analysis. These phases are described in the following sections.

3.1 Phase I: Data Acquisition

The first phase starts by extracting time headways of LFL and LFA platoon structures. More precisely, we extract LFA time headways from the Lyft L5 dataset, and LFL time headways from both Lyft L5 and NGSIM datasets. Once the time headways are extracted, We use Bayesian change point analysis to filter out the portions of time headway data in the LFA platoon structure where the legacy vehicle is not aware of following an AV.

3.1.1 Change Point Analysis

Our objective in this study is to make a determination on whether the presence of an AV affects the car following behaviour of its following vehicle in the LFA platoon structure. Consequently, we first need to identify scenarios in the Lyft L5 dataset where a legacy vehicle is following an AV, and more importantly, is *aware* that it is following an AV. To identify such scenarios, we first identify scenes from the Lyft L5 dataset where a legacy vehicle is immediately following an AV. Next, for each scene we use change point analysis to mark any changes in the time headway sequence of the legacy vehicle, and the velocity sequence of its leading AV Ruggieri and Antonellis (2016). The adopted Change point analysis is an online detection approach that provides uncertainty bounds on the number and location of change points across observations Ruggieri and Antonellis (2016). This method strives to make fast inferences on the occurrence of new change points based on each new observation Ruggieri and Antonellis (2016).

Let us denote by c_L^h the time instance when a change point is detected in the time headway time series of the legacy vehicle, and by c_A^v the time instance when a change point identified in the velocity time series of the AV. Let us denote by t^r_{min} and t^r_{max} the minimum and maximum reaction time of the legacy vehicle, i.e., the time period lapsed from the moment the AV changes its acceleration and the moment the acceleration of the legacy vehicle changes in response. When $t^r_{\text{min}} \leq c_L^h - c_A^v \leq t^r_{\text{max}}$, the change in the time headway of the legacy vehicle can be attributed to its car-following behaviour. However, when c_L^h is not proceeded with a c_A^v within the time interval $[t^r_{\text{min}}\ ,\ t^r_{\text{max}}]$, i.e.,, the change point analysis detects a change in time headway of the legacy vehicle that cannot be attributed to its car-following behaviour, we postulate that this change can be attributed to the legacy vehicle having identified its proceeding vehicle as an AV, and only consider the trajectory of the legacy vehicle after this change point. In other instances where no such change point is detected, we assume that the legacy vehicle is aware of its leading AV due to the distinctive appearance of AVs in the Lyft L5 study.

In the final step of phase I, the collected and filtered time headways from both Lyft L5 and NGSIM datasets are integrated and associated. In this step, each time headway is labeled based on platoon structure, road type, time period, data source, and lane as shown in Figure 2.

3.2 Phase II: Analysis

Phase II focuses on analysis. In the first step, two samples of equal sizes are taken from LFA and LFL datasets. Next, partial autocorrelation analysis is employed to detect autocorrelation lags. Using these identified lags, differencing is applied to stationarize the randomly selected time series.

Next, we define the factors of interest, which alongside time headway will be used for fitting the ANOVA model. For human drivers, there is an empirical preferable time headway interval towards the preceding vehicle (Fuller (1981); Das and Maurya (2017)). When time headway is shorter than the lower bound, drivers are more likely to slow down, while when the time headway is longer than the upper bound, drivers may either keep the current speed or accelerate to catch up with the preceding vehicle. The basic idea is that when the time headway is inside the interval, human drivers will feel comfortable and will not overreact unless there is an external disturbance. This preferable time headway is also influenced by many factors (e.g., road configuration, lane, etc.). Generally, there is no universal standard, and this interval can be determined from the observed data itself. We use the distribution of time headway in the LFA dataset to define the preferable time headway.

Once the factors of interest are identified and before fitting the nested model, we first create balanced datasets. To obtain balanced datasets we sample time headways without replacement from LFL and LFA datasets so that the same number of data points will be available in each branch of the nested design. Next, the ANOVA model is fitted using balanced datasets. Finally, we confirm the adequacy of the fitted model, and conduct follow-up pair-wise comparisons to isolate the effects that are significantly different, as displayed in Figure 1. The major steps of the analysis are detailed in the following.

3.3 Analysis of Variance

Analysis of Variance (ANOVA) is one of the most well-known statistical tools for evaluating the existence of significant differences between factor levels on a continuous measurement (Tabachnick and Fidell, 2013). A factorial ANOVA can be implemented to examine the impacts of independent categorical factors on a continuous target variable. Factorial ANOVA is an appropriate approach to study whether there exists a statistically significant difference in the time headway patterns of LFA and LFL platoon structures based on different factors and their levels. One of the main requirements of ANOVA is the independence of observations. The underlying sequential and time dependant nature of time series data is a direct violation of this requirement. To address this issue, we apply a two-step data processing procedure. First, we randomly (without replacement) down-sample the time series to remove any potential dependencies. Next, we render the randomly selected time series approximately stationary through differencing to remove auto-correlation.

