Contents lists available at ScienceDirect

Transportation Research Part F

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



Influence of following vehicle's tailway and classification on subject driver's response to the circular yellow indication



Hameed A. Mohammed^a, David S. Hurwitz^{a,*}, David A. Noyce^b, Xuesong Wang^c

^a School of Civil and Construction Engineering, Oregon State University, United States

^b Department of Civil and Environmental Engineering, University of Wisconsin Madison, United States

^c College of Transportation Engineering, Tongji University, China

ARTICLE INFO

Article history: Received 4 February 2020 Received in revised form 18 August 2020 Accepted 29 November 2020

Keywords: Dilemma zone Driver behavior Tailway Circular yellow indication

ABSTRACT

Traffic crashes at signalized intersections are frequently linked to driver behavior at the onset of the circular yellow (CY) indication. To better understand behavioral factors that influence a driver's decision to stop or go at an intersection, this study analyzed the behavior of the driver of a subject vehicle at the onset of the CY indication. Driver performance data from 53 participants were collected in the Oregon State University Driving Simulator, simulating scenarios of driving through high-speed intersections under various conditions. Data included interactions where the driver stopped at the stop line (n = 644) or proceeded through the intersection (n = 628) in response to a CY indication. Data were analyzed as panel data while considering 12 indicator variables related to the driver's stop/go decision. These indicator variables included time to stop line (TTSL), tailway time, following vehicle type, vehicle speed at the onset of the CY indication, and demographics (age, gender, driving experience, level of education, personal vehicle type, number of times driving per week, number of miles driving last year, participation in previous simulation studies. A randomparameter binary logit model was used to determine contributing factors for driver decision making at the onset of CY indication while accounting for unobserved heterogeneity. Four indicator variables were significantly related to the driver's stop/go decision, but three factors varied across observations. Findings showed that a driver's stop/go decision in response to a CY indication was associated with the time to the stop line (TTSL), tailway time to the following vehicle, subject vehicle speed at the onset of the CY indication, and driver's age (20-36 years), but was not significantly associated with classification of the following vehicle. Also, the findings indicated that a shorter tailway increased a subject driver's red-light running frequency. These findings provide insights into variables that affect driver decisions in a vehicle-following situation at the onset of the CY indication. This information can help make better decisions in smart traffic control systems such as to extend/decrease the green interval slightly to avoid decisions that are more difficult. © 2020 Elsevier Ltd. All rights reserved.

1. Introduction

Placement of traffic signals at intersections helps ensure the safety and efficiency of conflicting traffic movements. The Federal Highway Administration estimates that the U.S. is home to > 3 million intersections, including at least 300,000

* Corresponding author.

https://doi.org/10.1016/j.trf.2020.11.016 1369-8478/© 2020 Elsevier Ltd. All rights reserved.

E-mail address: david.hurwitz@oregonstate.edu (D.S. Hurwitz).

signalized intersections (FHWA, 2014). Approximately 40% of all crashes in the U.S. occur at intersections (Awadallah, 2009), particularly at signalized intersections where dilemma-zone conflicts related to the onset of the circular yellow (CY) may occur (Elhenawy et al., 2015). For example, in 2016, there were 3,145 fatalities resulting from crashes at signalized intersections in the U.S. (NHTSA, 2017).

The circular yellow (CY) indication warns drivers that the circular green indication has ended and the circular red indication will be presented next (FHWA, 2009). Specifically, the CY indication aims to provide a smooth transition during termination of the right-of-way for a particular movement (Brehmer et al., 2003; McGee et al., 2012). However, language in state laws on the correct driver response to a CY indication sometimes differs from language in driver training manuals, creating unnecessary driver confusion (Mohammed et al., 2018). At the onset of a CY indication, driver behavior can be considered a binary choice between coming to a safe stop before the stop line or proceeding through the intersection before termination of the yellow change interval (Elhenawy et al., 2015). Although these two options may seem straightforward, there are challenges associated with faulty driver decision making at the onset of the CY indication at isolated high-speed signalized intersections (Rakha et al., 2011). The inability of drivers to make correct decisions in the dilemma zone contributes to crashes at signalized intersections (Gates et al., 2007; Gates et al., 2010; Abbas et al., 2016; Gates et al., 2016).

This research focused on the Type II dilemma zone, also referred to as the indecision or option zone, which occurs as a result of complex driver behavior at the onset of the CY indication. Generally, the Type II dilemma zone describes the area where the driver has difficulty making the correct decision to stop or go (Hurwitz et al., 2011). This dilemma zone comprises the area upstream of the signalized intersection where 10% to 90% of drivers will stop in response to the CY indication (Gates et al., 2007). An incorrect decision may lead to various crash types. If the driver goes through the intersection when the correct decision is to stop, the driver may run the red indication, leading to a right-angle or head-on crash with another vehicle. If the driver stops when the correct decision is to clear the intersection, a rear-end crash could occur (Gates et al., 2016; Hurwitz et al., 2012; Rakha et al., 2007).

At the onset of a CY indication, driver behavior and decision making are affected by: (1) driver characteristics (e.g., perception-reaction time, age, gender), (2) intersection characteristics (e.g., type of intersection control, time to the intersection at onset of the CY indication, signal coordination, approach grade, pavement and environmental conditions), (3) vehicle characteristics (e.g., classification, approach speed, safe acceleration/deceleration rates, (4) signal control settings (e.g., yellow change interval duration), and (5) traffic-flow characteristics (e.g., headway and travel time) (Abbas et al., 2006; Rakha et al., 2007; Bonneson et al., 2002). For example, driver behavior is generally affected by the vehicle classification, with heavy vehicles being less likely than passenger vehicles to stop when caught in a dilemma zone. Heavy vehicles cannot stop as quickly as passenger vehicles, and their operational cost is higher when delayed. Moreover, drivers of heavy vehicles try to use gentler deceleration rates during emergency situations to prevent cargo from shifting (Gates et al., 2010).

