

Investigating the factors affecting traffic violations based on electronic enforcement data: A case study in Shangyu, China

Fan Haoxuan Ren Gang Li Haojie Ma Jingfeng

(School of Transportation, Southeast University, Nanjing 211189, China)

Abstract: To study the influencing factors of traffic violations, this study investigated the effects of vehicle attribution, day of week, time of day, location of traffic violations, and weather on traffic violations based on the electronic enforcement data and historical weather data obtained in Shangyu, China. Ten categories of traffic violations were determined from the raw data. Then, chi-square tests were used to analyze the relationship between traffic violations and the potential risk factors. Multinomial logistic regression analyses were conducted to further estimate the effects of different risk factors on the likelihood of the occurrence of traffic violations. By analyzing the results of chi-square tests via SPSS, the five factors above were all determined as significant factors associated with traffic violations. The results of the multinomial logistic regression revealed the significant effects of the five factors on the likelihood of the occurrence of corresponding traffic violations. The conclusions are of great significance for the development of effective traffic intervention measures to reduce traffic violations and the improvement of road traffic safety.

Key words: traffic violations; road traffic safety; electronic enforcement data; multinomial logistic regression; influencing factors

DOI: 10.3969/j.issn.1003-7985.2021.02.014

The number of annual road traffic deaths has reached 1.35 million worldwide. In fact, road traffic-related injuries are now the leading killer of people aged 5-29 years and are the 8th leading cause of death for people of all ages^[1]. In China, a total of 244 937 traffic accidents occurred in 2018, resulting in 63 194 deaths and a direct property loss of 1 384.559 million yuan^[2]. The traffic violation of motor vehicle drivers is one of the main causes of traffic accidents, accounting for 83.2%, 83.6%, and 85.2% of all accidents in 1996, 1997, and 1999 in China, respectively. By 2010, the traffic violation of motor vehicle drivers has accounted for 91.08% of all ac-

cidents^[3]. For this reason, traffic violation, as one of the major risks influencing traffic accidents^[4], is worthy of further study. Zhang et al.^[5] conducted research using the data of traffic accidents and traffic violations. They found that traffic violation was one of the major risks threatening road safety and that obeying traffic rules can effectively reduce the frequency of traffic accidents, consistent with the results of a Canadian study^[6]. Therefore, studying the characteristics and influencing factors of traffic violations is of great necessity, as it can help establish a theoretical foundation for the promotion of effective traffic intervention measures and improvement of road traffic safety.

Many studies have analyzed the factors affecting traffic violations over the years. For example, younger drivers were found to be more likely to be involved in traffic violations^[7]. Oppenheim et al.^[8] revealed the effect of gender on drivers' violation tendencies. One study identified male drivers, private vehicles, the lack of street lighting at night, and poor visibility as significant risk factors associated with speeding and drunk driving^[9]. Akaateba et al.^[10] investigated the relationship between drivers' educational attainment and driving experience and traffic violations. License plate, season, speeding area, position, and rainfall were also identified as factors affecting speeding violation^[11]. Martínez-Ruíz et al.^[12] evaluated the impact of camera enforcement for traffic violations, while Zhang et al.^[13] considered the factors of weather conditions in their research. Wang et al.^[14] concluded that the common traffic violations of bus drivers could be attributed to the date, weather, and presence of traffic cameras at bus stations. Precht et al.^[15] found that drivers are more likely to commit traffic violations on rainy days. Fu et al.^[16] selected vehicle factors, temporal factors, weather, and road and traffic management conditions as potential risk factors of wrong-way driving, red-light-running, violating traffic markings, and driving in the inaccurate oriented lane at signalized intersections.

Earlier studies on traffic violations mainly relied on data collected from questionnaires^[8,10,17-19]. Then, subsequent studies used structural equation modeling^[19-20], chi-square tests^[18], and regression techniques^[17] to analyze the self-reported traffic violations based on the data collected from the questionnaires. However, this method of data collection has a limitation, and the inaccuracy of

Received 2020-11-14, **Revised** 2021-04-07.

Biographies: Fan Haoxuan (1997—), male, master; Ren Gang (corresponding author), male, doctor, professor, rengang@seu.edu.cn.

Foundation item: The National Key Research and Development Program of China (No. 2019YFB1600200).

Citation: Fan Haoxuan, Ren Gang, Li Haojie, et al. Investigating the factors affecting traffic violations based on electronic enforcement data: A case study in Shangyu, China [J]. Journal of Southeast University (English Edition), 2021, 37(2): 227 – 236. DOI: 10.3969/j.issn.1003-7985.2021.02.014.

respondents’ self-reported behaviors may affect the analysis results to some extent^[19,21].

There are also some other data sources of traffic violations and methods of data analysis in existing studies. For example, logistic regression was utilized to analyze the significant factors influencing speeding and drunk driving based on data extracted from the traffic management sector-specific incident case data report^[9]. Some traffic violations cause incidents, whereas some do not. Therefore, it is not accurate to study the influencing factors of traffic violations based on incident case data alone. Andreuccetti et al.^[22] conducted a cross-sectional observational study by collecting data on drunk driving and speeding. Cheng et al.^[11] adopted the violations data collected by the electronic enforcement equipment and analyzed them using binary logistic regression. Tavakoli Kashani et al.^[23] analyzed crash data in Iran from 2012—2016. They used the classification regression tree and the quasi-induced exposure method to identify the important driver and vehicle features associated with the red-light running frequencies.

Based on the above literature review, the main research contents of this article are as follows:

1) Ten categories of traffic violations are extracted from the raw data instead of analyzing only one or two kinds of typical traffic violations.

2) The risk factors of traffic violations are determined based on the information included in the data and existing studies, including vehicle attribution, day of the week, time of day, location of violations, and weather.