3.3.1 Stationarity and Partial Auto-Correlation

In time series, auto-correlation is the correlation between two observations at different time stamps, where these observations correlate with themselves repetitively at certain lags. Auto-correlation and partial auto-correlation plots can be used to study the auto-correlation of time series. Although auto-correlation plots can measure and visualize the correlation between observations for a predefined set of lags, they fail to account for the propagation of correlation among successive lags. Partial auto-correlation analysis addresses this problem by isolating the auto-correlation lag. In this work, we use partial auto-correlation plots to identify auto-correlation lags, and apply differencing at the identified lags to stationarize the time headway time series. We discard data points that cannot be stationarized by first level differencing.

3.3.2 Nested Fixed Effect Model

The design of the fitted factorial ANOVA is highly dependent on the structure of the collected data. Fig. 2 displays the factors of interest. A total of five factors are considered in this study. The first factor, platoon structure, models whether the reported time headway profiles belong to an LFL or an LFA pair. The second factor, road type, represents whether the data is collected from an urban road network (i.e., Palo Alto, CA and Lankershim Blvd, CA) or a freeway (i.e., US 101, CA and I-80, CA). The third factor, time period, models whether the data in collected during the morning (i.e., 7:50am - 9:00am) or afternoon (i.e., 4:00pm - 5:30pm) peak period.

The fourth factor studies whether the source of the collected data has any significant impact on human driving behavior. Data source is defined as a factor to account for the impact of different data collection techniques and locations in NGSIM and Lyft L5 datasets. The final factor, lane, represents the lane at which the data has been collected. This factor is considered because the lane in which a vehicle travels could impact its car following behaviour. As the number of lanes is

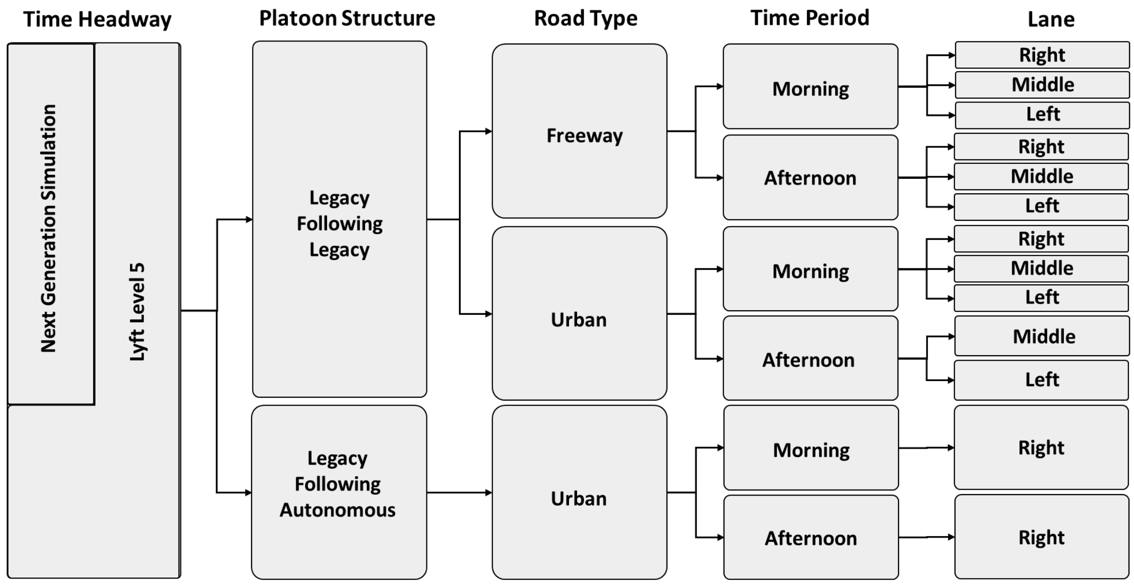


Figure 2: The Structure of the proposed nested model.

different across data collection sites, we used one-way ANOVA to group lanes that failed to show a statistically significant difference in their car following behaviour based on time headway analysis. As a result, the lane levels simplified to the left (i.e., speeding) lane, the middle lanes, and the right (merging) lane. Note that the high occupancy vehicle lanes were filtered out in this study when present.

