

Research Article

Analysis of Driver Decisions at the Onset of Yellow at Signalized Intersections

Juan Li ^{1,2}, Boyu Jiang,^{1,2} Chunjiao Dong ^{1,2}, Jue Wang ³ and Xuan Zhang ⁴

¹School of Traffic and Transportation, Beijing Jiaotong University, Beijing, China

²Key Laboratory of Transport Industry of Big Data Application Technologies for Comprehensive Transport, Beijing Jiaotong University, Beijing, China

³Center for Forecasting Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing, China

⁴Sam's Club Technology, Austin, TX, USA

Correspondence should be addressed to Juan Li; juanli@bjtu.edu.cn

Received 11 October 2019; Revised 15 January 2020; Accepted 13 February 2020; Published 5 August 2020

Guest Editor: Inhi Kim

Copyright © 2020 Juan Li et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Drivers' decisions to either slow and stop or go at the onset of yellow signal impact on intersection safety. This novel study contributes to the new classification scheme for drivers. Two driving style indexes (i.e., the driving reliability index and dangerous driving index) are adopted, along with other known factors to analyze stop/go decision-making. Initially, the driving reliability index is extracted using a Hidden Markov Model (HMM). The dangerous driving index is calculated based on statistics extracted from dangerous driving records. A latent class logit model is then proposed to explore the factors which influence drivers' decisions. Drivers are classified for analytical purposes into "low-risk" and "high-risk" categories according to driving styles and age. Results indicate that those considering "low-risk" tend to stop, while drivers considering "high-risk" are inclined to pass intersections. Furthermore, distractions from cell phones have different influences on each group of drivers. These findings help to determine driver preferences and may be used to formulate strategies to reduce unsafe driving occurring at signalized intersections.

1. Introduction

Signalized intersections are areas where traffic accidents occur frequently. In 2017, there were 34,247 fatal traffic accidents in the United States, involving 52,274 drivers and 37,133 deaths [1]. Over 50 percent of all injuries and fatalities occur at or near signalized intersections [2] and driver errors are the leading cause of intersection-related crashes [3]. At the onset of yellow signal, it is a challenge for drivers to make immediate decisions, especially in yellow light dilemma zones, which are where drivers may neither stop safely nor pass through intersections [4, 5]. Unfortunately, the decision to either stop or go may increase the number of angle crashes and rear-end collisions, as well as injuries and fatalities, which necessitates research into stop/go decision-making at the onset of yellow at signalized intersections.

In order to understand drivers' stop/go decision-making completely, researchers have tried to find demographic attributes besides age and gender. For example, Elhenawy et al. proposed a variable to measure drivers' aggressiveness level at the onset of yellow signal [6]. In this study, we adopt two new variables, i.e., the driving reliability index and dangerous driving index to develop this further. It is hoped that implementing these two new variables will provide a more comprehensive understanding of driving styles. We also conduct sophisticated analysis of heterogeneity, which is necessary considering diversity within any sample involving human participants.

In the past, researchers have used various statistical approaches to gain insight into driving heterogeneity. For example, Savolainen [7] adopted the latent class logit model instead of the fixed logit model. Constant terms were used to formulate classification probabilities. In this study, we

adopted a latent class model with newly proposed variables which contributes in three aspects: (1) involving heterogeneity into drivers' decisions; (2) evaluating the impact of factors on different-class drivers; (3) analyzing the relationship between decision property and driving styles.

The contribution of this study is twofold. Firstly, to uncover class features of drivers' behavior, two novel indexes associated with driving styles (i.e., the driving reliability index and dangerous driving index) are proposed. Unlike previously considered factors (i.e., drivers' demographic attributes, speed, accelerate, and distant to stop line), these two factors are extracted from historic driving records and able to reflect driving styles. Secondly, the latent class logit model categorizes drivers into two groups according to driving style indexes. Compared with traditional logit models, this model considers the group heterogeneity of drivers and thus can distinguish their choice-preference differences.

The remainder part of this paper is organized as follows: Section 2 presents relevant research around factors which influence stop/go decision-making and driving heterogeneity. Section 3 introduces experimental data and provides a description of the analytical dataset. Section 4 provides a detailed description of the methodologies adopted in this study. Section 5 presents model results and the quantified impact of variables. Section 6 discusses the driving risks, distracted driving behavior among different driver classes, and comparative results. The final section, Section 7, provides our conclusions based upon the findings.

2. Literature Review

2.1. Indicators Influencing Driver Decisions. Previous researches focusing on driver decisions at the onset of yellow signal have adopted various data-gathering methods, such as field studies [8, 9], driving simulator studies [6, 10], and video-capture studies [11–14]. Researchers have identified a number of factors that influence driver decisions, such as age, gender, and time to the stop line. The dilemma zone is one of the most important factors involved in stop/go decisions in this scenario.

Type I dilemma zone describes the region where drivers can neither stop safely nor pass through intersections. It has been postulated that this occurs due to insufficient yellow light time caused through high approaching speeds [15]. Type I dilemma zone is therefore formulated using vehicular information and intersection data, such as approach speed, acceleration, and yellow interval [4, 16].

Type II dilemma zone is defined as the zone in which drivers may have trouble making stop/go decisions at the onset of yellow signal [17]. This area can be thought of as a corridor of uncertainty, which is the area ahead of the stop line between the point where 90 percent of drivers will stop and the point where 90 percent of drivers will continue [18]. The existence of what has been described as Types I and II of dilemma zones which occur as drivers approach intersections adds to the complexity of driver's decision-making. However, Types I and II of dilemma zones are not the only

situations that increase cognitive load, clouding decision-making.

Distractions are another factor that may influence driving behaviors at the onset of yellow signals. In 2017 in the United States, 3,166 lives were claimed related to driving distractions. Of this total number, 434 people died in fatal crashes involving cell phone use or other cell phone related behaviors [1]. By splitting visual, manual, and cognitive attention, speaking on the phone distracts drivers from the road and therefore increases the likelihood of having an accident and risk of injury to drivers, passengers, and pedestrians [19, 20]. Interestingly, drivers appear more likely to stop when using handheld and hands-free devices, but less likely to stop when using headsets [7].

Several studies have examined other factors associated with drivers' decisions at the onset of yellow, including demographic characteristics such as age and gender, and vehicular conditions such as speed, distance to the stop line, and time to the intersection [6, 12, 21, 22]. Savolainen [7] found that driver decisions are mainly determined by the estimated time to the stop line.

2.2. Driving Heterogeneity. Previous researches have used the binary logit model to analyze drivers' behaviors at signalized intersections. Köll et al. [12], for example, adopted a binary logit model to examine the effect of speed, distance, and potential time at the point at which drivers decide to either stop or proceed. They found that both high speeds and short distances to the stop line decreased the probability of stopping. Papaioannou [23] extended the binary logit model intercalating gender, and evidence suggests that female drivers were less aggressive and more likely to stop when encountering yellow signals. Gates et al. [24] discovered that heavy vehicles were more likely to pass the intersection than passenger vehicles.

Further research has also looked to analyze the influence of the traffic environment, such as pavements and traffic control devices. For example, Yan et al. [25] examined the impact of pavement markings and found that they appear to reduce the probability of both conservative-stop and risky-go decisions. Long et al. [26] employed a binary logit model to identify the influence of countdown devices, which increased the probability of passing the stop line after the onset of yellow signal. Importantly, in the conventional binary logit model, utilities of indicators remain constant across individuals, which means the logit model is incapable of accommodating unobserved heterogeneity between individuals. Therefore, alternative analytical methods which are capable of intercalating individual heterogeneity should be considered.