3) Then, the chi-square tests and multinomial logistic regression method were used to analyze the relationships between traffic violations and the identified risk factors as well as the effects of different factors on the likelihood of the occurrence of traffic violations.

4) Finally, corresponding traffic intervention measures are proposed according to the analysis results.

1 Materials and Methods

1.1 Data sources

Studies on the influencing factors of traffic violations suffer from the limitation of inaccurate self-reported violations and incident case data. Furthermore, the observational study of traffic violations is time-consuming and laborious. However, due to restrictions on the acquisition of relevant data, there are still infrequent empirical studies on the factors affecting traffic violations. The current research is conducted from the perspective of traffic violations in practical driving situations rather than self-reported violations and uses electronic enforcement data of traffic violations collected in Shangyu, Zhejiang Province, China. Furthermore, historical weather data of Shangyu were extracted from the website <http://lishi.tianqi.com/shangyu.html>.

The electronic enforcement data of traffic violations

used in this paper were obtained from Shangyu, a city in China with a population of approximately 780 000 and an area of 1 401.68 km². The data were extracted from the electronic police system for the period of January 1—31, 2019. The data consisted of 57 441 records of traffic violations. After eliminating 1 444 invalid records without traffic violation codes, we obtained 55 997 valid records, including vehicle license plate numbers, times and locations of traffic violations, and traffic violation codes. The specific descriptions of traffic violations can be obtained according to the traffic violation codes.

There were 23 kinds of violations recorded in the data, of which 14 were classified as low-frequency violations whose frequencies were lower than 2 000 times a month. From these, ten types of violations were determined (see Tab. 1).

Tab. 1 Traffic violation statistics

Traffic violations	Description of traffic violations	Frequency
V1	Use of an improper oriented lane	9 775 (17.46%)
V2	Illegal parking	9 670 (17.27%)
V3	Violating traffic signals	8 425 (15.05%)
V4	Violating traffic markings	7 312 (13.06%)
V5	Driving in the reverse direction	5 292 (9.45%)
V6	Speeding by 20% but less than 50%	3 702 (6.61%)
V7	Spraying and pasting materials on the vehicle number plate	3 663 (6.54%)
V8	Failure to yield to pedestrians	3 345 (5.97%)
V9	Driving without a driving license	2 021 (3.61%)
V10	Low-frequency traffic violations	2 792 (4.99%)

Note: The data in parentheses indicates the percentage of various traffic violations to the total number of traffic violations.

1.2 Risk factors

Risk factors were determined based on the information included in the electronic enforcement data of traffic violations and the weather data. Vehicle attribution, day of week, time of day, location of violations, and weather were selected as potential influencing factors based on the information included in the data and existing studies^[9,13-16,24-25]. Although the driver’s personal characteristics may be closely related to traffic violations, and many researchers have made useful explorations into them, they are not considered as the influencing factors, because they are not recorded in the data source. Many studies have been carried out on the effect of vehicle characteristics on traffic violations. However, they ignored the information included in vehicle license plate numbers. We can actually obtain vehicle attribution from vehicle license plate numbers to identify whether a vehicle is local or non-local. Meanwhile, existing studies on weather-related factors mainly studied their impacts on traffic accidents or focused only on one or two kinds of typical traffic violations. In comparison, the current work studied their impacts on ten categories of traffic violations and conducted

a more comprehensive analysis.

As shown in Tab. 1, traffic violations were classified into ten categories. Vehicle attribution can be obtained from vehicle license plate numbers. Specifically, vehicles were divided into local and non-local vehicles according to vehicle attribution. The variable of day of week contained two categories: weekday (Monday, Tuesday, Wednesday, Thursday, and Friday) and weekend (Sat-

urday and Sunday). Then, the data were processed according to time, as shown in Fig. 1. The values in Fig. 1 are average values. The variable of time of day was divided into six categories: 00:00–06:59 (midnight to dawn), 07:00–08:59 (peak hours in the morning), 09:00–11:59 (morning working hours), 12:00–16:59 (afternoon working hours), 17:00–18:59 (peak hours in the evening), and 19:00–23:59 (night time).

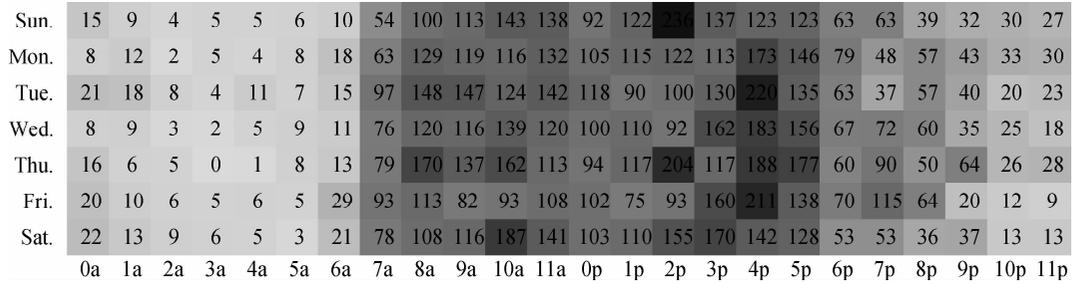


Fig. 1 Frequencies of traffic violations by the time

The location of traffic violations was used to identify whether a driver is committing a traffic violation on the road section or at the intersection. The weather variable was divided into five categories according to the weather records of January 2019 in Shangyu: sunny, cloudy, light rain, moderate rain, and snowy. Combined with the information above, the research variables were determined (see Tab. 2). Here, the dependent variable was traffic violations, and the remaining ones were the independent variables.