The factorial ANOVA relies on the underlying relationships between these different factors. Note that AVs are only present in the Lyft L5 dataset and the Lyft L5 data is limited to an urban environment. Furthermore, AV trajectories only appear on the right lane. As such, the values of factors data source, lane, and road type are restricted to the values of the factor platoon structure, leading to the choice of a nested factorial ANOVA as shown in Equation (1).

$$\begin{aligned} Y_{l(ijknm)} &= \mu + \alpha_i + \beta_j + (\alpha \times \beta)_{ij} + \gamma_{k(j)} \\ &\qquad + \lambda_{m(j)} + \theta_{n(j)} + \epsilon_{l(ijknm)}, \\ &\qquad \text{for } i, j, k, m \in \{1, 2\} \text{ and } n \in \{1, 2, 3\} \end{aligned} \tag{1}$$

where μ represents the overall mean, and $\alpha_i, \beta_j, \gamma_{k(j)}, \lambda_{m(j)}$, and $\theta_{n(j)}$ capture the effects of time period, platoon structure, data source, road type, and lane, respectively. The parenthetical subscriptions illustrate the nesting structure of the model. The $(\alpha \times \beta)_{ij}$ models the interaction effects between factors time period and platoon structure. Here, $\epsilon_{l(ijknm)}$ represents the error term, which is assumed to follow $N(0, \sigma^2)$. In addition to the normality and constant assumptions regarding the error term, the fitted model should also satisfy the following constraints:

$$\sum_i \alpha_i = 0 \tag{2a}$$

$$\sum_j \beta_j = 0 \tag{2b}$$

$$\sum_i (\alpha \times \beta)_{ij} = 0, \qquad \forall j \in \{1, 2\} \tag{2c}$$

$$\sum_j (\alpha \times \beta)_{ij} = 0, \qquad \forall i \in \{1, 2\} \tag{2d}$$

$$\sum_k \gamma_{k(j)} = 0, \qquad \forall j \in \{1, 2\} \tag{2e}$$

$$\sum_m \lambda_{m(j)} = 0, \qquad \forall j \in \{1, 2\} \tag{2f}$$

$$\sum_n \theta_{n(j)} = 0, \qquad \forall j \in \{1, 2\} \tag{2g}$$

As the nested factorial model in Equation (1) is not identifiable, the additional sets of constraints in Equation (2) help narrow down the solution space to a unique set of fitted parameters. Using a single ANOVA model, we define several hypotheses tests to assess the significance of each factor, with the null hypothesis in each case indicating that the mean time headways are similar for different values of a given factor, and the alternative hypothesis indicating otherwise. Nested factors (i.e., data source, lane, and road type) are added to absorb some of the unexplained variability. As a result, specific hypothesis tests associated with nested factors are of lesser importance.

Although a rejection of the null hypothesis in the ANOVA analysis signals the existence of a significant effect (i.e., factor), it fails to identify the factor level that is significantly different, specifically in the presence of interaction effects. As a result, ANOVA analyses are usually followed by pairwise comparisons. While studying the effects of multiple factor levels, comparisons between the individual means of either factor may be made using any pairwise comparison technique. We use Least Square Means to investigate the significance of the factors and apply Tukey's HSD method to adjust the significance level Abdi and Williams (2010).

Multiple assumptions are made prior to fitting the nested fixed effect model. As a result, the adequacy of the model relies on whether these assumptions hold true. These assumptions include

1) the normality of the residuals, i.e., $\epsilon_{l(ijknm)} \sim N(0, \sigma^2)$, and 2) the homogeneity of the residuals. Many mathematical tests are developed for checking the normality and homogeneity of the residuals (e.g., the Shapiro-Wilk test). One problem with such tests is that as the sample size increases, the test results are more likely to fail for even minor departures from normality or homoscedasticity. Therefore, in this paper we rely on visualization approaches instead.

4 Data

The raw data within both repositories are collected using different sensors such as digital video cameras, radars and lidars.

4.1 Lyft L5 Dataset

The Lyft L5 Prediction data repository was released by the Lyft Level 5 team in June 2020 Houston et al. (2020). This data repository contains raw camera/lidar/radar data collected from a fleet of 23 AVs operating along a fixed high-demand route in Palo Alto, CA, from October 2019 to March 2020. An internal perception stack has already been applied to report information such as the vehicle position based on a global coordinate system, velocity, and a unique ID for each agent. We extract the time headway series of each legacy vehicle in an LFA platoon structure for the purpose of this study.