Latent class analysis was proposed to overcome the limitations of the fixed model [27–29]. The latent class logit model is a semiparametric extension of the logit model which accommodates heterogeneity across individuals with a set of classes and without parametric distribution definition [30]. Various fields of study have adopted this model to evaluate human preferences [31–33]. In the field of transportation, Hess et al. [34] employed the latent class

model to analyze rail and bus travel behavior. Shen [35] used the latent class logit model to predict transport mode choices.

In this study, we make an initial attempt to introduce driving style indexes (i.e., the driving reliability index and dangerous driving index) to examine drivers' stop/go decision-making. A latent class logit model with these new indicators is developed to investigate the influence of each factor on driver subgroups. Also, each class is labeled with a unique driving style, which is necessary for analyzing the relationship between decision preference and driving styles.

3. Data and Analysis

The dataset was collected through the National Advanced Driving Simulator (NADS) at the University of Iowa [10]. In driving simulations, each driver partook in three "drives," where each "drive" consisted of three "segments." Each "segment" contained one rural zone and one urban zone. Each driver encountered five signal-controlled intersections in each "segment," only two of which triggered yellow signals when the driver was approaching. Each driver therefore encountered 18 points which required stop/go decisions. In each "segment," every driver accomplished one of the following three tasks: baseline (no phone call), outgoing (making calls), or incoming (answering calls). These tasks were randomly arranged within each driving "segment" and began prior to the arrival at each segment. This simulation experiment focused on identifying differences in driver decisions at the onset of yellow signal with and without cell phone distractions.

Data were recorded from 49 participants and contained 1157 trials across 17 variables. After deleting the missing and invalid data which were beyond predefined ranges, the final dataset included 829 complete runs. Deleted trials were consistent across age, gender, drive number, and phone status due to the simulator settings. The description of variables in the final dataset is presented in Tables 1 and 2.

According to the data in Table 1, most drivers decided to stop in the simulation experiment. The number of trials in the old age group was the least among all the ages with 304 occurring in the young age group, 295 in the middle-aged group, and 230 in the old age group. Trials with male drivers ($n = 432$) slightly outnumbered those of female drivers ($n = 397$). Also, the 829 trials were roughly balanced in the drive number and phone status. Stop times increased as the number of trials increased with 162 in the first drive, 176 in the second drive, and 194 in the third drive. In addition, the number of stop times under incoming ($n = 181$) or outgoing calls ($n = 178$) was slightly larger than under baseline conditions ($n = 173$).

In Table 2, speeds at the onset of yellow ranged from 11 m/s to 24.12 m/s. The distance to the stop line also ranged from 33.49 m to 86.76 m. Under experimental settings, the duration of the yellow light was triggered at either 3 seconds or 3.75 seconds from the stop line [36]. Given variations in the simulator algorithm, the yellow time varied from 2.78 seconds to 4.38 seconds.

TABLE 1: Distribution of category variables by drivers' decisions in the dataset.

Variables	Go	Stop	Total	
Age	Young group (18–25 years)	100	204	304
	Middle group (30–45 years)	98	197	295
	Old group (50–60 years)	99	131	230
Gender	Male	142	290	432
	Female	155	242	397
Drive number	First drive	114	162	276
	Second drive	98	176	274
	Third drive	85	194	279
Phone status	Baseline	105	173	278
	Incoming	96	181	277
	Outgoing	96	178	274

4. Methodology

A Hidden Markov Model (HMM) and a latent class logit model are devised to analyze drivers' decisions at the onset of yellow at signalized intersections. Initially, the HMM was established to compute a driving reliability index. Then, the dangerous driving index was obtained from the driving records. Finally, a latent class logit model was developed and calibrated based on age and the newly acquired factors of driving styles (i.e., driving reliability index and dangerous driving index). A complete description of model development is provided in Figure 1.

4.1. Hidden Markov Model. During simulations, driver behaviors were continuously tracked to contain a full set of stop/go decisions. Each driver made 18 decisions, where each decision was related to the previous one. The HMM is therefore the rational model to analyze drivers' stop/go decisions because each driver's decision regarding stop and go is unobservable and is therefore a constituent of a hidden state. Obtained vehicular data is denoted as an observable variable.

The HMM is a statistical model, which consists of N hidden states, M observable states, an initial state probability distribution π , a state transition probability matrix A , and an emission matrix B [37]. Model details are described as follows:

$$A = [a_{ij}]_{N \times N},$$

$$a_{ij} = P(i_{t+1} = q_j | i_t = q_i), \quad i = 1, 2, \dots, N; j = 1, 2, \dots, N,$$
(1)

where q represents the individual state, i_t is the state symbol at time t , and a_{ij} represents the transitional probability from state q_i to state q_j :

$$B = [b_{jk}]_{N \times M},$$

$$b_{jk} = P(o_t = v_k | i_t = q_j), \quad j = 1, 2, \dots, N; k = 1, 2, \dots, M,$$
(2)

TABLE 2: Descriptive statistics for continuous variables in the dataset.

Variables	Min.	Max.	Mean	Std. dev.
Speed at the stop line (m/s)	0.09	23.17	6.79	7.25
Speed at the onset of red signal (m/s)	0	23.26	8.10	6.71
Speed at the onset of yellow signal (m/s)	11.00	24.12	19.01	2.24
Distance to stop line at the onset of yellow signal (m)	33.49	86.76	62.22	10.20
Time to stop line at the onset of yellow signal (s)	2.48	3.81	3.27	0.38
Yellow signal duration (s)	2.78	4.08	4.00	0.22
Average acceleration rate (m/s ²)	-7.13	3.33	-1.37	1.95

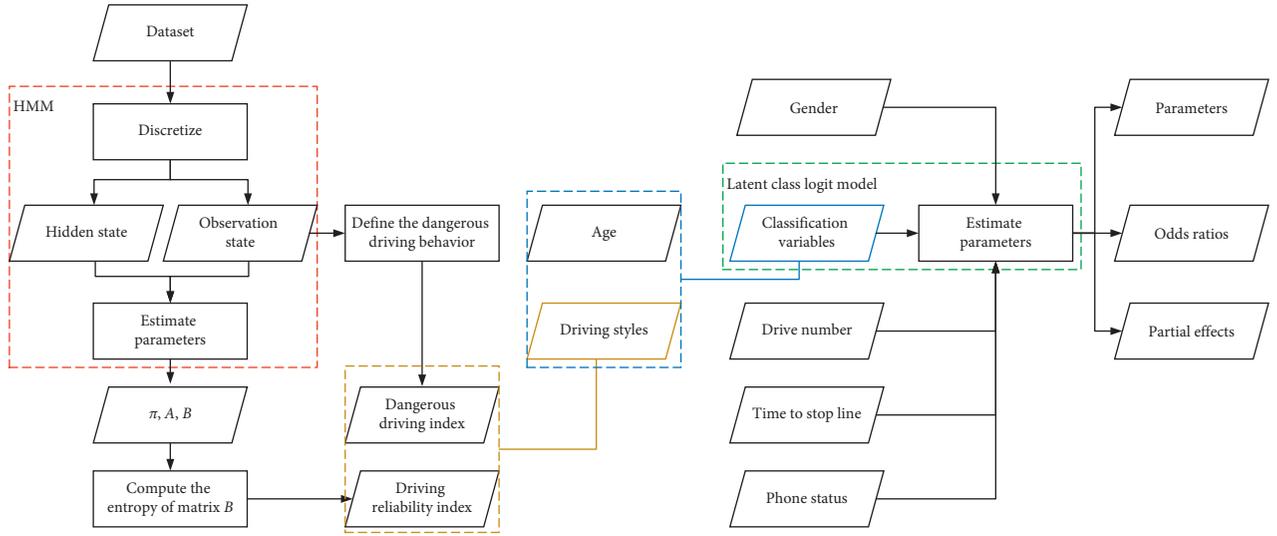


FIGURE 1: Model development flow chart.