Tab. 2 Value descriptions of variables

Variables	Description
Vehicle attribution	Local vehicles = 1; otherwise = 0
Day of week	Weekday = 1; otherwise = 0
Time of day	00:00–06:59 = 0; 07:00–08:59 = 1; 09:00–11:59 = 2; 12:00–16:59 = 3; 17:00–18:59 = 4; 19:00–23:59 = 5
Location of traffic violations	Road section = 1; otherwise = 0
Weather	Sunny = 0; cloudy = 1; light rain = 2; moderate rain = 3; snowy = 4
Traffic violations	V1 = 1; V2 = 2; V3 = 3; V4 = 4; V5 = 5; V6 = 6; V7 = 7; V8 = 8; V9 = 9; V10 = 10

1.3 Chi-square test and multinomial logistic regression model

A chi-square test was conducted to analyze the relationship between traffic violations and the potential risk factors. The factors that were statistically significant at the 95% confidence levels were included in the final multinomial model fitted.

The categories in the dependent variable were discrete, nominal, and unordered. Multinomial logistic regression is an effective method in dealing with multi-category de-

pendent variables^[26]. Although multinomial logistic regression is not the only estimation methodology available for situations with unordered categorical dependent variables, it has two advantages: the estimates can be quickly calculated, and the model results are easily interpreted^[27]. Thus, multinomial logistic regression was utilized in the current study to analyze the varying influence of each factor on the identified traffic violations.

When the dependent variable has k categories, logistic regression analysis is conducted in $k - 1$ categories, and the k -th category is defined as the reference category. Then the multinomial logistic regression model corresponding to the j -th traffic violations can be formulated as

$$\ln\left(\frac{p_j}{1 - p_k}\right) = \left(\beta_{j0} + \sum_{i=1}^n \beta_{ji}x_{ji}\right)$$

where p_j and p_k are the probabilities of the occurrence of the j -th and k -th kind of categories, respectively; β_{j0} is constant; β_{ji} denotes the regression coefficient of the i -th independent variable x_{ji} in the multinomial logistic regression model of the j -th category; and n represents the number of independent variables^[28].

2 Results

2.1 Descriptive statistics

After determining the research variables, the descriptive statistics of possible risk factors are summarized in Tab. 3. The percentages of traffic violations by each category within a factor are shown in Tab. 3. The percentages are the proportions of various types of traffic violations in each category within the risk factors. The percentage accumulation distributions of traffic violations under each category are shown in Fig. 2.

Tab.3 Traffic violation percentages in various influencing factors

Variables	Percentages of traffic violations									
	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10
Local vehicles	18.40	17.64	15.04	12.38	9.65	6.43	7.04	6.44	3.69	3.30
Non-local vehicles	14.01	15.90	15.07	15.54	8.72	7.27	4.73	4.27	3.32	11.17
Weekday	18.47	16.85	15.03	13.00	9.77	6.34	6.51	5.82	3.11	5.09
Weekend	14.51	18.49	15.10	13.23	8.52	7.39	6.62	6.42	5.05	4.67
00:00—06:59	11.84	7.01	30.43	8.59	19.38	13.81	0.00	0.26	0.04	8.64
07:00—08:59	23.61	11.90	12.45	14.50	7.42	4.35	4.93	7.99	4.09	8.75
09:00—11:59	10.37	15.84	14.05	14.63	8.14	7.71	13.15	7.38	4.54	4.20
12:00—16:59	17.18	16.80	13.20	14.88	8.38	6.16	7.10	8.47	3.49	4.34
17:00—18:59	32.82	24.50	14.53	8.25	7.86	3.38	1.54	1.32	1.59	4.21
19:00—23:59	12.20	22.80	20.57	9.23	15.46	8.94	1.28	0.03	5.03	4.45
Road section	12.21	26.83	0.00	0.24	9.91	11.78	14.64	11.21	8.08	5.10
Intersection	21.69	9.54	27.20	23.42	9.08	2.543	0.00	1.74	0.00	4.89
Sunny	17.85	14.49	16.06	12.87	9.35	5.73	8.34	6.72	3.46	5.13
Cloudy	17.26	17.36	15.30	12.62	8.49	7.35	6.13	6.42	4.22	4.87
Light rain	17.89	18.39	14.02	13.56	11.35	5.60	6.89	4.78	2.64	4.90
Moderate rain	15.92	18.23	15.18	17.44	11.45	5.55	1.77	5.55	3.14	5.75
Snowy	16.59	19.58	14.44	11.51	6.36	9.61	6.30	5.88	3.61	6.12