4.2 NGSIM Dataset

The Next Generation Simulation (NGSIM) is a well known dataset published by the U.S. Department of Transportation Intelligent Transportation Systems Joint Program Office (JPO) NGS (2021). This dataset includes detailed vehicle trajectory data collected in four sites: southbound US 101 and Lankershim Boulevard in Los Angeles, CA, eastbound I-80 in Emeryville, CA, and Peachtree Street in Atlanta, Georgia. The data is collected in different time periods from April 20, 2005 to November 9, 2006. The dataset contains vehicle ID, global coordinates of the vehicle, vehicle type, velocity, acceleration, space headway, and time headway, among other attributes. We extract the time headway series of each vehicle in each regular (non-carpool) lane at each site for the purpose of this study.

4.3 Data Processing Pipeline

To fully leverage the abundant data in the Lyft L5 and NGSIM datasets for ANOVA, a modular data processing pipeline is developed with three blocks: time headway calculation, change point analysis, and down-sampling and filtering. A detailed explanation of the processing pipeline is given for the Lyft L5 dataset.

- Time headway calculation: Realizing that the driving behavior in different lanes on the same road may be different, the lane-specific time headway data is of our interest. To stay consistent with the NGSIM dataset, all the raw data in the L5 dataset is taken from the multi-lane roads. By utilizing the provided semantic map with 8.500 discrete lane segments, a customized semantic map is constructed by connecting any lanes that physically belong to the same continuous lane (multiple lane segments in the original semantic map may correspond to the same lane in the real world), referred as the augmented map. In the multi-lane roads, three lanes are identified (right lane, middle lane, and left lane). Given the position information of vehicles, the augmented map can immediately match the vehicles to the corresponding lanes. The time headway in the car-following mode is calculated as the travel time from the centroid of the following vehicle to the centroid of the preceding/leading vehicle based on the following vehicle's speed.

- Change point analysis: In investigating an AV's effect on the following behaviour of human drivers, we need to construct a dataset in which the following human driver is aware that the leading vehicle is an AV. To this end, we conduct a change point analysis as described in section 3.1.1.

- Down-sampling and filtering: The sampling frequency in both datasets is 10 Hz, and a high correlation among data points is present under such a high-frequency sampling regime. To ensure independence of observations, autocorrelation and partial autocorrelation are evaluated, and down-sampling of the time headway sequence is implemented. According to our evaluation results, 1 Hz is selected to be the updated sampling frequency. Furthermore, a filtering step is introduced to ensure that the time headway sequence satisfies the minimum length of containing at least 10 data points or 10-seconds of observation.

For the NGSIM dataset, as the lane information is readily available, only the down-sampling and filtering module will be used.

5 Results and Discussion

In this section, we present the results of our proposed framework. In accordance with the flow of the framework, we first stationarize the time headway time series through differencing and partial auto-correlation analysis. Then, we balanced our dataset. Next, we test our hypotheses using nested factorial ANOVA, followed by pairwise comparisons.

5.1 Down-sampling and Auto-correlation Analysis

Since the sample frequency in Lyft L5 and NGSIM datasets is high (10 Hz), data points may correlate with each other at such high frequency and thus introduce unnecessary bias into the results. A common approach to reduce autocorrelation is to down-sample the data at a slower frequency. We test Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF) at down-sampling frequencies of 2Hz and 1Hz, in comparison with the original data. Decreasing sample frequency can significantly reduce both ACF and PACF at higher lags. Down-sampling at 1 Hz can reduce the magnitude of the auto-correlation lags. Differencing at lag one further stationalizes the time series. As the majority of the time series are not significantly auto-correlated after lag 1 differencing, the non-stationary ones are dropped at this step.

Some interesting takeaways may be discussed before presenting the ANOVA results. In a freeway driving environment, e.g., US 101 and I-80, after down-sampling at 1 Hz, there is still a significant autocorrelation at lag 1 and neutrally-distributed partial autocorrelation (PAC) after lag 2. In an urban driving environment, Lankershim Blvd and Lyft L5, a similar pattern can be observed; however, at lag 1, a relative smaller ratio of data is correlated. An interpretation for this difference is that in freeways, human drivers encounter fewer external disturbances and therefore their behavior is more consistent and predictable. A neutral-distributed outbound PAC after lag 2 indicates that the behaviors tend to be random in 2 seconds into the future. If we view a human driver as a controller, s/he will control the time headway to the leading vehicle roughly at some period, which can be determined by the lag where outbound PAC values are approximately neutral-distributed.

5.2 Factorial Analysis

The processed dataset contains a total of 537,060 data points, out of which 5,774 (i.e., 1%) of data points represent the LFA structure while the remaining 531,285 (i.e., 99%) belong to the LFL platoon structure. In order to maximize the power of the factorial analysis, the dataset should be balanced. In addition, balancing helps protect the analysis against small departures from the assumptions. Although the balancing effort reduces the total size of the dataset (i.e., 25 data points per each leaf in Figure 2) through random sampling, it improves the the distribution of the data within different factor levels, including platoon structure: 85% for LFL and 15% LFA; Road type: 46% for freeway and 54% urban; Time period: 53% for morning and 45% afternoon; Lane: 31% for left, 31% for middle and 38% right.