where v is the possible observation state, o_t is the observation symbol at time t , and b_{jk} represents the probability of the state q_j at time t with the observation v_k :

$$\begin{aligned} \pi &= (\pi_i), \\ \pi_i &= P(i_1 = q_i), \quad i = 1, 2, \dots, N, \end{aligned} \quad (3)$$

where π_i is the probability of q_i being the state at time $t = 1$. π and A produce the state sequence, and B derives the observation sequence. HMM can be described as π, A , and B . Although in an HMM, the hidden state cannot be observed directly, the output, derived from the state, can be observed. State estimation is therefore derived through a probability distribution across the full range of possible outputs. In HMM, the sequence of outputs provides essential information for model estimations [38]. There are two kinds of methods to estimate π, A , and B : the supervised learning algorithm and the unsupervised learning algorithm. In this study, the parameters (π, A , and B) were estimated using an unsupervised approach, known as the expectation maximization (EM) algorithm.

4.2. Indicators Associated with Driving Styles

4.2.1. Driving Reliability Index. In order to account for the level of driving reliability, we use driving reliability index to characterize driving styles. When encountering yellow

signals, driver's manipulation (e.g., acceleration and deceleration) appears similar, which means that his or her behavior is relatively stable and predictable.

An HMM was employed to calculate the driving reliability index. Each row of the estimated emission probability matrix B represents the probabilities of each observation state under one hidden state. Therefore, if a single driver presented many driving states in the same traffic scene, this would mean that his/her behavior was unstable, and the corresponding B would become a dense matrix. The denser a matrix B is, the larger the driver's driving reliability index would be. Information *entropy* within each row of a matrix was used to describe the sparsity of matrix B . The normalized entropy of the matrix B was calculated as follows:

$$\begin{aligned} H_j &= -(\log_2 M)^{-1} \sum_{k=1}^M (b'_{jk} \times \log_2(b'_{jk})), \\ b'_{jk} &= \begin{cases} b_{jk}, & b_{jk} \neq 0 \\ 1, & b_{jk} = 0 \end{cases}, \end{aligned} \quad (4)$$

$$j = 1, 2, \dots, N; k = 1, 2, \dots, M,$$

where H_j is the entropy of row j in matrix B and b_{jk} stands for the probability of the observation state k under the hidden state j in B . Although each driver had more than one hidden

state (i.e., a decision), the driving reliability index was defined according to arithmetic means of the entropy of each row in matrix B :

$$\text{driving reliability index} = \frac{1}{N} \sum_{j=1}^N H_j \quad (5)$$

4.2.2. Dangerous Driving Index. Dangerous driving index is proposed to represent the level of driving risk, which is also an indicator of driving style. Traffic records imply that drivers' accident proneness exists, and the proneness appears to be sustainable under various traffic circumstances [39]. Therefore, the risk of drivers' behavior can be identified by examining driving records. According to Farmer and Chambers [39], one drive trial can be categorized as either dangerous or safe. Dangerous driving behavior is more likely to result in traffic accidents, and therefore driver behavior can be evaluated by the probability of dangerous driving using driving histories found within records, which creates the dangerous driving index.

In this study, the dangerous driving index was defined as the probability of their dangerous behavior during the experiment, which was computed as follows:

$$\text{dangerous driving index} = \frac{1}{h} \sum_{d=1}^h x'_d \quad (6)$$

where h represents the number of trials for each driver and x_d equals 1 when the driver's behavior in the d th trial is considered dangerous and 0 when the behavior is relatively safe.

4.3. Latent Class Logit Model. The binary logit model is a classical method used to study drivers stop/go decisions. While in the conventional binary logit model, there is a potential problem with the estimation of parameters. The impact of heterogeneity makes drivers more likely to either stop or go at the onset of yellow signal. Also, this hidden heterogeneity can lead to biased parameter estimations. Hence, the latent class logit model was proposed to intercalate heterogeneity across individuals.

The latent class logit model is an improved model of the conventional logit model. The utility function of the logit model is written as

$$S_{ij} = \beta' x'_{ij} + \varepsilon_{ij} \quad (7)$$

where S_{ij} is the utility function that determines the probability of decision outcome j for individual i ; β' is a vector of parameters; x'_{ij} is a vector of observed variables; and ε_{ij} is the error term, which is independent and follows Gumbel distribution. The probability (P_{ij}) of one alternative j (stop/go) for individual i is defined as

$$P_{ij} = \frac{\exp(\beta' x'_{ij})}{1 + \exp(\beta' x'_{ij})} \quad (8)$$

The logit model was estimated with maximum likelihood estimation (MLE) procedures.

In the latent class model, heterogeneity is modeled using a set of groups, otherwise known as classes. Specifically, each individual is distributed to a "latent" class. It is assumed that parameters within each class contain the same effects, but there are different effects across classes. Model estimations are split into parameters related to each class, and a set of class probabilities [40]. Within the class, choice probabilities are calculated using the multinomial logit model:

$$P_{j|q} = \frac{\exp(\beta'_q x'_{ij})}{\sum_{j=1}^J \exp(\beta'_q x'_{ij})}, \quad q = 1, 2, \dots, Q, \quad (9)$$

where Q is the number of latent classes, J represents the alternatives, x'_{ij} is the vector of all variables in the utility function, and β'_q is the class-specific parameter vector. The class for a specific individual is unobservable. Class probabilities are therefore generated using the multinomial logit form, as follows:

$$C_{iq} = \frac{\exp(\theta'_q z'_i)}{\sum_{q=1}^Q \exp(\theta'_q z'_i)}, \quad \theta'_Q = 0, \quad (10)$$

where z_i is an optional set of personal invariant characteristics. In this study, z_i denotes demographic and driving styles. For estimation, the last of parameter θ_Q was fixed as zero. For an individual, the estimation of the probability of a specific choice is the expected value (over classes) of the class-specific probabilities:

$$P_{ij} = \sum_{q=1}^Q C_{iq} P_{j|q} \quad (11)$$

The number of classes, Q , is fixed, which is generally determined by setting an initial value and adjusting it. However, extended classes may not yield the best estimations and can sometimes create model instability and divergence [30]. In this study, the latent class logit model provides the best fit when Q was set to 2.

4.4. Model Development

4.4.1. HMM Estimates. HMM was established to extract the driving reliability index, one of driving style indexes. To establish the observation sequence, three variables were selected: speed, average acceleration rate, and phone status, which are proven to influence drivers' decision-making behaviors [41–44]. As illustrated in Table 3, the data of observed variables were discretized into three categories. Discretization thresholds for speed and average acceleration rate were the 50th and 85th percentiles. The observable state in HMM was the combination of the discrete data including 27 values, signed by a number from 1 to 27, respectively. Table 4 shows part of the ordered sequence of this observed data combination. Each trial was denoted with a corresponding figure which created a sequence. After coding the input sequence, the HMM for each driver was estimated using the EM algorithm.