Vehicle-following behavior is another important element of transportation safety. Drivers can avoid crashes by maintaining safe temporal and spatial separation with vehicles in their same lane (Qu et al., 2014). There are two distinct aspects of car-following behavior: (1) determining an acceptable tailway when following a subject vehicle, and (2) controlling acceleration in response to shifts in trajectory of the subject vehicle. Rear-end crashes occur when the distance between the following and subject vehicles reduces to zero as a result of the deceleration or acceleration of the subject or following vehicle, respectively (Sato and Akamatsu, 2012). Differences typically exist in the following behaviors of drivers of passenger cars vs. heavy vehicles due to differences in physical and operational characteristics of the vehicles, which can significantly affect traffic stream characteristics. Acceleration, relative speed, and free space between the subject and following vehicles are all variables that significantly influence the following behavior of heavy-vehicle drivers (Aghabayk Eagely et al., 2012).

The objective of this research was to investigate the influence of following vehicle tailway and following vehicle classification on the subject driver's response to the CY indication at a high-speed signalized intersection. A driving simulator study was conducted to address this objective. A random-parameter binary logit model was developed to examine the probability that a driver will stop at the stop line or proceed through the intersection at the onset of the CY indication. The model was developed as a function of the main decision-making factors, including tailway time, classification of the following vehicle (see Fig. 1), the time to the stop line (TTSL) of the subject vehicle, subject vehicle speed at the onset of the CY indication, and demographic variables, such as driver's age and gender.

2. Background and problem statement

Substantial previous research has addressed driver behavior in response to the CY indication. In 1978, Zegeer and Deen defined the boundaries of the indecision zone in terms of distance from the stop line, starting where 90% and ending where 10% of drivers stopped. Subsequently, Chang et al. (1985) attempted to define the boundaries in terms of travel time to the stop line, finding that 85% of drivers stopped if they were \geq 3 s from the stop line. Almost all drivers continued through the intersection if they were \leq 2 s from the stop line. Webster and Elison (1965) and Bonneson et al. (1994) similarly defined the indecision zone in relation to the stop line. Synthesizing results from several of the previously mentioned studies, Bonneson et al. (1994) developed the popular definition of the indecision zone as the area between 5.5 and 2.5 s from the stop line, as measured from the onset of CY indication.

Rakha et al. (2007) performed a field study of 60 participants to characterize driver behavior at the onset of the CY indication. Probability of stopping varied from 100% at a TTSL of 5.5 s to 9% at a TTSL of 1.6 s. Some previous driving simulator



Fig. 1. Vehicle-following behavior at onset of the CY indication.

studies showed that TTSL (Caird et al., 2007) and driver age (Senserrick et al., 2007) had significant effects on the driver's stop/go decision at the CY indication.

Elmitiny et al. (2010) indicated that operating speed, vehicle distance from the stop line at the onset of the CY indication, and vehicle position (subject /following) in the traffic flow were important factors in the stop/go decision and red-light running (RLR) violations. Gates et al. (2007) indicated that drivers were more likely to result in a go- through event with shorter travel time to the intersection at the onset of the yellow indication, with a longer yellow-interval, a shorter cycle length, and if the subject vehicle was a heavy vehicle. Hurwitz et al. (2012) developed a binary logistic regression model for the probability that a driver would stop or go at the onset of a CY indication based on empirical vehicle position data. Model input was generated from a fuzzy subset that required fewer data than similar models. Moore and Hurwitz (2013) used driving simulator data to build a fuzzy logic model based on the TTSL, which is a function of vehicle position and speed.

Li et al. (2016) conducted a study to predict the stop/go decisions of drivers and RLR violations during a CY indication. Data were gathered by a vehicle data collection system and analyzed by a sequential logit model. TTSL at the onset of the CY indication was an important factor for stop/go decisions and RLR violations, and the speed of the approaching vehicle was a contributing variable for stop/go decisions. Vehicle acceleration after onset of the CY indication was positively correlated with RLR violations. Xiong at el. (2016) examined driver behaviors during CY indication maneuvers in a driving simulator. Driver stop/go decisions in response to a CY indication were associated with TTSL, age, and distractions.

As a research tool, a driving simulator offers a safe and controlled environment and a cost-effective way of investigating transportation-related issues (Burnett, 2008). However, there are some limitations of using driving simulators. First, the generalizability of findings from driving simulator studies to the real world has been questioned. For example, some studies have shown that people drive more cautiously in driving simulators than they do on the road (Dutta et al., 2004; Young et al., 2008). Second, although driving simulators vary in sophistication, concerns have been raised for certain use cases (Santos et al., 2005; Burnett, 2008; Young et al., 2008). Third, it can be difficult to create trials of on-road and simulator driving that have comparable conditions in term of tasks, participants, variables, measures, environment, etc. (Sharples et al., 2016). Despite these limitations, driving simulators have robustly demonstrated relative behavioral validity for many measures of driving performance (Noyce and Elango, 2004; Young et al., 2008; Parkes, 2012; Swake et al., 2013). In addition, a driving simulator was effective in the evaluation of driver comprehension of signal displays (Knodler et al., 2005).

In summary, although extensive research has been conducted into the driver's response to the CY indication, there is a lack of empirical data related to vehicle-following behavior while the CY indication is displayed. There is an important gap in research related to effect of the following vehicle tailway and following vehicle classification on the subject driver's response to the CY indication at a high-speed signalized intersection in a simulator environment. This research fills these knowledge gaps. Additionally, previous research suggests that several factors, including TTSL, approach speed of the vehicle, and driver age, affect the decision of drivers to stop or proceed in response to the CY indication. These factors were also examined in this study.

3. Method

For the experiment, each participant traveled along a virtual highway with several signalized intersections in a driving simulator environment. Driver behavior was observed during the onset of the CY indication to investigate the effects of TTSL, tailway, and classification of the following vehicle on the driver's decision to stop at the stop sign or proceed through the intersection.

3.1. Driving simulator

The OSU Driving Simulator consists of a fully functional full-size 2009 Ford Fusion cab mounted on an electric pitch motion system that accurately reproduces acceleration and braking events. The cab is surrounded by screens projecting the simulated environment. Three liquid crystal on silicon (LCOS) projectors with a resolution of 1400×1050 are used to project a front view of 180° by 40° . These front screens measure 11 ft \times 7.5 ft. A DLP projector is used to display a rear image for the driver's center mirror. The two side mirrors have embedded LCD displays. Sound is provided by 500 W surround sound speakers.

The vehicle cab instruments are fully functional and include a steering control loading system to accurately represent steering torques based on vehicle speed and steering angle. The production instrument panel has been replaced with a configurable LCD instrument panel. The data update rate for graphics is 60 Hz. Fig. 2 shows views of the simulated environment created for this experiment from inside (right) and outside (left) the vehicle.