The nested factorial ANOVA introduced in Equation 1 is fitted and its results are displayed in Table 1. The fitted model allows us to study whether there are statistically significant associations between the time headway and the factors introduced in Figure 1. Table 1 reports findings on the main effects (i.e., time period and platoon structure factors), nested effects (i.e., data source, road type, and lane factors), as well the interaction effects between the time period and platoon structure factors.

Table 1: Results of the nested fixed model

Factor	DoF	SSE	MSE	F Statistics	P-Value	α
Time Period	1	1.46	1.46	1.55	0.21	
Platoon Structure	1	49.86	49.86	52.81	2.88e-12	0.001
Platoon Structure × Time	1	1.09	1.09	1.16	0.28	
Platoon Structure: Data Source	1	0.03	0.03	0.04	0.85	
Platoon Structure: Road Type	1	1.92	1.92	2.03	0.15	
Platoon Structure: Lane	2	0.01	0.006	0.006	0.99	
Residuals	317	299.28	0.94			

The first three rows in Table 1 correspond to hypotheses on time period, platoon structure, and the interaction effect between time period and platoon structure factors. The next three rows display the impact of data source, road type, and lane as nested factors of platoon structure, respectively. The last row provides some information regarding the residuals. For each one of the hypotheses of interest, Table 1 reports the degree of freedom (DoF) of the test, sum of squared errors (SSE), mean square errors (MSE), as well as the F-statistics, its corresponding p-value, and the significance level at which a conclusion is made. The reported p-values can assess the null hypotheses and determine whether the association between the time headway and the factors of interest are statistically significant. Table 1 reports that only the platooning structure is of significance at $\alpha = 0.001$. The results also highlights the fact that the collected time headway data are not impacted by the differences in data collection techniques and locations in NGSIM and Lyft L5 datasets at a statistically significant level. To further study the results reported in Table 1, multiple follow up pairwise comparisons are conducted to understand which levels of the platoon structure factor are significantly different given the nested structure. Table 2 illustrates the results of the pairwise comparisons.

Table 2: Pairwise comparisons using least square means

	Estimate	SE	T Ratio	P-Value	α
Time Period (Morning vs Afternoon) : Platoon Structure (LFL vs LFA)					
Morning LFL - Afternoon LFL	-0.132	0.132	-0.996	0.7519	
Morning LFL - Morning LFA	0.944	0.215	4.39	0.0001	0.001
Morning LFL - Afternoon LFA	1.055	0.215	4.91	<.0001	0.001
Afternoon LFL - Morning LFA	1.075	0.218	4.93	<.0001	0.001
Afternoon LFL - Afternoon LFA	1.187	0.218	5.44	<.0001	0.001
Morning LFA - Afternoon LFA	0.112	0.275	0.40	0.9774	
Lane (Left vs Middle vs Right) : Platoon Structure (LFL vs LFA)					
Left LFL - Middle LFL	0.013	0.138	0.098	0.9997	
Left LFL - Right LFL	0.016	0.158	-0.103	0.9996	
Left LFL - Right LFA	1.073	0.0171	6.279	<.0001	0.001
Middle LFL - Right LFL	-0.003	0.160	-0.022	1.000	
Middle LFL - Right LFA	1.061	0.171	6.279	<.0001	0.001
Right LFL - Right LFA	1.057	0.191	6.20	<.0001	0.001

Although the platoon structure is the only significant factor as reported in Table 1, the interaction effect between time period and platoon structure and the nesting factors may have obscured the comparisons between the means of different levels of the platoon structure. As a result, the least squared method is applied to the means of one of the factors, with the remaining factor set at a particular level. In addition, as pairwise comparisons lead to inflation of the significance level, the p-values within Table 2 are adjusted based on the Tukey method for comparing a family of multiple estimators.

Table 2 reports the estimated difference between means (i.e., estimate), the standard error of that estimate (i.e., SE), the T ratio, and its corresponding p-value along with the reported level of significance α. The top half of Table 2 studies the pairwise comparisons between time period and platoon structure. Here, results are averaged over the levels of lane (i.e., left, middle, and right), road type (i.e., freeway and urban), and data source (i.e., NGSIM and Lyft L5). As shown in Table 2, other than the Morning LFL - Afternoon LFL and Morning LFA - Afternoon LFA, the remaining levels between time period and platoon structure are significant.