TABLE 3: Rules of data discretization.

Speed Before (m/s)	Average acceleration rate		Phone status	
	After	Before (m/s ²)	After	Before After
[11, 19.49)	1	(-2.24, 1.34]	1	Baseline 1
[19.49, 21.14)	2	(-4.02, -2.24], (1.34, 2.24]	2	Incoming 2
[21.14, 24.12]	3	[-7.13, -4.02], (2.24, 3.33]	3	Outgoing 3

TABLE 4: Rules of the observed combination sequence (partial).

Sequence number	Speed	Average acceleration rate	Phone status
1	1	1	1
2	1	1	2
3	1	1	3
4	1	2	1
5	1	2	2
6	1	2	3
7	1	3	1
8	1	3	2
9	1	3	3

4.4.2. Calculate Indicators Associated with Driving Styles.

After estimating HMM, it becomes possible to calculate the driving reliability index. Before calculating the dangerous driving index, the dangerous drive needs to be identified. Telephone conversations (i.e., incoming calls or outgoing calls) and sudden braking/accelerating (i.e., average acceleration rate greater than 1.34 m/s^2 or less than -2.24 m/s^2) were distinguished as dangerous behaviors. There were 12 dangerous states described with the observation combination sequence as shown in Table 4: (1, 2, 2), (1, 2, 3), (1, 3, 2), (1, 3, 3), (2, 2, 2), (2, 2, 3), (2, 3, 2), (2, 3, 3), (3, 2, 2), (3, 2, 3), (3, 3, 2), and (3, 3, 3). Corresponding sequence numbers were 5, 6, 8, 9, 14, 15, 17, 18, 23, 24, 26, and 27. Therefore each trial which involved dangerous behavior was identified. This key analytical step is the basis of calculating the dangerous driving index.

4.4.3. Estimation of the Latent Class Logit Model.

Table 5 presents variables in the latent class logit model, including drivers' decisions, basic demographics (i.e., gender and age), driving styles (i.e., driving reliability index and dangerous driving index), and experimental variables such as drive number, phone status, and predicted time to stop line. It should be noted that the "predicted time to the stop line" was calculated by dividing the current distance from the stop line with the instant speed at the onset of yellow signal. According to the definition of the latent class logit model, variables for latent classification need to present demographics and driving characteristics. In this study, three variables were selected for latent classification—age, driving reliability index, and dangerous driving index.

5. Results

Table 6 provides estimates for the two new indicators of driving styles. The driving reliability index ranged from 0.465 to 0.692, while the dangerous driving index fluctuated from 0 to 0.611. For instance, the driving reliability index for Driver #2 had a maximum value of 0.692 which means that his/her driving decision was relatively unpredictable. Driver #2's dangerous driving index was 0.222 which was lower than the average. This means the probability of dangerous driving behavior was only 22.2%; therefore it can be understood that this driver's behavior in these trials was comparatively safe. Given that each indicator was different across drivers, these two indicators can be regarded as driver identifiers and further used for latent classification.

Estimation results are provided in Table 7, including a binary logit model without new indicators, a binary logit model with new indicators, and a latent class logit model with new indicators. When investigating fitting criterion indicators, the latent class model got 0.105 for the McFadden R -squared and 996.3 for the Akaike information criterion (AIC). These suggest that the latent class model provides the best fit and additionally explained attributes for each of the members within the class. Also, the McFadden R -squared of the binary logit model without the new predictors was 0.02, while this figure rose to 0.028 after adding the new indicators. The AIC of the binary logit model without new indicators is 1076.1, and it drops to 1071 after adding the new indicators. This suggests the derived estimations validate the precision of two indicators.

5.1. Direction and Magnitude of Parameters. When considering the results in Table 7, one might conclude that directions of parameters in both binary logit models and the latent class logit model are generally consistent. However, there are significant differences in the magnitude of these effects. Parameter estimations for the latent class model highlight the differences between drivers. Statistically significant parameters of the latent classification variables (i.e., age and dangerous driving index) indicate that these indicators influenced class probabilities. Also, the indicator, "DrR_Index," derived using HMM is statistically significant ($\alpha = 0.01$) in the binary logit model, which indicates that this is an additional attribute to affect the decision-making process of the participants.

The directions of most parameter estimation in two classes are opposing, which highlights the differences between groups. Class I accounts for 53.3% of all participants. Visually, the parameter of the indicator "Time" is negative. The parameter of "Time_Baseline" (-2.395) is less than that of "Time_Income" (-2.468) or "Time_Outgo" (-2.493). This may have occurred because nonsignificant parameters for the second/third driving experiment had no apparent inclination of stopping or going as the experiment commenced.

Class II accounted for 46.7% of the participants. The parameter of indicator "Time" is positive and the parameters for "Time" under different conditions are almost the same

TABLE 5: Descriptions for selected variables.

Variables	Description	Values
<i>Dependent variables</i>		
Drivers' decisions	Stop Go	0 = stop 1 = go
<i>Independent variables</i>		
Gender	Male Female	0 = male 1 = female
Age	Young group (18–25 years) Middle-aged group (30–45 years) Old group (50–60 years)	0 = young and middle-aged groups 1 = old group
2 nd _drive	Driver's second "drive" experiment	0 = others 1 = the second "drive" experiment
3 rd _drive	Driver's third "drive" experiment	0 = others 1 = the third "drive" experiment
Time_Baseline	Expected time to stop line without phone calls	Continuous
Time_Income	Expected time to stop line with incoming phone calls	Continuous
Time_Outgo	Expected time to stop line with outgoing phone calls	Continuous
DrR_Index	Driving reliability index	Continuous
Ddr_Index	Dangerous driving index	Continuous

TABLE 6: Indicators associated with driving styles (partial).

	Min.	Max.	Mean	Driver ID									
				1	2	3	4	5	6	7	8	9	10
Driving reliability index	0.465	0.692	0.601	0.558	0.692	0.619	0.533	0.560	0.649	0.603	0.686	0.465	0.587
Dangerous driving index	0	0.611	0.233	0.167	0.222	0.154	0.056	0.111	0.529	0.167	0.471	0	0.118

TABLE 7: Parameter estimation results for models.

Parameter	Binary logit model without new indicators	Binary logit model with new indicators	Latent class logit model	
			Class I	Class II
Constant	0.460 (0.651)	-2.504 (1.214)**	1.498 (2.419)	2.695 (2.755)
Gender	0.306 (0.147)**	0.388 (0.151)**	2.038 (0.551)***	-1.303 (0.325)***
2 nd _drive	-0.233 (0.177)	-0.252 (0.178)	0.002 (0.355)	-0.709 (0.334)**
3 rd _drive	-0.465 (0.180)***	-0.474 (0.181)***	-0.142 (0.362)	-0.996 (0.349)***
Time_Baseline	-0.316 (0.196)	-0.328 (0.198)*	-2.395 (0.688)***	1.578 (0.424)***
Time_Income	-0.353 (0.198)*	-0.367 (0.199)*	-2.493 (0.719)***	1.560 (0.430)***
Time_Outgo	-0.340 (0.198)*	-0.353 (0.199)*	-2.468 (0.705)***	1.594 (0.433)***
Age	0.460 (0.161)***	0.553 (0.165)***	-3.889 (1.480)***	0
DrR_Index	—	5.144 (1.837)***	-10.479 (8.695)	0
Ddr_Index	—	-0.630 (0.597)	-11.513 (3.486)***	0
Constant_Class	—	—	10.186 (6.019)*	0
Class probabilities	—	—	0.533 (0.403)	0.467 (0.403)
Log likelihood	-530.06	-525.50	-478.17	
McFadden R-squared	0.020	0.028	0.105	
AIC	1076.1	1071.0	996.3	

Tabular values indicate parameter estimates (standard errors) for each model. The symbols ***, **, and * mean statistical significance at $\alpha = 0.01, 0.05, 0.1$, respectively.