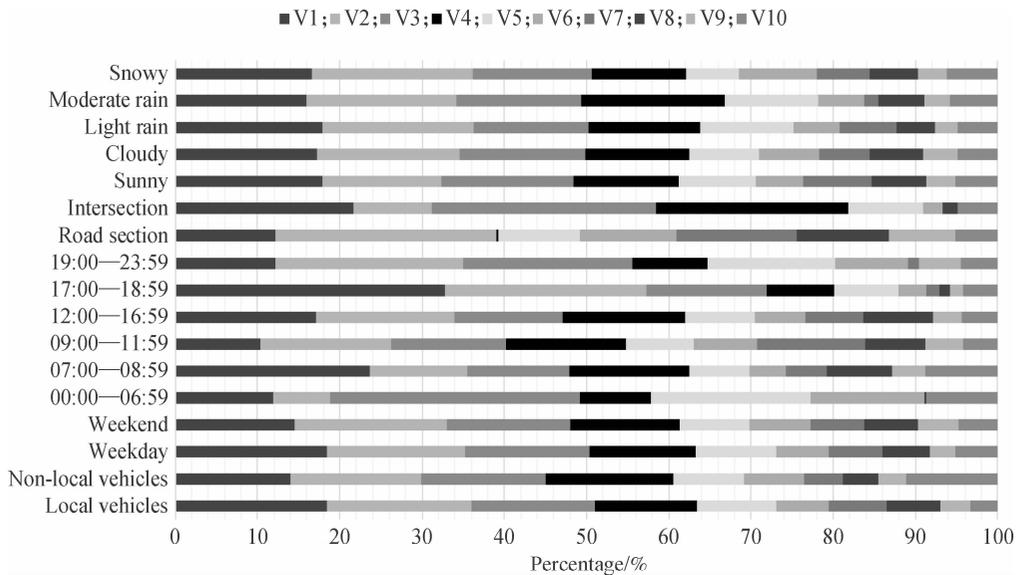


Fig.2 Percentage accumulation distributions of traffic violations

From Tab.3 and Fig.2, driving in the reverse direction and speeding by 20% but less than 50% account for a larger proportion at 00:00—06:59 (V5 19.38%, V6 13.81%) than at 07:00—08:59 (V5 7.42%, V6 4.35%), 09:00—11:59 (V5 8.14%, V6 7.71%), 12:00—16:59 (V5 8.38%, V6 6.16%), 17:00—18:59 (V5 7.86%, V6 3.38%), and 19:00—23:59 (V5 15.46%, V6 8.94%). Furthermore, violating traffic signals (V3), as the main traffic violation at 00:00—06:59, accounts for more than 30%.

Use of an improper oriented lane (V1 21.69%), violating traffic signals (V3 27.20%) and violating traffic markings (V4 23.42%) are the main violations at the intersection, while illegal parking (V2 26.83%) accounts for the largest proportion on the road section.

2.2 Factors affecting traffic violations

The chi-square test was conducted at the 5% level of significance. Tab.4 shows the results of the chi-square test. The chi-square statistic (χ^2) of each variable is large and $p < 0.001$, indicating that vehicle attribution, day of week, time of day, location of traffic violations, and weather are statistically significant factors associated

Tab.4 Chi-square test results

Independent variables	χ^2	df
Vehicle attribution	1 541.036 ***	9
Day of week	269.112 ***	9
Time of day	7 177.538 ***	45
Location of traffic violations	26 564.460 ***	9
Weather	575.609 ***	36

Note: *** $p < 0.001$.

with traffic violations.

To further estimate the effects of different influencing factors on the likelihood of the occurrence of various traffic violations, logistic regression analyses were conducted. The values of each variable are defined in Tab. 2. The data collected are imported into SPSS. Traffic violations was set as the dependent variable, and the other variables were chosen as factors in the model. According to the

principle of multinomial logistic regression, a category was first selected as the reference category. Low-frequency traffic violations (V10) was chosen as the reference category in the multinomial logistic regression model of traffic violations. The reference categories of each factor were local vehicles, weekday, 19:00–23:59, road section, and snowy. The main output results are shown in Tab. 5.

Tab. 5 Parameter estimates of the multinomial logistic regression model

Traffic violations ^a	Factors	OR	Traffic violations ^a	Factors	OR	Traffic violations ^a	Factors	OR
Use of an improper oriented lane (V1)	Non-local vehicles	0.216 ***	Illegal parking (V2)	Non-local vehicles	0.284 ***	Violating traffic signals (V3)	Non-local vehicles	0.244 ***
	Weekend	0.806 ***		Weekend	1.286 ***		Weekend	0.965
	00:00–06:59	0.515 ***		00:00–06:59	0.188 ***		00:00–06:59	0.726 *
	07:00–08:59	1.037		07:00–08:59	0.197 ***		07:00–08:59	0.415 ***
	09:00–11:59	0.928		09:00–11:59	0.672 ***		09:00–11:59	0.789 *
	12:00–16:59	1.537 ***		12:00–16:59	0.620 ***		12:00–16:59	0.818 *
	17:00–18:59	3.156 ***		17:00–18:59	0.871		17:00–18:59	1.067
	Intersection	2.305 ***		Intersection	0.387 ***		Intersection	7 638.088 ***
	Sunny	1.361 *		Sunny	0.816		Sunny	1.386 *
	Cloudy	1.410 *		Cloudy	1.028		Cloudy	1.312 *
Light rain	1.395 *	Light rain	1.158	Light rain	1.198			
Moderate rain	0.968	Moderate rain	0.990	Moderate rain	1.021			
Violating traffic markings (V4)	Non-local vehicles	0.320 ***	Driving in the reverse direction (V5)	Non-local vehicles	0.255 ***	Speeding by 20% but less than 50% (V6)	Non-local vehicles	0.357 ***
	Weekend	1.046		Weekend	1.001		Weekend	1.246 ***
	00:00–06:59	0.450 ***		00:00–06:59	0.706 ***		00:00–06:59	1.001
	07:00–08:59	1.099		07:00–08:59	0.217 ***		07:00–08:59	0.166 ***
	09:00–11:59	1.833 ***		09:00–11:59	0.548 ***		09:00–11:59	0.799 *
	12:00–16:59	2.063 ***		12:00–16:59	0.527 ***		12:00–16:59	0.530 ***
	17:00–18:59	1.356 *		17:00–18:59	0.496 ***		17:00–18:59	0.275 ***
	Intersection	116.501 ***		Intersection	0.937		Intersection	0.191 ***
	Sunny	1.375 *		Sunny	1.733 ***		Sunny	0.655 *
	Cloudy	1.330 *		Cloudy	1.676 ***		Cloudy	0.891
Light rain	1.486 *	Light rain	2.219 ***	Light rain	0.718 *			
Moderate rain	1.474 *	Moderate rain	1.893 ***	Moderate rain	0.637 *			
Spraying and pasting materials on the vehicle number plate (V7)	Non-local vehicles	0.239 ***	Failure to yield to pedestrians (V8)	Non-local vehicles	0.225 ***	Driving without driving license (V9)	Non-local vehicles	0.318 ***
	Weekend	1.156 *		Weekend	1.176 *		Weekend	1.768 ***
	00:00–06:59			00:00–06:59	5.930 *		00:00–06:59	0.006 ***
	07:00–08:59	1.115		07:00–08:59	90.361 ***		07:00–08:59	0.232 ***
	09:00–11:59	9.280 ***		09:00–11:59	232.173 ***		09:00–11:59	0.764 *
	12:00–16:59	3.860 ***		12:00–16:59	217.717 ***		12:00–16:59	0.460 ***
	17:00–18:59	0.744		17:00–18:59	31.724 ***		17:00–18:59	0.194 ***
	Intersection			Intersection	0.186 ***		Intersection	
	Sunny	1.560 *		Sunny	1.353		Sunny	0.992
	Cloudy	1.155		Cloudy	1.312		Cloudy	1.238
Light rain	1.381 *	Light rain	1.046	Light rain	0.900			
Moderate rain	0.302 ***	Moderate rain	1.022	Moderate rain	1.000			