The virtual environment was created with Internet Scene Assembler (ISA) and SimCreator (Realtime Technologies, Inc.), and Blender. The simulated environment was developed in ISA by using Java Script-based sensors to change the signal indication and display dynamic objects, such as a following vehicle responding to the subject vehicle while approaching the intersection.

3.2. Scenario layout and intersection control

Roadway cross-sections consisted of one lane in each direction of travel separated by solid double yellow lines. Lane widths were 12 ft in each direction and included a white edge line to delineate the right edge of the travel way. The experiment required participants to drive a virtual roadway at a posted speed limit of 45 mph. Regulatory speed limit signs displayed the posted speed limit at several locations along the roadway. Curve warning signs were added to provide drivers warnings about upcoming horizontal curves. This experiment was modeled with a suburban land use pattern, which included retail stores, houses, gas stations, schools, and service buildings, as well as green zones with various vegetation. Duration of the yellow change interval was 4.5 s, consistent with the duration suggested by the ITE kinematic equation (FHWA, 2009).

A within-group counterbalanced partially randomized factorial experimental design was used in this study. Participants were presented with combinations of 3 independent variables (Table 1): TTSL (2.5, 3.5, 4.5, or 5.5 s), tailway (0.5, 1, or 2 s), and classification of the following vehicle (passenger car or heavy vehicle). Based on FHWA vehicle classifications, heavy vehicle was class 5 (single unit 2-axle trucks) (see Table 1). Examples of following vehicle scenarios as shown in Fig. 3.

The $4 \times 3 \times 2$ factorial design resulted in 24 scenarios being presented to participants across 6 grids. A total of 53 participants were exposed to various conditions to measure their response to the onset of the CY indication. In each grid, 4 signalized intersections, each separated by roughly 2,000 ft of roadway, were modeled. Fig. 4 shows an example grid layout as presented to drivers.

In this experiment, 24 scenarios were manipulated within-subject. This within-subject design provides advantages of greater statistical power and reduced error variance associated with individual differences (Cobb, 1998). To control for prac-



Fig. 2. OSU Driving Simulator.

Table 1

Transportation Research Part F 77 (2021) 73-86

Experimental factors and levels. Variable name Acronym Level Description 0 Time to Stop Line TTSL 2.5 s 1 3.5 s 2 4.5 s 3 5.5 s Tailway TW 0 0.5 s 1 1 s 2 2 s Following Vehicle Classification FVC 0 PC 1



Fig. 3. Examples of following vehicle scenarios.

tice and carryover effects, the factorial design was fully counterbalanced and partially randomized. Additionally, the duration of test drives was designed to be relatively brief.

3.3. Subject recruitment and sample size

Participants between the ages of 18 and 75 years were recruited for the experiment from the area surrounding Corvallis, Oregon. Participants were required to possess a valid driver's license and to be able to be calibrated with the eye tracker. In



Fig. 4. Grid layout.

total, 54 participants (30 men, 24 women) participated in the study. Only 1 person (1 woman) experienced simulator-based discomfort. Data from that participant were excluded from the analysis. Participant ages ranged from 18 to 70 years (M_{age} = 31.2 years, SD_{age} = 13.7 years). Efforts were made to recruit participants of all ages and varying backgrounds. Recruitment efforts included flyers posted and distributed around the OSU campus and the city of Corvallis. The research design and all study documentation were reviewed and approved by the OSU Institutional Review Board (IRB) (Study # 8080).

3.4. Procedure

Upon a participant's arrival to the laboratory, the approved informed consent document was presented and explained to the participant. Each participant was given \$10 compensation in cash for participating in an experimental trial after signing the informed consent document. After providing informed consent, participants were asked to complete a prescreening and demographics survey, which asked questions about their prior experience with motion discomfort, simulator discomfort, and driver simulators, age, gender, driving experience, and highest level of education. After the prescreening survey, each participant completed a test drive for a 3- to 5-min calibration drive to acclimate the participant to the operational characteristics of the driving simulator and to determine whether they were susceptible to simulator discomfort.

After the participant's eyes were calibrated to the driving simulator screens, they drove 6 grids. The virtual driving course itself was designed to take the participant approximately 30 min to complete. After the experimental drives, participants responded to several questions in an online Qualtrics survey. The entire experiment lasted approximately 1 h.

3.5. Data analysis

The statistical software for data analysis was NLOGIT Version 5.0. There were 1,272 observations from the 53-participant dataset. A value of 0.05 was used as the criterion for statistical significance. A statistical model of drivers' stopping probability at a signalized intersection was developed, which considered both the response and explanatory variables.

3.5.1. Response variable

An objective of this study was to examine whether drivers stopped or proceeded through the intersection on the CY indication under different conditions. Therefore, a binary dependent variable represented the drivers' response (1 = Stop or $0 = G_0$) to the onset of the CY signal indication.

3.5.2. Explanatory variables

Predictors included TTSL, tailway, classification of the following vehicle, and vehicle speed at the onset of the CY indication. TTSL represents the number of seconds it takes for a vehicle travelling at a certain speed to reach the stop line, starting from the onset of the CY indication. To accommodate potential sources of variability from individual subjects, demographics (age, gender, driving experience, level of education, personal vehicle type, number of times driving per week, number of miles driving last year, participation in previous simulation studies) were included in the analysis.

4. Statistical method

A driver's decision to stop or proceed through an intersection when the traffic signal turns from a circular green to a CY indication is a dichotomous variable. Discrete outcome models (e.g., binary logit and probit models) are well suited for such data (Savolainen, 2016) and can be used to examine driver decisions associated with factors such as driver demographic features and driving simulator events (e.g., TTSL, vehicle speed, tailway, etc.). Although several studies (Gates et al., 2007; Papaioannou, 2007; Rakha et al., 2008; Sharma et al., 2011; Liu et al., 2012) have adopted logit or probit models to study driver decision making at signalized intersections, authors typically assumed that estimated parameters were fixed across observations for participants. Fixing parameters that actually vary across observations could lead to biased and inefficient parameter estimates (Washington et al., 2011; Greene, 2012; Agbelie, 2016). Using models that allow some or all of the estimated parameters to vary across participants could provide more robust results, thereby improving understanding of the parameters that influence driver decision making at signalized intersections (Agbelie, 2016).