(1.578, 1.560, and 1.594). Drivers in Class II were perhaps influenced by the simulator environment because "2nd_drive" and "3rd_drive" indicators were significant at $\alpha = 0.05$ and 0.01, respectively.

5.2. Quantified Impact of Variables. To quantify the impact of these variables, partial effects were calculated as follows:

$$\frac{\partial E[y | x]}{\partial x} = \frac{\partial F(\beta'x)}{\partial x} = \frac{dF(\beta'x)}{d\beta'x} \beta = f(\beta'x)\beta, \quad (12)$$

where $f(\beta_1 x)$ is the density function of x . When the variable in x is a dummy variable, the alternative method is

$$\begin{aligned}\Delta F_{x_k} &= \text{Prob}[y = 1 | x_k = 1] - \text{Prob}[y = 1 | x_k = 0] \\ &= F(\beta_1 x + \alpha x_k | x_k = 1) - F(\beta_1 x + \alpha x_k | x_k = 0) \\ &= F(\beta_1 x + \alpha) - F(\beta_1 x).\end{aligned}\quad (13)$$

Partial effects were calculated by averaging the function over the sample observations. Findings suggest an average change in the probability of going for a 1-unit change in scale factor, or the change in the probability of going compared with the baseline category in dummy factor. Table 8 presents partial effects and odds ratios for each model.

From Table 8, the effects of indicators on different classes can be identified. In Class I, the odds ratios of “Time” indicators which ranged from 0.08 to 0.09, are all less than 1. This means that these drivers were more likely to stop, as the time to the stop line increases. Odds ratios of women in Class I making the “go” decision were 7.68 times those of men in the same class.

In Class II, odds ratios of “Time” indicators ranged from 4.76 to 4.92 and were larger than 1. This suggests that these drivers tend to pass the intersection as the time to stop line increases. Odds ratios for “go” during the 2nd experimental phase were 0.49 times those during the first one, and this figure decreased to 0.37 in the 3rd experiment. This suggests that drivers in Class II generally prefer to stop as experiment proceeds. Odds ratios for women in Class II making the “go” decision were 0.27 times those of men in the same class.

In terms of partial effects of three models, several conclusions can be made. Female drivers are 6.9%–10.1% more likely to choose to go compared to male drivers. Old drivers are 10.6%–12.6% more likely to choose to go than young and middle-aged drivers. Compared with 1st “drive,” drivers were 5.2%–7.0% and 10.2%–11.5% less likely to choose to go in their 2nd and 3rd “drive” experience. There was also a 114% increase in the probability of going which was estimated with every 1-unit increase in the driving reliability index. Each 1-second increase in the time to stop line manifested in a 7.1%–12.7% decreased probability of going.

Considering the phone as a distraction related to stop/go behavior in these circumstances, there was relative conformity with empirical judgment. Partial effects of the three models suggest that compared with uninterrupted driving, incoming/outgoing calls increase the probability of stopping slightly. When investigating the effect on drivers in a specific class, the odds ratios of “Time” under different calling tasks were almost equivalent. Further details about the effect of phone distractions will be critically discussed in the following section.

6. Discussion

6.1. Driving Risk of Different Classes. Analysis using the latent class model indicates that two classes of drivers with different driving styles exist. The negative parameters of

latent classification variables (−3.889 for “Age” and −11.513 for “Ddr_Index,” significant at $\alpha=0.01$) in Table 7 suggest that increasing these values contributes to a higher probability of Class II categorization. Additionally, the data provided in Figure 2 validates that as drivers’ age or the value of drivers’ dangerous driving index increases, the probability of Class II increases. Ultimately, this means that those who were categorized as Class II were old and more likely to display risky behaviors, such as sudden acceleration and braking.

Before discussing drivers’ decision properties, experimental settings of triggering yellow signals need to be specified. As drivers approached simulated intersections, traffic signal began to change to yellow at one of the two preset intervals (i.e., 3 seconds or about 3.75 seconds). From the designers’ perspective, the first interval of 3 seconds was intended to elicit a “go” response from the participant, whereas the second interval of 3.75 seconds was intended to elicit a “stop” response. In other words, the probability of going at around 3 seconds was expected to be larger than that at 3.75 seconds. This enabled us to determine what was considered *safe* decisions.

Drivers’ decision properties can be identified and evaluated using data provided in Figure 3. The distribution of dots for Class I individuals shows that these drivers are more likely to go through the intersection when the time to stop line is less than 3 seconds. Although the same drivers are also more likely to halt further when the time to stop line is 3.75 seconds, this decision appears identical to the expected safe decision probably because these drivers appear to prefer the low-risk choice. Decisions of Class II drivers are comparatively inconsistent. The probability of going under the 3.75 seconds interval is equal or larger than that observed in the 3-second interval. This result may be interpreted according to their respective high-risk driving characteristics.

Considering driving styles and stop/go decision property discussed above, drivers in Class I can be confirmed as “low-risk drivers,” while drivers in Class II are defined as “high-risk drivers.”

6.2. Distraction Effects on Different Classes. The effects of cell phone distractions on stop/go decisions were evaluated using changes in probability. For instance, when drivers were near the intersection with a time to the stop line of fewer than 3 seconds, the probability of going with an incoming call decreased compared with no call scenarios. This would suggest that drivers receiving a call are more likely to slow and stop even though passing the intersection was actually the safer choice.

Figure 4 demonstrates the effects of calling distraction on drivers’ decisions. “Low-risk” drivers, categorized as Class I, were disturbed only when the time to stop line was less than 3 seconds. Compared with uninterrupted driving, drivers in Class I with phone calls appeared to prefer to stop, which may actually be improper because passing the intersection was the safer choice at this point. This effect was the same under both incoming and outgoing calls. The influence of

TABLE 8: Quantified impact of variables.