Note: ^a The reference category is V10. * $p < 0.05$, *** $p < 0.001$.

Odds ratio (OR) denotes the adjusted odds ratio for each factor. By analyzing OR, the effects of different influencing factors on the likelihood of the occurrence of various violations can be estimated.

Non-local vehicles ($p < 0.001$, $OR < 1$) exhibit a lower risk of all traffic violations compared to local vehicles, and V1 is less likely to occur on weekends ($p < 0.001$, $OR = 0.806$) than on weekdays. An increased probability

of V1 is related to 17:00–18:59 ($p < 0.001$, $OR = 3.156$) and intersection ($p < 0.001$, $OR = 2.305$). V1 behaviour is less likely to occur in snowy compared to sunny ($p < 0.05$, $OR = 1.361$), cloudy ($p < 0.05$, $OR = 1.410$), and light rain ($p < 0.05$, $OR = 1.395$) weather conditions.

The probability of V2 occurring at the weekend ($p < 0.001$, $OR = 1.286$) is 1.286 times that in the weekday,

and lower probabilities of V2 are exhibited at 00:00—06:59 ($p < 0.001$, OR = 0.188), 07:00—08:59 ($p < 0.001$, OR = 0.197), 09:00—11:59 ($p < 0.001$, OR = 0.672), and 12:00—16:59 ($p < 0.001$, OR = 0.620) than at 19:00—23:59. In addition, V2 is less likely to occur at the intersection ($p < 0.001$, OR = 0.387) than at the road section.

Vehicle attribution, time of day, and location of traffic violations are found to be significantly associated with V3 and V4. The effect of vehicle attribution has been discussed above. The intersection exhibits a particularly high risk of the occurrence of V3 and V4 compared to road sections, and lower probabilities of V3 and V5 are exhibited at 00:00—16:59 than at 19:00—23:59. The periods of 09:00—11:59 ($p < 0.001$, OR = 1.833), 12:00—16:59 ($p < 0.001$, OR = 2.063), and 17:00—18:59 ($p < 0.05$, OR = 1.356) contribute significantly to the occurrence of V4. In addition, V3, V4, and V5 are less likely to occur in snowy weather.

Among the factors of day of week, results show that higher risks of V6 ($p < 0.001$, OR = 1.246), V7 ($p < 0.05$, OR = 1.156), V8 ($p < 0.05$, OR = 1.176), and V9 ($p < 0.001$, OR = 1.768) are more likely to be associated with weekends. In addition, an increased risk of V6 is associated with snowy weather. Sunny ($p < 0.05$, OR = 1.560) and light rain ($p < 0.05$, OR = 1.381) contribute significantly to the occurrence of V7.

V6 is less likely to occur at 07:00—08:59 ($p < 0.001$, OR = 0.166), 09:00—11:59 ($p < 0.05$, OR = 0.799), 12:00—16:59 ($p < 0.001$, OR = 0.530), and 17:00—18:59 ($p < 0.001$, OR = 0.275) than at 19:00—23:59. Conversely, V7 is more likely to occur at 09:00—11:59 ($p < 0.001$, OR = 9.280) and 12:00—16:59 ($p < 0.001$, OR = 3.860). In addition, the higher probabilities of V8 are exhibited at 00:00—06:59 ($p < 0.05$, OR = 5.930), 07:00—08:59 ($p < 0.001$, OR = 90.361), 09:00—11:59 ($p < 0.001$, OR = 232.173), 12:00—16:59 ($p < 0.001$, OR = 217.717), and 17:00—18:59 ($p < 0.001$, OR = 31.724) than at 19:00—23:59, which is contrary to the probability of V9.

Considering the location of traffic violations, the intersection is found to have lower risks of V6 ($p < 0.001$, OR = 0.191) and V8 ($p < 0.001$, OR = 0.186).

3 Discussions

3.1 Effects of vehicle attribution

Consistent with the conclusions of previous studies^[16,29], local vehicles are more likely to become involved in various traffic violations. In particular, compared to non-local vehicles, local vehicles are more likely to commit red-light-running^[16,29] and driving in inaccurately oriented lanes^[16]. This may be attributed to the fact

that non-local drivers are not familiar with the local traffic environment and are more cautious when driving, resulting in fewer traffic violations, while local drivers may appear more casual while driving in familiar conditions. Therefore, it is necessary to carry out effective safety publicity and training for drivers of local vehicles to improve their attention to traffic safety.