The bottom half of Table 2 studies the interaction between the nested factor lane and the main factor platoon structure. Here, results are averaged over the levels of time period (i.e., morning and afternoon), road type (i.e., freeway and urban), and data source (i.e., NGSIM and Lyft L5). This table demonstrates that while the LFL behavior does not significantly differ within the middle, left, and right lanes, it does significantly differ within the left and right, middle and right, as well as right and right lanes when compared to LFA. It also shows that although LFL and LFA display statistically different behaviors in different lanes, they do exhibit statistically significantly different behaviors even within the same right lane. Although the proposed nested factorial model recognizes platooning structure leads to a statistically significant different car-following behaviour, and the follow-up pair-wise comparisons further confirm this, none of these approaches can identify whether the THW of LFA is less than or greater of LFL's THW. Figure XX demonstrates that LFL has higher mean and variance THW values when compared to LFA.

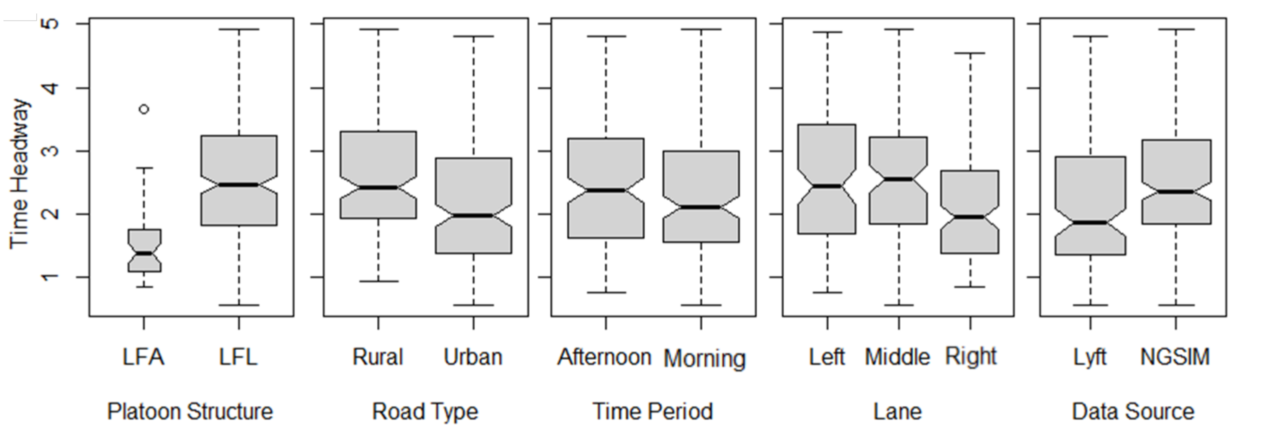
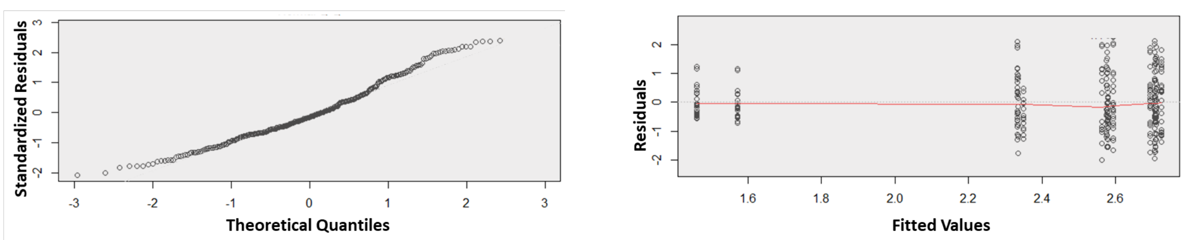


Figure 3: The distribution of time headway over factor levels.

As displayed in Figure 2, LFA has lower median (1.38), mean (0.41), and variance (o.31) values in comparison to LFL which has scores of median (2.48), mean (0.85), and variance (1.05). The reduction in the mean time headway manifests in less bumper-to-head distance, enabling more vehicles to operate on the road and increasing the road capacity. The reduction in the variance of time headway leads to a more stable traffic flow.

The final step is the verification of the fitted model's adequacy through Q-Q and residuals plots as shown in Figure 4. To check the adequacy of the model, Q-Q plots of residuals and residuals versus fitted values are shown in Figure 4. Q-Q plots are commonly used to confirm the normality of the residuals, i.e., $\epsilon_{l(ijknm)} \sim N(0, \sigma^2)$. As a Q-Q plot is a scatter plot created by plotting the actual quantiles of the residuals of the fitted model against the theoretical normally distributed ones, a diagonal line is a confirmation that both sets of quantiles came from the same distribution. In the Q-Q plot in Figure 4, the residuals roughly lie around the 45-degree line, suggesting that the they are approximately normally distributed. The homogeneity of the residuals can be validated using the residuals plot. If the variance of the error term is homogeneous, not only should the residuals plot show no pattern, but also the spread of residuals should be equal per group across corresponding fitted values. The residuals plot in Figure 4 show that the variances are approximately homogeneous since the residuals are distributed approximately equally above and below zero.