Variables	Binary logit model without new indicators		Binary logit model with new indicators		Latent class logit model		
	Partial effect (%)	Odds ratio	Partial effect (%)	Odds ratio	Partial effect (%)	Odds ratio	
						Class I	Class II
Gender	+6.9	1.36	+8.6	1.47	+10.1	7.68	0.27
2 nd _drive	-5.2	0.79	-5.5	0.78	-7.0	1.00	0.49
3 rd _drive	-10.2	0.63	-10.3	0.62	-11.5	0.87	0.37
Time_Baseline	-7.1	0.73	-7.3	0.72	-11.4	0.09	4.85
Time_Income	-7.9	0.70	-8.1	0.69	-12.7	0.08	4.76
Time_Outgo	-7.6	0.71	-7.8	0.70	-12.1	0.08	4.92
Age	+10.6	1.58	+12.6	1.74	—	—	—
DrR_Index	—	—	+114.0	171.40	—	—	—
Ddr_Index	—	—	-14.0	0.53	—	—	—

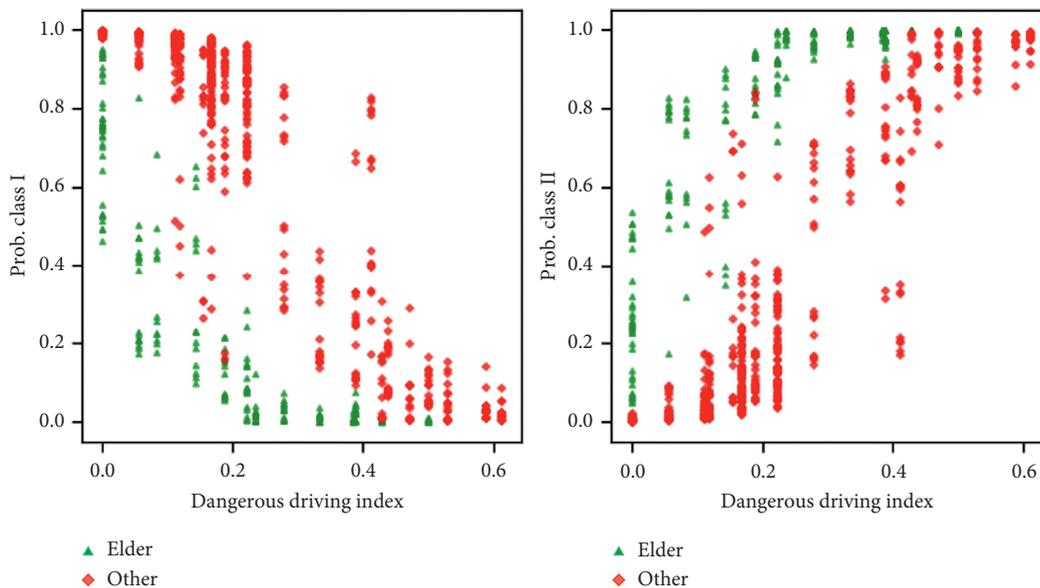


FIGURE 2: Effect of driving styles on the probability of Classes I and II.

phone calls on “high-risk” drivers in Class II changed under different circumstances. When they were close to the stop line with less than 2.95 seconds, both incoming and outgoing calls increased the likelihood probability of slowing and stopping, compared with uninterrupted driving. Although when the time to stop line was 3 seconds, incoming calls resulted in a greater decrease in the probability of going, compared to outgoing calls. While drivers were still at a distance from the stop line, with times longer than 3.6 seconds, both incoming and outgoing calls increase the probability of deciding to go, compared with uninterrupted driving.

It could be concluded that taking a phone call can initiate unsafe decisions in both Classes I and II. This could be in terms of both slowing unnecessarily and deciding to go when the time does not permit. Results show that Class II is more sensitive to this distraction having a greater change in the probability of deciding to go. This analysis suggests therefore that higher-risk drivers are more likely to respond

frivolously to distractions such as an incoming or outgoing phone call.

6.3. Results Comparison with Previous Research. The findings of this study add to the growing evidence base. For example, we find age and gender influence drivers’ stop/go decision-making, which agrees with previous studies [7, 45–47]. The results also suggest that most female drivers are more likely to run the yellow signal, compared with their male counterparts. Additionally, compared with adolescent and middle-aged drivers, old drivers (over 50 years old) are more likely to make unsafe decisions, which may be due to the long perception-reaction time.

The results agree with the research which suggests that drivers’ decisions are influenced by their familiarity with the traffic environment [7]. We further find that most drivers have no apparent inclination of stopping or going as the simulation experiment carries on, while a small portion of

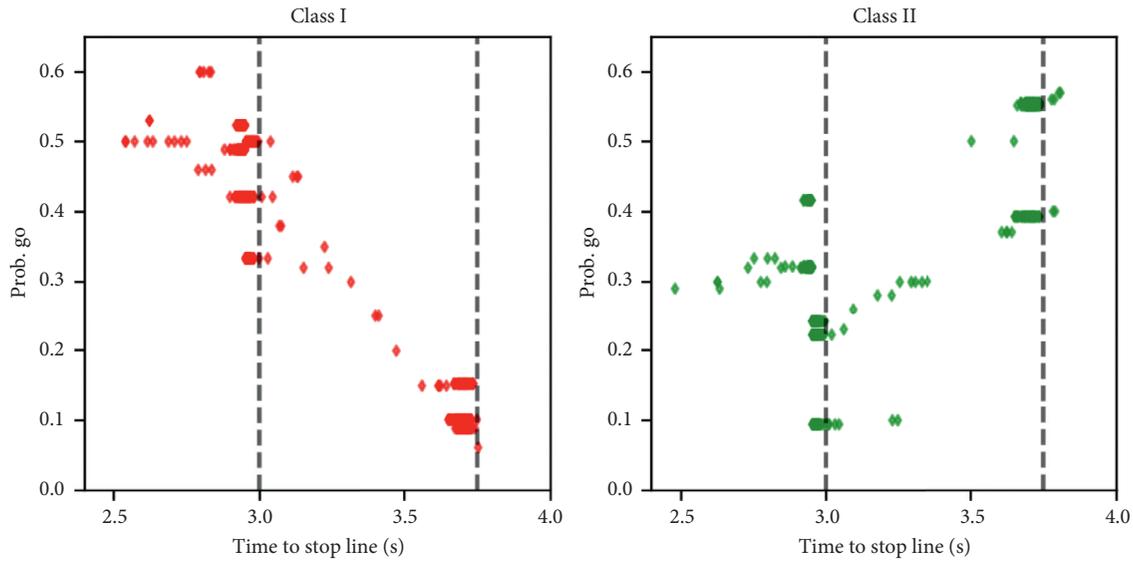


FIGURE 3: Decision properties for Classes I and II of drivers.

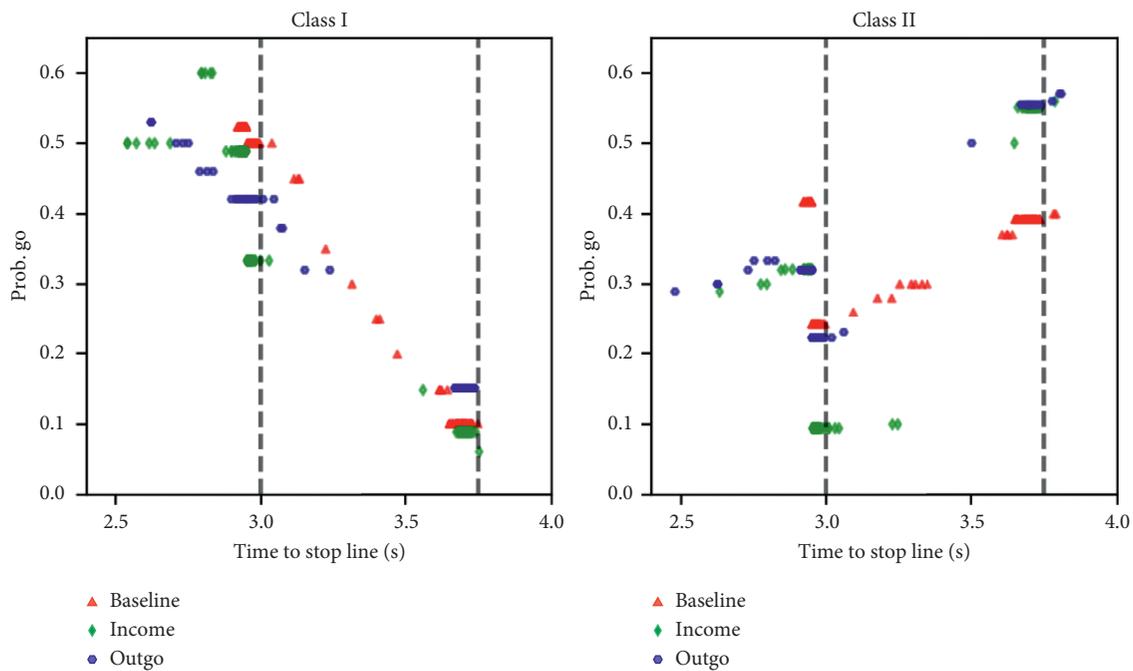


FIGURE 4: Effect of distraction on decisions in Classes I and II of drivers.