3.2 Effects of the day of week

It is noteworthy that illegal parking is more likely to occur on the weekend. However, Wang et al.^[14] concluded that illegal stopping was less common on weekends than on weekdays. This may be attributed to our different research objects. This paper considered all vehicles, while Wang et al.^[14] only investigated buses. As of December 2015, the number of motor vehicles in Shangyu had reached 193 000. The contradiction between the situation of the high-speed growth of motorization and resources is becoming increasingly serious. Most of the traveling done on weekdays is commuting, for which some motorists may choose public or shared transportation. In comparison, the travel purpose of travelers on the weekend is more for leisure, so cars are chosen by urban residents because of the advantages of high speed, high comfort, and flexibility. Therefore, the demand for parking is likely to increase on weekends. Furthermore, buildings at worksites are often equipped with complete parking facilities, but on weekends more travelers go to public places, thus highlighting the problem of large gaps in public parking spaces, which lead to more instances of illegal parking. Koohpayma et al.^[30] also concluded that the number of vehicle parking violations is directly related to points of interest. Therefore, it is necessary to diversify investments in the construction of public parking lots and standardize the parking space allocation indicators for various types of land.

Unlike the results of a previous study^[9], the higher likelihood of the occurrence of speeding by 20% but less than 50% was observed on weekends in the present study. Such inconsistency may be attributed to the fact that the previous literature studied the influencing factors of traffic violations based on incident case data. Some traffic violations cause incidents, whereas others do not. The higher probabilities of speeding by 20% but less than 50% can be explained by the fact that the driving purpose for leisure is associated with a higher speeding risk compared to the driving purpose for commuting^[7].

3.3 Effects of time of day

Time of day is also identified as a significant factor associated with traffic violations. Illegal parking is more likely to take place at night (19:00—23:59). Drivers are inclined to park illegally when their dwell time is short^[31]. The early morning (00:00—06:59) is still part

of bedtime, and 09:00—16:59 is devoted to working hours, which lead to longer parking times during these periods. During the morning rush hour (07:00—08:59), more drivers are on their way to work and fewer parking behaviors occur. The period of 19:00—23:59 is off-work time, so there may be greater demand for temporary parking, which leads to more illegal parking.

The probabilities of violating traffic signals and driving in the reverse direction are observed to be higher in the evening (19:00—23:59). This is consistent with previous studies^[16,24], which concluded that red-light-running was more likely to happen in the evening (18:00—23:59) than in the morning (06:00—11:59). When it is in the evening (19:00—23:59), the probabilities of driving in the reverse direction increase, which is in accordance with a previous conclusion^[16]. The higher probabilities of the occurrence of violating traffic signals and driving in the reverse direction may be due to bad visibility in the evening. Therefore, on the one hand, it is necessary to improve the lighting conditions; on the other hand, the setting of signal lights and pavement markings should be standardized. Drivers can better see the signal lights and traffic markings, thus reducing the rate of traffic violations. Certainly, the drivers' intentional behaviors should not be ruled out, and these should be solved by strengthening the traffic safety education for drivers.

Furthermore, a higher probability of speeding by 20% but less than 50% is related to 19:00—23:59. There is no significant difference between the probabilities of traffic violations at 00:00—06:59 and at 19:00—23:59. The traffic volumes in these two periods are lower than those in the other four periods, which may result in higher average vehicle speed at 00:00—06:59 and 19:00—23:59. It is possible that police enforcement intensity at dawn and at night is lower than that during the day due to poor visibility; thus, drivers are likely to rely only on their luck, resulting in more traffic violations mentioned above. Thus, the improvement of lighting conditions and enforcement intensity may be helpful in reducing the above traffic violations^[9]. Interval-based speed enforcement systems^[32], social media campaigns, advocacy for legislative change, and law enforcement training^[33] are also effective in reducing the percentage of speeding.

In addition, 09:00—11:59, 12:00—16:59, and 17:00—18:59 are periods of higher rates of violating traffic markings. The risk of failure to yield to pedestrians is especially high during working hours (09:00—11:59 and 12:00—16:59), with OR = 232.173 and OR = 217.717, respectively. The probabilities of traffic violations in the morning rush hour (07:00—08:59, OR = 90.361) and evening rush hour (17:00—18:59, OR = 31.724) are also higher than at night time (19:00—23:59). This may be attributed to the fact that there is a greater volume of pedestrian flow during the daytime. In

order to create a good walking environment for pedestrians, it is necessary to impose heavier punishment on drivers who do not yield to pedestrians. Briefly speaking, targeted traffic intervention measures can be proposed for significant traffic violations with a higher probability of occurrence in different periods.

3.4 Effects of the locations of traffic violations

The risks of use of an improper oriented lane, violating traffic signals, and violating traffic markings at the intersection are much higher than at the road section. Previous studies have also shown that these traffic violations are common behaviors at the intersection^[16,24,29,34-35]. This may be related to the fact that most of the oriented lanes, traffic signals, and traffic markings are placed at the intersection. In view of these common traffic violations at intersections, it is necessary to standardize the setting of traffic signs and markings, improve lighting conditions, reasonably regulate signal timing at signalized intersections, or take reasonable channelization measures. Strengthening the traffic safety education for drivers is also of great importance.

The road section is more likely to be involved in illegal parking, speeding by 20% but less than 50%, and failure to yield to pedestrians. On account of the high-speed growth of motorization, cities frequently permit parking or temporary parking at road sections without heavy traffic. Curb parking influences road capacity and has adverse influences on the operation safety of dynamic traffic^[36]. Moreover, curb parking violations directly cause traffic congestions and interrupt the normal flow of cars on the roads^[30]. There is also a large gap in public parking spaces in Shangyu, and the supply and demand for equipped parking spaces are uneven.