(a) Q-Q plot **(b)** Residuals v.s. fitted values

Figure 4: Adequacy check of the fitted nested fixed effect model

6 Conclusions

In this study we proposed a nested factorial model to study the potential effects of autonomous vehicles on human drivers' car-following behavior using naturalistic driving data (i.e., NGSIM and Lyft L5 prediction datasets). The objective of this study was to bridge the gap between anticipated and real-world impacts of AVs on traffic streams and roadway safety and capacity. The proposed nested model studied the impact of different factors such as platoon structure (i.e., whether a human driver follows a legacy vehicle or an AV), time period, traveling lane, and road type on the time headway between two vehicles, which is considered as a proxy for the car following behaviour of the following vehicle. The results indicate that the platoon structure affects the car following behavior of human drivers in a statistically significant manner, allowing us to conclude that in a real-world setting, a human driver's car following behaviour when following a legacy vehicle is different from following an autonomous vehicle. Furthermore, our analysis illustrates that the difference in car following behaviour is significantly present regardless of the traveling lane or the time period.

Data Availability Statement

Some of models, or code that support the findings of this study are available from the corresponding author upon reasonable request; All data used during the study are available in repositories online in accordance with funders data retention policies.

Acknowledgments

This work has been supported by Midwest US-DOT Center for Connected and Automated Transportation (award number: 69A3551747105) and National Science Foundation (award number: 1837245).

References

(2021). U.s. department of transportation federal highway administration. (2016) next generation simulation (ngsim) vehicle trajectories and supporting data. [dataset]. provided by its datahub through data.transportation.gov. http://doi.org/10.21949/1504477 Accessed May 1.

Abdi, H. and Williams, L. J. (2010). Tukey's honestly significant difference (hsd) test. *Encyclopedia of research design*, 3(1):1–5.

Amini, E., Tabibi, M., Khansari, E. R., and Abhari, M. (2019). A vehicle type-based approach to model car following behaviors in simulation programs (case study: Car-motorcycle following behavior). *IATSS research*, 43(1):14–20.

Angkititrakul, P., Miyajima, C., and Takeda, K. (2011). Modeling and adaptation of stochastic driver-behavior model with application to car following. In *2011 IEEE Intelligent Vehicles Symposium (IV)*, pages 814–819. IEEE.

Boer, E. R. (1999). Car following from the driver's perspective. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2(4):201–206.

Chen, X. M., Li, L., and Shi, Q. (2015). A markov model based on headway/spacing distributions. In *Stochastic Evolutions of Dynamic Traffic Flow*, pages 49–79. Springer.

Da Lio, M., Bortoluzzi, D., and Rosati Papini, G. P. (2020). Modelling longitudinal vehicle dynamics with neural networks. *Vehicle System Dynamics*, 58(11):1675–1693.

Das, S. and Maurya, A. K. (2017). Time headway analysis for four-lane and two-lane roads. *Transportation in developing economies*, 3(1):9.

Durrani, U., Lee, C., and Maoh, H. (2016). Calibrating the wiedemann's vehicle-following model using mixed vehicle-pair interactions. *Transportation research part C: emerging technologies*, 67:227–242.

Evans, L. and Wasielewski, P. (1983). Risky driving related to driver and vehicle characteristics. *Accident Analysis & Prevention*, 15(2):121–136.

Fuller, R. G. (1981). Determinants of time headway adopted by truck drivers. *Ergonomics*, 24(6):463–474.

Gallelli, V., Iuele, T., Vaiana, R., and Vitale, A. (2017). Investigating the transferability of calibrated microsimulation parameters for operational performance analysis in roundabouts. *Journal of Advanced Transportation*, 2017.

Greenshields, B. D., Thompson, J., Dickinson, H., and Swinton, R. (1934). The photographic method of studying traffic behavior. In *Highway Research Board Proceedings*, volume 13.

Helbing, D. and Tilch, B. (1998). Generalized force model of traffic dynamics. *Physical review E*, 58(1):133.

Hjelkrem, O. A. (2015). Determining influential factors on the threshold of car-following behavior. Technical report.

Houchin, A. J. (2015). An investigation of freeway standstill distance, headway, and time gap data in heterogeneous traffic in iowa.