In the present study, data were classified as a panel dataset because multiple observations were collected from each participant. The final model accounted for potential correlations across observations. A random-parameter logit model approach was previously shown to be useful for accounting for unobserved heterogeneity across observations (Savolainen, 2016). When a panel dataset is modeled with this approach, parameter estimates are allowed to vary between participants, but each estimate is restricted to a fixed value for observations from the same participant (Lavrenz et al., 2014). Given the possibility of heterogeneity in observed and unobserved variables for driver behavior data, the random-parameter model can be an appropriate methodology for studying the driver's go/stop decision at the onset of the CY indication.

4.1. Modeling framework

Binary logistic regression was applied due to the binary nature of the selected response variable. The response variable had two possible outcomes: 1, if the driver stopped at the stop line during the CY indication; and 0, if the driver went through the intersection. As such, the following binary logit formulation was used to determine the probability *Pij* of driver *i* stopping during event *j* as a function of covariates (Washington et al., 2011):

$$Pij = \frac{EXP(\overset{K}{X}ij\beta)}{1 + EXP(\overset{K}{X}ij\beta)}(1)$$

where β is a vector of estimable parameters, and *Xij* is a vector of explanatory variables (e.g., characteristics of driver, vehicle, and simulation), used to determine the outcome probability of *Pij* being equal to 1 and associated with driver *i* and simulator event*j*.

There are two important methodological concerns with a standard logit model. The first concern relates to the structure of data for the estimation of a standard logit model. The within-group experimental design resulted in the same 53 individuals being observed multiple (24) times, once during each scenario. As such, it is reasonable to expect there to be correlations in the decisions that are made by the same participants across simulator events. If these correlations are not considered, then the resulting parameter estimates will be inaccurate due to biased standard errors (Savolainen, 2016). To mitigate this concern, the 53 participants were treated as a panel, with parameter estimates assumed to be equal for each participant and allowed to vary across participants. The second concern related to the potential influence of unobserved heterogeneity. Specifically, each participant may show unique characteristics that make them more (or less) prone to stop at the stop line or proceed through the intersection at the onset of the CY indication. These concerns were addressed by using a flexible model, which is explained by relevant recent research in this area (Lavrenz et al., 2014; Savolainen, 2016).

One alternative model to account for heterogeneity across individuals is the random-parameter logit model. The model captures heterogeneity resulting from unobserved factors that are common to each study participant by allowing the constant term to vary across participants. This heterogeneity among participants is assumed to follow a parametric distribution (e.g., normal, lognormal, triangular, etc.) and can reflect those unobserved factors (e.g., tendency of risk, driving style, etc.) that may affect driver decision making. Another, yet-unsolved concern relates to the heterogeneity of covariate effects, as the model implicitly assumes that covariates have a compatible influence across participants. Heterogeneity is also caused by unobserved features of the participants or scenarios, which are not captured by the model. These concerns can be accommodated by allowing all parameters to vary across participants, while also holding parameters at the same value for each participant. Not accounting for this heterogeneity can lead to inaccurate or biased model estimates and corresponding inferences.

Therefore, to account for the heterogeneity, constants and covariates were allowed to vary across participants by applying a random parameters technique. Equation (2) is now written as (Washington et al., 2011):

$$Pij = \int_{X} \frac{EXP(\hat{X} \, ij\beta i)}{1 + EXP(\hat{X} \, ij\beta)} f(\beta/\phi) d\beta(2)$$

where $(\beta i/\phi)$ is the density function of β , with distributional parameter ϕ . All other terms are as previously defined. Density function $f(\beta i/\phi)$ is defined as having a distribution, which depends on the analysis (e.g., normal, uniform, etc.) and which parameters are permitted to vary across observations. This approach permits β to account for observation-specific variations of the effect of X on *Pij* (Washington et al., 2011). A simulation-based maximum likelihood approach with 200 Halton draws was used to estimate the random-parameter logit model as recommended by previous research (Bhat, 2003).

Normal, uniform, and triangular distributions were tested, but only the normal distribution was found to have statistically significant standard deviations. To evaluate the effects of the variables, inferences from partial effects were applied. Partial effects measure the effect on the response variable when there is a one-unit increase in an explanatory variable while holding all other variables are constant (i.e., equal to their means) (Anderson et al., 2018).

5. Results and discussion

5.1. Preliminary investigation: Driver's decision to stop/go

A driver's decision to stop before the stop line or proceed through the intersection is the foundation for developing models to describe the dilemma zone. The final dataset from this experiment contained a comprehensive set of variables for 1,272 vehicles approaching intersections during CY indications. Each vehicle had two choices: either to stop (n = 644) or go (n = 628), including cases of RLR (n = 46). Vehicle speed undoubtedly influences a driver's decision to stop or go; therefore, driver response was presented in relation to TTSL. As shown in Fig. 5, nearly all drivers (97%) went through an intersection when they were 2.5 s from the stop line at the onset of the CY indication. This finding is consistent with the findings of Chang et al. (1985), Gates et al. (2007), and Moore and Hurwitz (2013), who likewise reported that nearly all vehicles proceeded through the intersection when they were \leq 2.5 s from stop line at the CY onset. When TTSL was 5.5 s, most drivers decided to stop, and RLR violations started to increase.

Next, this study considered how the driver's decision to stop or go varied depending on the position of the vehicle relative to the stop line (Fig. 6). All vehicles proceeded through the intersection when they were \leq 100 ft from the stop line at the onset of the CY indication. By contrast, when drivers were 340–400 ft from the intersection at the onset of the CY indication, only 7% of drivers went through the intersection, 8% were RLR, and 85% stopped at the stop line. No vehicles except RLR vehicles (17%) proceeded through the intersection where they were at 400–460 ft at the onset of the CY.



Fig. 5. Probability of stopping based on TTSL.



Fig. 6. Probability of stopping based on distance from stop line.

5.2. Model results

To understand and evaluate factors affecting driver decision, 12 indicator variables were generated from the factorial design and demographic characteristics. All possible combinations of the continuous and categorical factors that influence the driver's decision making at the onset of the CY indication were examined to construct the model. The final combination of model factors was based on p-values at the 95% confidence interval. A stepwise procedure was used to test and determine statistically significant factors. Four parameters had statistically significant effects on the driver's decision to stop or go at the onset of the CY indication. Descriptive statistics for the significant parameters are shown in Table 2. Vehicle speeds at the onset of the CY indication varied from 19.20 to 64.70 mph. The mean speed was 46.5 mph, and the speed limit was 45 mph.