drivers tend to stop as the experiment proceeds. This may reflect that part drivers are more accustomed to the simulation environment and anticipate signal changes as the experiment commences.

The findings confirm that cell phone talk can induce drivers to make unsafe decisions [7, 10, 47]. Furthermore, we find drivers' behavioral differences in decision-making under distractions. Only close to the stop line are most drivers disturbed to make unsafe decisions. A small portion of drivers in distractions will make unsafe decisions

wherever they were, which may be due to their oldness and high-risk driving style.

7. Conclusions

Drivers' improper stop/go decisions at the onset of yellow signals may cause numerous problems at intersections. Improper decisions may lead to rear-end collisions or red-light running violations. To address the decision issues and accommodate group heterogeneity, a latent class logit model

was used to analyze drivers' decision-making processes in this study.

We explored two new variables associated with driving styles (i.e., the driving reliability index and dangerous driving index), which further calibrated the class probabilities in the latent class model. Indicators for the goodness of fit demonstrate that our model with driving styles is superior to binary logit models. Therefore, the latent class model considering driving styles can accurately evaluate drivers' decisions at the onset of yellow signal.

Drivers are classified into "low-risk" and "high-risk" categories according to driving styles. The driving reliability index appears to influence drivers' stop/go decision-making whereas the dangerous driving index influences the proportion of two categories. Results indicate that "low-risk" drivers are less likely to make risky decisions, while "high-risk" drivers are more likely to make improper decisions.

Similar with previous research, we found that driving while talking on the phone may cause drivers' inappropriate decisions at the onset of yellow signal. The effects of cell phone distractions were inconsistent within our sample of drivers. For "low-risk" drivers, it seemed that they were only slightly disturbed by the phone calls and more prone to stop when close to the stop line. However, there was little difference in the effect of incoming or outgoing calls. "High-risk" drivers presented obvious differences in the decision when using cell phones. This group appeared to behave completely differently to "low-risk" drivers meaning that our classification was meaningful.

Drivers in different groups have different preferences in stop/go decision-making. To improve safety at signalized intersections, policy-makers need to pay more attention to "high-risk" drivers, although critically speaking risk driving may be an impulse rather than a consistent factor. Nevertheless, "high-risk" drivers might be more intensively observed to ensure there are fewer risky behaviors. Imposing higher fines, reeducation, and higher charge rates of insurance may incentivize changes. Once the "high-risk" drivers develop stable and safe driving habits, the intensity of observation might be reduced.

Data Availability

The data used for this study are from experiments conducted at the University of Iowa-National Advanced Driving Simulator (NADS), which can be accessed from the following website: <http://depts.washington.edu/hfsm/upload.php>.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was supported by the Fundamental Research Funds for the Central Universities (no. 2019JBM036).

References

- [1] National Center for Statistics and Analysis, "Distracted driving in fatal crashes, 2017," Report No. DOT HS 812 700, National Highway Traffic Safety Administration, Washington, DC, USA, 2019.
- [2] The Federal Highway Administration (FHWA), *Intersection Safety: Background and Objectives*, The Federal Highway Administration (FHWA), Washington, DC, USA, 2018, <https://highways.dot.gov/research-programs/safety/intersection-safety>.
- [3] E.-H. Choi, "Crash factors in intersection-related crashes: An on-scene perspective (No. HS-811366)," National Center for Statistics and Analysis, National Highway Traffic Safety Administration, US Department of Transportation, Washington, DC, USA, 2010, <https://crashstats.nhtsa.dot.gov/Api/Public/ViewPublication/811366>.
- [4] D. Gazis, R. Herman, and A. Maradudin, "The problem of the amber signal light in traffic flow," *Operations Research*, vol. 8, no. 1, pp. 112–132, 1960.
- [5] C. Liu, R. Herman, and D. C. Gazis, "A review of the yellow interval dilemma," *Transportation Research Part A: Policy and Practice*, vol. 30, no. 5, pp. 333–348, 1996.
- [6] M. Elhenawy, A. Jahangiri, H. A. Rakha, and I. El-Shawarby, "Modeling driver stop/run behavior at the onset of a yellow indication considering driver run tendency and roadway surface conditions," *Accident Analysis and Prevention*, vol. 83, pp. 90–100, 2015.
- [7] P. T. Savolainen, "Examining driver behavior at the onset of yellow in a traffic simulator environment: comparisons between random parameters and latent class logit models," *Accident Analysis and Prevention*, vol. 96, pp. 300–307, 2016.
- [8] K. L. Young, P. M. Salmon, and M. Cornelissen, "Distraction-induced driving error: an on-road examination of the errors made by distracted and undistracted drivers," *Accident Analysis and Prevention*, vol. 58, pp. 218–225, 2013.
- [9] S. G. Klauer, F. Guo, B. G. Simons-Morton, M. C. Ouimet, S. E. Lee, and T. A. Dingus, "Distracted driving and risk of road crashes among novice and experienced drivers," *The New England Journal of Medicine*, vol. 370, no. 1, pp. 54–59, 2014.
- [10] A. D. Ohlhauser, L. N. Boyle, D. Marshall, and O. Ahmad, "Drivers' behavior through a yellow light: effects of distraction and age," *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, vol. 55, no. 1, pp. 1919–1923, 2011.
- [11] B. E. Porter and K. J. England, "Predicting red-light running behavior: a traffic safety study in three urban settings," *Journal of Safety Research*, vol. 31, no. 1, pp. 1–8, 2000.
- [12] H. Köll, M. Bader, and K. W. Axhausen, "Driver behaviour during flashing green before amber: a comparative study," *Accident Analysis and Prevention*, vol. 36, no. 2, pp. 273–280, 2004.
- [13] Y. Liu, G.-L. Chang, R. Tao, T. Hicks, and E. Tabacek, "Empirical observations of dynamic dilemma zones at signalized intersections," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2035, no. 1, pp. 122–133, 2007.
- [14] N. Elmitiny, X. Yan, E. Radwan, C. Russo, and D. Nashar, "Classification analysis of driver's stop/go decision and red-light running violation," *Accident Analysis and Prevention*, vol. 42, no. 1, pp. 101–111, 2010.
- [15] Y. Wu, M. Abdel-Aty, Y. Ding, B. Jia, Q. Shi, and X. Yan, "Comparison of proposed countermeasures for dilemma zone at signalized intersections based on cellular automata simulations," *Accident Analysis and Prevention*, vol. 116, pp. 69–78, 2018.