Houston, J., Zuidhof, G., Bergamini, L., Ye, Y., Chen, L., Jain, A., Omari, S., Iglovikov, V., and Ondruska, P. (2020). One thousand and one hours: Self-driving motion prediction dataset. *arXiv preprint arXiv:2006.14480*.

Jiang, R., Wu, Q., and Zhu, Z. (2001). Full velocity difference model for a car-following theory. *Physical Review E*, 64(1):017101.

Jin, S., Huang, Z.-y., Tao, P.-f., and Wang, D.-h. (2011). Car-following theory of steady-state traffic flow using time-to-collision. *Journal of Zhejiang University-SCIENCE A*, 12(8):645–654.

Koutsopoulos, H. N. and Farah, H. (2012). Latent class model for car following behavior. *Transportation research part B: methodological*, 46(5):563–578.

Li, T., Han, X., Ma, J., Ramos, M., and Lee, C. (2021). Operational safety of automated and human driving in mixed traffic environments: A perspective of car-following behavior. *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, page 1748006X211050696.

Liu, T. et al. (2019). Comparison of car-following behavior in terms of safety indicators between china and sweden. *IEEE Transactions on Intelligent Transportation Systems*, 21(9):3696–3705.

Mathew, T. V. and Radhakrishnan, P. (2010). Calibration of microsimulation models for nonlane-based heterogeneous traffic at signalized intersections. *Journal of Urban Planning and Development*, 136(1):59–66.

Newell, G. F. (2002). A simplified car-following theory: a lower order model. *Transportation Research Part B: Methodological*, 36(3):195–205.

Rahmati, Y., Khajeh Hosseini, M., Talebpour, A., Swain, B., and Nelson, C. (2019). Influence of autonomous vehicles on car-following behavior of human drivers. *Transportation research record*, 2673(12):367–379.

Ro, J. W., Roop, P. S., Malik, A., and Ranjitkar, P. (2017). A formal approach for modeling and simulation of human car-following behavior. *IEEE transactions on intelligent transportation systems*, 19(2):639–648.

Ruggieri, E. and Antonellis, M. (2016). An exact approach to bayesian sequential change point detection. *Computational Statistics & Data Analysis*, 97:71–86.

Stern, R. E., Cui, S., Delle Monache, M. L., Bhadani, R., Bunting, M., Churchill, M., Hamilton, N., Pohlmann, H., Wu, F., Piccoli, B., et al. (2018). Dissipation of stop-and-go waves via control of autonomous vehicles: Field experiments. *Transportation Research Part C: Emerging Technologies*, 89:205–221.

Sugiyama, Y. (1999). Optimal velocity model for traffic flow. *Computer Physics Communications*, 121:399–401.

Tabachnick, B. G. and Fidell, L. S. (2013). Using multivariate statistics upper saddle river.

Treiber, M., Hennecke, A., and Helbing, D. (2000). Congested traffic states in empirical observations and microscopic simulations. *Physical review E*, 62(2):1805.

Vogel, K. (2002). What characterizes a “free vehicle” in an urban area? *Transportation research part F: traffic psychology and behaviour*, 5(1):15–29.

Vogel, K. (2003). A comparison of headway and time to collision as safety indicators. *Accident analysis & prevention*, 35(3):427–433.

Wang, J., Xiong, C., Lu, M., and Li, K. (2015). Longitudinal driving behaviour on different roadway categories: an instrumented-vehicle experiment, data collection and case study in china. *IET Intelligent Transport Systems*, 9(5):555–563.

Wang, W., Xi, J., and Chen, H. (2014). Modeling and recognizing driver behavior based on driving data: A survey. *Mathematical Problems in Engineering*, 2014.

Zhang, L., Wang, J., Li, K., Yamamura, T., Kuge, N., and Nakagawa, T. (2008). Driver car-following behavior modeling using neural network based on real traffic experimental data. In *15th World Congress on Intelligent Transport Systems and ITS America's 2008 Annual MeetingITS AmericaERTICOITS JapanTransCore*.

Zhang, X., Sun, J., Qi, X., and Sun, J. (2019). Simultaneous modeling of car-following and lane-changing behaviors using deep learning. *Transportation research part C: emerging technologies*, 104:287–304.

Zhao, X., Wang, Z., Xu, Z., Wang, Y., Li, X., and Qu, X. (2020). Field experiments on longitudinal characteristics of human driver behavior following an autonomous vehicle. *Transportation research part C: emerging technologies*, 114:205–224.

Zhu, M., Wang, X., and Wang, Y. (2018). Human-like autonomous car-following model with deep reinforcement learning. *Transportation research part C: emerging technologies*, 97:348–368.