An estimated parameter is considered random across participants when the SD of the parameter density is statistically significant (Agbelie, 2014). If the estimated SD is not statistically significant (not statistically different from zero), then the estimated parameter can be considered fixed across participants. Three parameters were statistically significant and varied significantly across observations (random parameters) (Table 3). These parameters included TTSL, speed at onset of CY indication and driver age (20–36) year. The estimated constant also varied across participants.

An important aspect in interpreting parameter estimates in a random-parameter logit model relates to the concept of driver stop/go decision making. The dependent variable of driver decision can take a positive or a negative value. In this model, for the possibility of stopping at the stop line, a positive (or negative) parameter estimate should be interpreted as an increased (or decreased) probability that the driver will stop.

Partial effects of indicator variables were examined to understand their quantitative effects (Table 3). For continuous variables, such as TTSL, partial effects represent the percent increase in the probability that the driver will stop in response to the CY indication associated with a one unit increase in the covariate. Interpretation of the partial effects indicated a relationship between TTSL and stopping, in that an increase in TTSL increased the probability of stopping. Among the variables, TTSL had the largest effect on the probability of stopping, with the partial effect suggesting an increase in the probably of stopping of 0.46 for each 1 s increase in TTSL. Tailway increased the probability of stopping by 0.07 for each 1 s increase in tailway, and driver age of 20–36 years increased the probability of stopping by 0.25 for each 1 year increase in driver age. Conversely,

Table 2

Descriptive statistics.

Variable	Mean	SD	Min	Max	Cases
TTSL (s)	3.99	1.11	2.50	5.50	1,145
Tailway (s)	1.17	0.62	0.50	2.00	1,145
Speed at onset of CY indication (mi/h)	46.53	4.77	19.18	64.65	1,145
Driver age (20–36) (year)	26.55	4.68	20.00	36.00	38

Table 3

Random-parameter binary logit model estimates for response variable (driver decision making either go or stop).

Variable	Coef.	SE	t-Stat	P-value	Partial Effects
Constant	-3.80	0.68	-5.60	0.000	-
(Standard deviation of parameter)	0.49	0.11	4.64	0.000	-
TTSL	1.89	0.10	19.34	0.000	0.46
(Standard deviation of parameter)	0.47	0.03	13.74	0.000	-
Tailway	0.29	0.11	2.59	0.009	0.07
(Standard deviation of parameter)	0.02	0.08	0.21	0.833	-
Speed at onset of CY indication	-0.10	0.01	-7.55	0.000	-0.03
(Standard deviation of parameter)	0.02	0.00	6.40	0.000	-
Driver age (1 if 20-36 years, else 0)	1.06	0.18	6.01	0.000	0.25
(Standard deviation of parameter)	0.40	0.12	3.29	0.001	-
No. observations	1145				
Log likelihood at zero	-491.29				
Log likelihood at convergence	-384.08				
McFadden Pseudo R ²	0.22				

faster vehicle speed at the onset of the CY indication reduced the probability of stopping by 0.03 for each 1-mile increase in vehicle speed. Three estimated parameters were found to be random based on the statistical significance of SD as shown in Table 3. The model had a McFadden pseudo R^2 value of 0.22, which is considered a very good fit (Louviere et al., 2000) and sufficiently strong to predict driver behavior.

One of the main findings of this study was that the probability of stopping increased when TTSL increased. This finding is consistent with previous literature (Bonneson and Son, 2003; Caird et al., 2007; Elmitiny et al., 2010; Gates et al., 2012; Moore and Hurwitz, 2013; Savolainen, 2016; Li et al., 2016; Xiong et al., 2016; Pathivada and Perumal, 2017). Vehicles are more likely to stop in response to a CY indication when they have greater time (a product of distance and speed) from the stop line. Another key finding was that the probability of stopping decreased when vehicle speed was higher at the onset of the CY indication. Similarly, previous studies (Bonneson and Son, 2003; Papaioannou, 2007; Elmitiny et al., 2010; Gates et al., 2012; Li et al., 2016; Pathivada and Perumal, 2017) found that higher vehicle speed at the onset of the CY indication reduced the probability of stopping. The higher the vehicle speed is at the onset of the CY indication, the less time there is for the driver to react and brake, and the more likely it is that the driver will go through the intersection.

Tailway was hypothesized to play a role in the subject driver's decision to stop or go. In this study, the number of stop decisions by the driver decreased when the subject vehicle was closely followed by another vehicle during the CY indication. For example, when the tailway between vehicles was 0.5 s, the likelihood of drivers deciding to stop was reduced for all TTSLs (2.5, 3.5, 4.5, or 5.5 s) (see Fig. 7). In particular, many RLR violations were observed when the following vehicle was close to the subject vehicle at the onset of CY indication. The total number of RLR violations was 46 (3.6% of total observations), with 50% of RLR violations occurring when the tailway was 0.5 s. This fact may relate to the short time gap between



Fig. 7. Relationship between driver decision to stop and tailway.

the following and subject vehicles. A shorter tailway may influence a driver's decision to run the red light, due to the pressure experienced by the subject vehicle. Of the RLR violations, 26% occurred when the tailway was 1.0 s, which provides additional evidence about the relationship between the driver's decision to run the red light and the tailway. Interestingly, the classification of the following vehicle did not significantly contribute to the driver's stop/go decision. This may be related to the visual attention of drivers. Results from a previous study showed that while the CY indication was displayed, total fixation durations (TFDs) were highest on the traffic signal (626 s), lower for the rear view mirror (50 s), and lowest for the side view mirrors (3 s) (Aswad Mohammed, 2020). In other words, drivers spent significant time looking at the traffic signal and comparatively much less time inspecting the vehicles behind them during the CY indication. When the CY indication was displayed, a driver has limited time to make the correct decision because the yellow change interval was 4.5 s in this study and driver allocated more visual attention to the process of determining the status of the traffic signal than for evaluating the following vehicle's speed and type. This may indicate greater perceived risk on the part of the driver for misinterpreting the right-of-way direction as provided by the traffic signal, when compared to the risk posed by the following vehicle. Furthermore, drivers' visual search is controlled by higher order processes such as driving style, habit, experience, skill, and familiarity with the road and the vehicle. Therefore, most drivers were unaware of the classification of the following vehicle behind them at the onset of the CY indication, which could explain this result.