- [16] Urbanik T., Koonce P., The dilemma with dilemma zones, 2007, http://redlightrobber.com/red/links_pdf/The-Dilemma-with-Dilemma-Zones.pdf.
- [17] C. V. Zegeer and R. C. Deen, "Green-extension systems at high-speed intersections," *Ite Journal-Institute of Transportation Engineers*, vol. 48, no. 11, pp. 19–24, 1978.
- [18] L. A. Klein, M. K. Mills, and D. R. P. Gibson, *Traffic Detector Handbook*, US. Department of Transportation, Federal Highway Administration, Washington, DC, USA, 3rd edition, 2006.
- [19] N. K. Chaudhary, T. D. Casanovapowell, L. Cosgrove, I. Reagan, and A. Williams, "Evaluation of NHTSA distracted driving demonstration projects in connecticut and New York," 2014, http://www.nhtsa.gov/staticfiles/nti/pdf/811635_Eval_NHTSA_Distracted_Driving_Demo_Proj_Comm_CT_and_NY.pdf.
- [20] B. J. Russo, J. J. Kay, P. T. Savolainen, and T. J. Gates, "Assessing characteristics related to the use of seatbelts and cell phones by drivers: application of a bivariate probit model," *Journal of Safety Research*, vol. 49, pp. 137.e1–142, 2014.
- [21] I. El-Shawarby, H. A. Rakha, V. Inman, and G. Davis, "Effect of yellow-phase trigger on driver behavior at high-speed signalized intersections," in *Proceedings of the IEEE Intelligent Transportation Systems Conference*, pp. 683–688, Toronto, Canada, September 2006.
- [22] H. Rakha, I. El-Shawarby, and J. R. Setti, "Characterizing driver behavior on signalized intersection approaches at the onset of a yellow-phase trigger," *IEEE Transactions on Intelligent Transportation Systems*, vol. 8, no. 4, pp. 630–640, 2007.
- [23] P. Papaioannou, "Driver behaviour, dilemma zone and safety effects at urban signalized intersections in Greece," *Accident Analysis and Prevention*, vol. 39, no. 1, pp. 147–158, 2007.
- [24] T. J. Gates, D. A. Noyce, L. Laracuent, and E. V. Nordheim, "Analysis of driver behavior in dilemma zones at signalized intersections," *Transportation Research Record*, vol. 2030, no. 1, pp. 29–39, 2007.
- [25] X. Yan, E. Radwan, D. Guo, and S. Richards, "Impact of "signal ahead" pavement marking on driver behavior at signalized intersections," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 12, no. 1, pp. 50–67, 2009.
- [26] K. Long, Y. Liu, and L. D. Han, "Impact of countdown timer on driving maneuvers after the yellow onset at signalized intersections: an empirical study in Changsha, China," *Safety Science*, vol. 54, pp. 8–16, 2013.
- [27] C. R. Bhat, "An endogenous segmentation mode choice model with an application to intercity travel," *Transportation Science*, vol. 31, no. 1, pp. 34–48, 1997.
- [28] K. Roeder, K. G. Lynch, and D. S. Nagin, "Modeling uncertainty in latent class membership: a case study in criminology," *Journal of the American Statistical Association*, vol. 94, no. 447, pp. 766–776, 1999.
- [29] J. S. Uebersax, "Probit latent class analysis with dichotomous or ordered category measures: conditional independence/dependence models," *Applied Psychological Measurement*, vol. 23, no. 4, pp. 283–297, 1999.
- [30] W. H. Greene and D. A. Hensher, "A latent class model for discrete choice analysis: contrasts with mixed logit," *Transportation Research Part B: Methodological*, vol. 37, no. 8, pp. 681–698, 2003.
- [31] J. W. Milon and D. Scrogin, "Latent preferences and valuation of wetland ecosystem restoration," *Ecological Economics*, vol. 56, no. 2, pp. 162–175, 2006.
- [32] J. L. Walker and J. Li, "Latent lifestyle preferences and household location decisions," *Journal of Geographical Systems*, vol. 9, no. 1, pp. 77–101, 2007.
- [33] A. R. Hole, "Modelling heterogeneity in patients' preferences for the attributes of a general practitioner appointment," *Journal of Health Economics*, vol. 27, no. 4, pp. 1078–1094, 2008.
- [34] S. Hess, A. Stathopoulos, and A. Daly, "Allowing for heterogeneous decision rules in discrete choice models: an approach and four case studies," *Transportation*, vol. 39, no. 3, pp. 565–591, 2012.
- [35] J. Shen, "Latent class model or mixed logit model? A comparison by transport mode choice data," *Applied Economics*, vol. 41, no. 22, pp. 2915–2924, 2009.
- [36] TRB Statistics Committee (ABJ80), "Questions and answers for 2014 TRB contest," 2013, http://trbstats.weebly.com/uploads/6/6/6/4/6664035/q_a_for_2014_trb_contest_nov_18_2013.pdf.
- [37] L. E. Baum and T. Petrie, "Statistical inference for probabilistic functions of finite state Markov chains," *The Annals of Mathematical Statistics*, vol. 37, no. 6, pp. 1554–1563, 1966.
- [38] L. R. Rabiner, "A tutorial on hidden markov models and selected applications in speech recognition," *Proceedings of the IEEE*, vol. 77, no. 2, pp. 267–296, 1989.
- [39] E. Farmer and E. G. Chambers, "A study of accident proneness among motor drivers," *Journal of the Royal Statistical Society*, vol. 103, no. 2, p. 254, 1940.
- [40] Greene W. H., NLOGIT 5, Version 5.0, 2012, <http://www.limdep.com>.
- [41] A. Sharma, "Integrated behavioral and economic framework for improving dilemma zone protection systems," Purdue University, West Lafayette, IN, USA, Doctoral Dissertation, 2008.
- [42] A. Amer, H. Rakha, and I. El-Shawarby, "Agent-based behavioral modeling framework of driver behavior at the onset of yellow indication at signalized intersections," in *Proceedings of the 2011 14th International IEEE Conference on Intelligent Transportation Systems*, pp. 1809–1814, Washington, DC, USA, October 2011.
- [43] A. Sharma, D. Bullock, and S. Peeta, "Estimating dilemma zone hazard function at high speed isolated intersection," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 3, pp. 400–412, 2011.
- [44] J. Li, Q. He, H. Zhou, Y. Guan, and D. Wei, "Modeling driver behavior near intersections in hidden markov model," *International Journal of Environmental Research and Public Health*, vol. 13, no. 12, p. 1265, 2016.
- [45] S. M. Lavrenz, V. D. Pyrialakou, and K. Gkritza, "Modeling driver behavior in dilemma zones: a discrete/continuous formulation with selectivity bias corrections," *Analytic Methods in Accident Research*, vol. 3–4, pp. 44–55, 2014.
- [46] H. Xiong, P. Narayanaswamy, S. Bao, C. Flannagan, and J. Sayer, "How do drivers behave during indecision zone maneuvers?" *Accident Analysis & Prevention*, vol. 96, pp. 274–279, 2016.
- [47] M. M. Haque, A. D. Ohlhauser, S. Washington, and L. N. Boyle, "Decisions and actions of distracted drivers at the onset of yellow lights," *Accident Analysis & Prevention*, vol. 96, pp. 290–299, 2016.