The probability of stopping was highest among drivers aged 20–36 years. It was not clear whether this effect was due to differences in physiology, driving style, driving behavior, or familiarity and comfort in a simulated driving environment. This finding is consistent with a previous driving simulator study (Savolainen, 2016), which found that drivers aged 18–45 years were more likely to stop.

5.3. Model validation

Model validation involved a two-step process. (1) The model was cross-validated with 90% of the data set. (2) Stopping probabilities were compared with other models based on distance from the stop line.

5.3.1. Model cross-validation

To validate the model produced by this research, 90% of the dataset was randomly selected to develop the model, which was used to predict driver decisions (go or stop) for the remaining 10% of the data. Results of the comparison between the actual observed behavior for 10% of the data and the predictive power of this model are presented in Table 4. Overall prediction accuracy for the developed model was 85%.

5.3.2. Probability of stopping

Driver decision-making data were compared to empirical datasets from prior research, including experiments by Moore and Hurwitz (2013), Gates et al. (2007), and Van der Horst and Wilmink (1986). Fig. 8 presents the probability of stopping from the present study to these previous experiments, one of which was conducted in a driving simulator and two in the field.

To test whether the probability distribution of the present study was similar to that of any of the previous studies, twoindependent-samples Kolmogorov-Smirnov tests were applied between this study and Moore and Hurwitz (2013), Gates et al. (2007), and Van der Horst and Wilmink (1986). Results of the comparison revealed no statistically significant differences (failed to reject the null hypothesis, p > .05) among distributions at the 95% confidence level. The stopping curve for Moore and Hurwitz (2013) is shifted to the left, meaning that, for each TTSL, a subject vehicle is more likely to stop.

It could be due to the lower operating speed during data collection.

6. Conclusions

Driver behavioral data were obtained from a study performed in the OSU Driving Simulator Laboratory. This dataset included data from 53 participants. Data were extracted from a series of events during which participants traversed virtual intersections as a traffic signal changed from the circular green to CY indication. A total of 1,272 observations were collected, wherein 51% of drivers stopped and 49% proceeded through the intersection in response to the onset of the CY indication. All drivers continued through the intersection at the onset of the CY indication if their distance to the stop line was \leq 100 ft. If the distance was \geq 400 ft, drivers were more likely to stop. The resulting dataset provides potential insights into how driver

Table 4

Prediction accuracy of the model.

		Predicted		%Correct
		Stop	Go	
Observed	Stop	60	9	87%
	Go	10	48	83%
			Predictive Power	85%



Fig. 8. Probability of stopping versus TTSL at onset of CY indication.

behavior was influenced by various factors. Analysis revealed four parameters that directly affect a driver's decision to stop or proceed: TTSL, tailway, vehicle speed at onset of CY indication, and driver age (20–36 years).

This study provides important insights into how unobserved heterogeneity can be considered in simulator data. In general, a logistic regression (i.e., logit) model is the most widely used approach in many studies in this area. However, using this model in a study of panel data can introduce methodological concerns, due to the use of multiple driving events conducted by the same 53 participants, unobserved heterogeneity, and behavioral correlations among participants. To address these concerns and produce more accurate estimates, a random-parameter logit model was applied. Heterogeneity was found within three variables: TTSL, vehicle speed at onset of the CY indication, and driver age (20–36 years). The estimated constant was found to be random. Probability of stopping decreased as vehicle speed increased at the onset of the CY indication. Moreover, the probability of stopping when the car was followed increased as the tailway and TTSL increased. Driver age (20–36 years) had a statistically significant effect on stopping probability, whereas the classification of the following vehicle did not significantly affect a driver's decision to go or stop in response to the CY indication.

When the developed model was validated by using 10% of extracted data, the predictive accuracy for the model was 85%. Additionally, data obtained from this study were statistically comparable to data from previous simulator, and in-field experiments in the literature. Taken together, our findings confirm that driving simulation is a valid mechanism for describing driver behavior under the given conditions.

This research has the potential to contribute to safety improvements at signalized intersections by considering the effect of vehicle-following behavior during the CY indication at high-speed signalized intersection. This work provides an effective framework to explore and understand the behavior of the subject vehicle when closely followed at the onset of CY indication. Drivers in this scenario have a higher risk of being involved in right-angle crashes if they incorrectly choose to proceed through the intersection during circular red indication. The results provide useful information for traffic engineers, revealing that a shorter tailway increases the subject driver's RLR frequency. This finding could be used in the development of appropriate countermeasures to avoid RLR violations. For example, traffic safety efforts could include the promotion of longer tailway between drivers, particularly on the approach to signalized intersections. Advanced in-vehicle technology could play a vital role in warning when a following vehicle is following the subject vehicle too closely. Additionally, adaptive signals and smart signal technologies can also benefit from this knowledge.

One of limitation of this study was the unbalanced sample with respect to age. Approximately 70% of participants were between 20 and 36 years old. Many of the recruited participants were students, which increased the number of younger participants.

For future studies, researchers could use on-road driving to investigate the influence of following tailway and classification of the following vehicle on driver behavior at the onset of the CY indication, to complement simulated driving. This approach would help to extend the results to the real world. In addition, other driving conditions that influence driving behavior at the onset of CY indication should be studied. These include, but are not limited to, environmental conditions (e.g., day or night driving), weather conditions (e.g., rain or fog), low speed condition (25 mph), yellow change intervals (e.g., 3 s) and different classifications of subject and following vehicles.

CRediT authorship contribution statement

Hameed A. Mohammed: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft, Writing - review & editing. **David S. Hurwitz:** Conceptualization, Formal analysis, Investigation, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing - original draft, Writing - review & editing. **David A. Noyce:** Conceptualization, Visualization, Writing - review & editing. **Xuesong Wang:** Conceptualization, Visualization, Writing - review & editing.

Acknowledgements

This research was supported by the Iraqi Ministry of Higher Education and Scientific Research and the University of Anbar. This work would not have been possible without the OSU Driving Simulator in the School of Civil and Construction Engineering at Oregon State University.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.trf.2020.11.016.

References

- Abbas, M., & Machiani, S. G. (2016). Modeling the Dynamics of Driver's Dilemma Zone Perception Using Agent Based Modeling Techniques. IJT, 4(2), 1–14.
 Agbelie, B. R. D. K. (2016). Random-parameters analysis of highway characteristics on crash frequency and injury severity. Journal of Traffic and Transportation Engineering (English Edition), 3(3), 236–242.
- Agbelie, B. R. D. K. (2014). An empirical analysis of three econometric frameworks for evaluating economic impacts of transportation infrastructure expenditures across countries. *Transport Policy*, 35, 304–310.
- Aghabayk Eagely, S. K., Sarvi, M., Young, W., & Wang, Y. (2012). Investigating heavy vehicle interactions during the car following process. *Transportation Research Board 91st Annual Meeting, Transportation Research Board of the National Academies*, 1–13.
- Anderson, J. C., Hernandez, S., Jessup, E. L., & North, E. (2018). Perceived safe and adequate truck parking: A random parameters binary logit analysis of truck driver opinions in the Pacific Northwest. International Journal of Transportation Science and Technology, 7(1), 89–102.
- Awadallah, F. (2009). A Legal Approach to Reduce Red Light Running Crashes. Transportation Research Record, 2096(1), 102–107.
- Bhat, C. R. (2003). Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. *Transportation Research Part B: Methodological*, 37(9), 837–855.

Bonneson, J. A., & Son, H. J. (2003). Prediction of Expected Red-Light-Running Frequency at Urban Intersections. Transportation Research Record, 1830(1), 38-47.

- Bonneson, J.A., McCoy, P.T., Moen, B.A., 1994. Traffic detector design and evaluation guidelines. Nebraska Dept. of Roads, Rep. TRP-02-31-93, Lincoln, NE. Bonneson, J.A., Zimmerman, K., Brewer, M.A., 2002. Engineering countermeasures to reduce red-light-running. Texas Transportation Institute, Texas A & M University System, Report No. FHWA/TX-03/4027-2, College Station, TX.
- Burnett, G. E. (2008). Designing and evaluating in-car user-interfaces. In J. Lumsden (Ed.), Handbook of Research on User-Interface Design and Evaluation for Mobile Technology (pp. 218–236). Hershey: Idea Group Inc.
- Brehmer, C. L., Kacir, K. C., Noyce, D. A., & Manser, M. P. (2003). NCHRP Report 493: Evaluation of Traffic Signal Displays for Protected/Permissive Left-Turn Control. Washington, DC: Transportation Research Board of the National Academies.
- Caird, J. K., Chisholm, S. L., Edwards, C. J., & Creaser, J. I. (2007). The effect of yellow light onset time on older and younger drivers' perception response time (PRT) and intersection behavior. *Transportation Research Part F: Traffic Psychology and Behaviour*, 10(5), 383–396.
- Chang, M. S., Messer, C. J., & Santiago, A. J. (1985). Timing traffic signal change intervals based on driver behavior. *Transportation Research Record*, 1027, 20–30.
- Cobb, G.W., 1998. Introduction to Design and Analysis of Experiments (Textbooks in Mathematical Sciences). Springer Verlag New York, Inc., New York, USA. Dutta, A., Fisher, D. L., & Noyce, D. A. (2004). Use of a driving simulator to evaluate and optimize factors affecting understandability of variable message signs. *Transportation Research Part F: Traffic Psychology and Behaviour*, 7(4-5), 209–227.
- Elhenawy, M., Jahangiri, A., Rakha, H. A., & El-Shawarby, I. (2015). Classification of Driver Stop/Run Behavior at the Onset of a Yellow Indication for Different Vehicles and Roadway Surface Conditions Using Historical Behavior. *Procedia Manufacturing*, 3, 858–865.
- Elmitiny, N., Yan, X., Radwan, E., Russo, C., & Nashar, D. (2010). Classification analysis of driver's stop/go decision and red-light running violation. Accident Analysis & Prevention, 42(1), 101–111.

FHWA (2009). Manual on Uniform Traffic Control Devices (MUTCD) for Streets and Highways. Department of Transportation, Washington, D.C.: U.S.

- FHWA, 2014. Safety at Signalized Intersections. U.S. Department of Transportation, Washington, D.C. http://safety.fhwa.dot.gov/intersection/signalized/ presentations/sign_int_pps051508/short/index.cfm.
- Gates, T. J., Noyce, D. A., Laracuente, L., & Nordheim, E. V. (2007). Analysis of Driver Behavior in Dilemma Zones at Signalized Intersections. Transportation Research Record, 2030(1), 29–39.
- Gates, T. J., McGee, H. W., Moriarty, K. D., & Maria, H. U. (2012). A comprehensive evaluation of driver behavior to establish parameters for timing of yellow change and red clearance intervals. *Transportation Research Record*, 2298(1), 9–21.

Gates, T. J., & Noyce, D. (2010). Dilemma zone driver behavior as a function of vehicle type. *Transportation Research Record*, 2149(1), 84–93.

- Gates, T. J., & Noyce, D. A. (2016). A conceptual framework for dynamic extension of the red clearance interval as a countermeasure for red-light-running. Accident Analysis & Prevention, 96, 341–350.
- Greene, W.H., 2012. Econometric Analysis, 7th Ed. Prentice Hall, Boston. https://www.intechopen.com/books/fuzzy-logic-algorithms-techniques-andimplementations/understanding-driver-car-following-behavior-using-fuzzy-logic-car-following-model.

Hurwitz, D. S., Knodler, M. A., Jr., & Nyquist, B. (2011). Evaluation of Driver Behavior in Type II Dilemma Zones at High-Speed Signalized Intersections. Journal of Transportation Engineering, 137(4), 277–286.

Hurwitz, D. S., Wang, H., Knodler, M. A., Jr., Ni, D., & Moore, D. (2012). Fuzzy sets to describe driver behavior in the dilemma zone of high-speed signalized intersections. *Transportation Research Part F: Traffic Psychology and Behaviour*, *15*(2), 132–143.

Knodler, M. A., Jr., Noyce, D. A., Kacir, K. C., & Brehmer, C. L. (2005). Evaluation of Traffic Signal Displays for Protected-Permissive Left-Turn Control Using Driving Simulator Technology. Journal of Transportation Engineering, 131(4), 270–278.

Lavrenz, S. M., Dimitra Pyrialakou, V., & Gkritza, K. (2014). Modeling driver behavior in dilemma zones: A discrete/continuous formulation with selectivity bias corrections. Analytic Methods in Accident Research, 3-4, 44–55.

Li, J., Jia, X., Shao, C., 2016. Predicting driver behavior during the yellow interval using video surveillance. Int. J. Environ. Res. Public Health 13(12), 1213.
 Liu, Y., Chang, G.-L., & Yu, J. (2012). Empirical Study of Driver Responses during the Yellow Signal Phase at Six Maryland Intersections. *Journal of Transportation Engineering*, 138(1), 31–42.

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2000). Stated Choice Methods: Analysis And Applications. UK: Cambridge University Press.

McGee, H., Moriarty, K., Eccles, K., Liu, M., Gates, T., Retting, R., 2012. Guidelines for Timing Yellow and All-Red Intervals at Signalized Intersections. NCHRP Report No. 731. Transportation Research Board of the National Academies, Washington, D.C.

Mohammed, H. A., Hurwitz, D. S., & Smaglik, E. (2018). Variable driver responses to yellow indications: An operational challenge and safety concern. ITE Journal, 88(1), 44-49.

Aswad Mohammed, H., Hurwitz, D. S., & Macuga, K. L. (2020). Drivers' visual attention during the onset of the circular yellow indication at high-speed signalized intersections. *Traffic Injury Prevention*, 21(4), 259–264.

Moore, D., & Hurwitz, D. S. (2013). Fuzzy Logic for Improved Dilemma Zone Identification: Driving Simulator Study. *Transportation Research Record*, 2384(1), 25–34.

National Highway Traffic Safety Administration (NHTSA) (2017). *Traffic Safety Facts 2017 Data*. Department of Transportation, Washington, D.C.: U.S. Noyce, D. A., & Elango, V. V. (2004). Safety Evaluation of Centerline Rumble Strips: Crash and Driver Behavior Analysis. *Transportation Research Record*, 1862

(1), 44–53. Papaioannou, P. (2007). Driver behaviour, dilemma zone and safety effects at urban signalised intersections in Greece. Accident Analysis & Prevention, 39(1),

147–158. Parkes, A. (2012). Automotive Ergonomics: Driver-Vehicle Interaction. CRC Press. https://doi.org/10.1201/b13017-10.

Pathivada, B. K., & Perumal, V. (2017). Modeling Driver Behavior in Dilemma Zone under Mixed Traffic Conditions. *Transportation Research Procedia*, 27, 961–968.

Qu, D., Chen, X., Yang, W., & Bian, X. (2014). Modeling of Car-Following Required Safe Distance Based on Molecular Dynamics. Mathematical Problems in Engineering, 2014, 1–7.

Rakha, H., Amer, A., & El-Shawarby, I. (2008). Modeling Driver Behavior within a Signalized Intersection Approach Decision–Dilemma Zone. Transportation Research Record, 2069(1), 16–25.

Rakha, H., El-Shawarby, I., & Amer, A. (2011). Development of a Framework for Evaluating Yellow Timing at Signalized Intersections. Final Report VCTIR 11–R12. Charlottesville, VA: Virginia Center for Transportation Innovation and Research.

Rakha, H., El-Shawarby, I., & Setti, J. R. (2007). Characterizing Driver Behavior on Signalized Intersection Approaches at the Onset of a Yellow-Phase Trigger. *IEEE Trans. Intell. Transport. Syst.*, 8(4), 630–640.

Santos, J., Merat, N., Mouta, S., Brookhuis, K., & de Waard, D. (2005). The interaction between driving and in-vehicle information systems: Comparison of results from laboratory, simulator and real-world studies. *Transportation Research Part F: Traffic Psychology and Behaviour*, 8(2), 135–146.

Sato, T., & Akamatsu, M. (2012). Chapter 13: Understanding driver car-following behavior using a fuzzy logic car-following model. In Fuzzy Logic-Algorithms (pp. 265–282). Japan: Techniques and Implementations. In Tech.

Savolainen, P. T. (2016). Examining driver behavior at the onset of yellow in a traffic simulator environment: Comparisons between random parameters and latent class logit models. Accident Analysis & Prevention, 96, 300–307.

Senserrick, T.M., Brown, T., Marshall, D., Quistberg, D.A., Dow, B., Winston, F.K., 2007. Risky driving by recently licensed teens: self-reports and simulated performance. In: Proceedings of the Fourth International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Stevenson, Washington, 541–548.

Sharma, A., Bullock, D., & Peeta, S. (2011). Estimating dilemma zone hazard function at high speed isolated intersection. *Transportation Research Part C:* Emerging Technologies, 19(3), 400-412.

Sharples, S., Shalloe, S., Burnett, G., & Crundall, D. (2016). Journey decision making: The influence on drivers of dynamic information presented on variable message signs. *Cogn Tech Work*, 18(2), 303–317.

Swake, J., Jannat, M., Islam, M., & Hurwitz, D. (2013). Driver response to phase termination at signalized intersections: are driving simulator results valid. Bolton Landing, NY: and Vehicle Design.

Van der Horst, R., and A., 1986. Wilmink. Drivers' decision-making at signalised intersections: An optimisation of the yellow timing. Traffic Engineering and Control, Vol. 27, No. 12, 615–622.

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). Statistical and Econometric Methods for Transportation Data Analysis (2nd Ed.). New York: Chapman & Hall/CRC.

Webster, F. V., & Elison, P. B. (1965). Traffic signals for high-speed roads. RRL Technical Paper 74. U.K Transport and Road Research Laboratory.

Xiong, Huimin, Narayanaswamy, Prabha, Bao, Shan, Flannagan, Carol, & Sayer, James (2016). How do drivers behave during indecision zone maneuvers? Accident Analysis & Prevention. 96, 274–279.

Young, Kristie, Regan, Michael, & Lee, John (2008). Driver Distraction: Theory, Effects, and Mitigation. CRC Press. https://doi.org/10.1201/9781420007497.ch7. Zegeer, C. V., & Deen, R. C. (1978). Green-extension systems at high-speed intersections. *ITE Journal*, 496, 19–24.