In addition, the parking fees in Shangyu are lower compared to cities with the same development scale. In addition to the consideration of the nature of land use, parking requirements should also be modified according to the specific conditions of the area to formulate more accurate and flexible parking space allocation standards. Moreover, we should give full play to the role of economic leverage, increase the charging standards for on-street parking, and raise the penalties for illegal parking. As for the two violations of speeding by 20% but less than 50% and failure to yield to pedestrians, slowing down before entering the intersection and yielding to pedestrians at the intersection may be normal for drivers, while they could be ignored by drivers at the road section. Therefore, the adjustment of the traffic monitoring system and police enforcement plan is necessary to overcome specific violations that require varying interventions at different locations. Strengthening traffic safety education for drivers can also help reduce these traffic violations.

3.5 Effects of weather

The weather has been found to significantly impact the likelihood of the occurrences of the following violations: use of an improper oriented lane, violating traffic signals, violating traffic markings, driving in the reverse direction, and speeding by 20% but less than 50%. Drivers are more likely to violate traffic signals on sunny and cloudy days. Although the p value of rainy days is not significant at the significance level of 95%, the probability of violating traffic signals is larger on rainy days than on snowy days. Similarly, the probability of violating traffic markings on snowy days is smaller than those under the four other weather conditions. This may be attributed to the fact that sunlight on sunny and cloudy days reduces the visibility of the traffic signal and traffic markings. The rain can also obstruct the driver's view of the traffic signal and markings. Relevant departments should thus improve the design of traffic signals and markings based on the drivers' cognitive characteristics of information.

Meanwhile, compared with other weather conditions, the probability of speeding on snowy days increases, which is in accordance with the conclusions of existing studies^[15-16], indicating that bad weather significantly increases traffic violations. Road skidding can be caused by snowy days, which can lead to unfavorable factors for driving and braking. In turn, these can lead to speeding and even traffic accidents. Therefore, in order to reduce related traffic violations under severe weather conditions, the traffic management department should set up relevant slogans or variable signs to remind drivers to be careful of slippery roads. Relevant departments should also clear the roads in a timely manner during snowy weather.

4 Conclusions

1) Local vehicles are more likely to be involved in various traffic violations. Illegal parking and speeding by 20% but less than 50% are more likely to occur on the weekend.

2) Illegal parking, violating traffic markings, driving in the reverse direction, and speeding by 20% but less than 50% frequently occurred at 19:00—23:59. The periods within 09:00—18:59 show higher probabilities of violating traffic markings, whereas the periods within 09:00—16:59 show higher rates of failure to yield to pedestrians.

3) The risks of use of an improper oriented lane, violating traffic signals, and violating traffic markings at the intersection are much higher, and road section is more likely to be involved in illegal parking, speeding by 20% but less than 50%, and failure to yield to pedestrians.

4) The probability of speeding in snowy weather is higher, while the probabilities of violating traffic signals

and violating traffic markings are higher in other weather conditions.

5) Some targeted traffic intervention measures have been proposed to resolve significant traffic violations with higher probabilities of occurrence in different situations.

References

- [1] World Health Organization. Global status report on road safety 2018[R]. Geneva: WHO, 2018.
- [2] National Bureau of Statistics of China. *China statistical yearbook*[M]. Beijing: China Statistics Press, 2019. (in Chinese)
- [3] Traffic Management Research Institute of the Ministry of Public Security. Road accidents white paper[R]. Beijing: Traffic Management Research Institute of the Ministry of Public Security, 2011. (in Chinese)
- [4] Rezapour Mashhadi M M, Wulff S S, Ksaibati K. Utilizing crash and violation data to assess unsafe driving actions[J]. *Journal of Sustainable Development of Transport and Logistics*, 2017, 2(2): 35 – 46. DOI:10.14254/jsdtl.2017.2-2.3.
- [5] Zhang G N, Yau K K W, Chen G H. Risk factors associated with traffic violations and accident severity in China[J]. *Accident Analysis & Prevention*, 2013, 59: 18 – 25. DOI:10.1016/j.aap.2013.05.004.
- [6] Peden M M, Scurfield R, Sleet D. World report on road traffic injury prevention[R]. Geneva: World Health Organization, 2004.
- [7] Tseng C M. Speeding violations related to a driver's social-economic demographics and the most frequent driving purpose in Taiwan Province's male population[J]. *Safety Science*, 2013, 57: 236 – 242. DOI:10.1016/j.ssci.2013.02.005.
- [8] Oppenheim I, Oron-Gilad T, Parmet Y, et al. Can traffic violations be traced to gender-role, sensation seeking, demographics and driving exposure? [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2016, 43: 387 – 395. DOI:10.1016/j.trf.2016.06.027.
- [9] Zhang G N, Yau K K W, Gong X P. Traffic violations in Guangdong Province of China: Speeding and drunk driving[J]. *Accident Analysis & Prevention*, 2014, 64: 30 – 40. DOI:10.1016/j.aap.2013.11.002.
- [10] Akaateba M A, Amoh-Gyimah R, Amponsah O. Traffic safety violations in relation to drivers' educational attainment, training and experience in Kumasi, Ghana [J]. *Safety Science*, 2015, 75: 156 – 162. DOI:10.1016/j.ssci.2015.02.010.
- [11] Cheng Z Y, Lu J, Zu Z S, et al. Speeding violation type prediction based on decision tree method: A case study in Wujiang, China [J]. *Journal of Advanced Transportation*, 2019, 2019: 1 – 10. DOI:10.1155/2019/8650845.
- [12] Martínez-Ruiz D M, Fandiño-Losada A, Ponce de Leon A, et al. Impact evaluation of camera enforcement for traffic violations in Cali, Colombia, 2008—2014 [J]. *Accident Analysis & Prevention*, 2019, 125: 267 – 274. DOI:10.1016/j.aap.2019.02.002.
- [13] Zhang G N, Tan Y, Jou R C. Factors influencing traffic signal violations by car drivers, cyclists, and pedestrians:

- A case study from Guangdong, China [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2016, **42**: 205 – 216. DOI:10.1016/j.trf.2016.08.001.
- [14] Wang Q, Zhang W, Yang R, et al. Common traffic violations of bus drivers in urban China: An observational study [J]. *PLoS One*, 2015, **10** (9): e0137954. DOI: 10.1371/journal.pone.0137954.
- [15] Precht L, Keinath A, Krems J F. Identifying the main factors contributing to driving errors and traffic violations—results from naturalistic driving data [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2017, **49**: 49 – 92. DOI:10.1016/j.trf.2017.06.002.
- [16] Fu C Y, Liu H. Investigating influence factors of traffic violations at signalized intersections using data gathered from traffic enforcement camera [J]. *PLoS One*, 2020, **15** (3): e0229653. DOI:10.1371/journal.pone.0229653.
- [17] Atombo C, Wu C Z, Zhong M, et al. Investigating the motivational factors influencing drivers intentions to unsafe driving behaviours: Speeding and overtaking violations [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2016, **43**: 104 – 121. DOI:10.1016/j.trf.2016.09.029.
- [18] Tajvar A, Jahangiri M, Aghamolaei T, et al. Investigating self-reported violations of the drivers of Bandar Abbas City and its relation with their knowledge and attitude regarding traffic regulations [J]. *Archives of Trauma Research*, 2019, **8**(4): 219. DOI:10.4103/atr.atr_31_19.
- [19] Wang X X, Xu L J, Hao Y P. What factors predict drivers' self-reported lane change violation behavior at urban intersections? A study in China [J]. *PLoS One*, 2019, **14** (5): e0216751. DOI: 10.1371/journal.pone.0216751.
- [20] Joewono T B, Susilo Y O. Traffic violations by young motorcycle riders on Indonesian urban roads [J]. *Journal of Transportation Safety & Security*, 2017, **9**(Supl): 236 – 261. DOI:10.1080/19439962.2016.1247123.
- [21] Iversen H H, Rundmo T. Changes in Norwegian drivers' attitudes towards traffic safety and driver behaviour from 2000 to 2008 [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2012, **15** (2): 95 – 100. DOI:10.1016/j.trf.2011.12.006.
- [22] Andreuccetti G, Leyton V, Carvalho H B, et al. Drink driving and speeding in Sao Paulo, Brazil: Empirical cross-sectional study (2015–2018) [J]. *BMJ Open*, 2019, **9** (8): e030294. DOI: 10.1136/bmjopen-2019-030294.
- [23] Tavakoli Kashani A, Amirifar S, Azizi Bondarabadi M. Analysis of driver and vehicle characteristics involved in red-light running crashes: Isfahan, Iran [J]. *Iranian Journal of Science and Technology, Transactions of Civil Engineering*, 2021, **45** (1): 381 – 387. DOI:10.1007/s40996-020-00453-2.
- [24] Al-Atawi A M. Characteristics of red light running violations in urban areas in Tabuk, Kingdom of Saudi Arabia [J]. *IATSS Research*, 2014, **37** (2): 119 – 123. DOI: 10.1016/j.iatssr.2013.08.001.
- [25] Zahid M, Chen Y Z, Khan S, et al. Predicting risky and aggressive driving behavior among taxi drivers: Do spatio-temporal attributes matter? [J]. *International Journal of Environmental Research and Public Health*, 2020, **17** (11): 3937. DOI:10.3390/ijerph17113937.
- [26] Jin M H, Shriar A J. Exploring the relationship between social capital and individuals' policy preferences for environmental protection: A multinomial logistic regression analysis [J]. *Journal of Environmental Policy & Planning*, 2013, **15** (3): 427 – 446. DOI: 10.1080/1523908X.2013.769415.
- [27] Kwak C, Clayton-Matthews A. Multinomial logistic regression [J]. *Nursing Research*, 2002, **51** (6): 404 – 410. DOI:10.1097/00006199-200211000-00009.
- [28] Li X, Zhang M M. *SPSS 22.0 statistical analysis from introduction to proficiency* [M]. Beijing: Publishing House of Electronics Industry, 2015. (in Chinese)
- [29] Wang X S, Yu R J, Zhong C J. A field investigation of red-light-running in Shanghai, China [J]. *Transportation Research Part F: Traffic Psychology and Behaviour*, 2016, **37**: 144 – 153. DOI:10.1016/j.trf.2015.12.010.
- [30] Koohpayma J, Tahooni A, Jelokhani-Niaraki M, et al. Spatial analysis of curb-park violations and their relationship with points of interest: A case study of Tehran, Iran [J]. *Sustainability*, 2019, **11** (22): 6336. DOI: 10.3390/su11226336.
- [31] Nourinejad M, Gandomi A, Roorda M J. Illegal parking and optimal enforcement policies with search friction [J]. *Transportation Research Part E: Logistics and Transportation Review*, 2020, **141**: 102026. DOI:10.1016/j.tre.2020.102026.
- [32] Bhalla K, Li Q F, Duan L, et al. The prevalence of speeding and drink driving in two cities in China: A mid project evaluation of ongoing road safety interventions [J]. *Injury*, 2013, **44**: S49 – S56. DOI: 10.1016/S0020-1383(13)70213-4.
- [33] Li Q, He H, Duan L, et al. Prevalence of drink driving and speeding in China: A time series analysis from two cities [J]. *Public Health*, 2017, **144**: S15 – S22. DOI: 10.1016/j.puhe.2016.11.024.
- [34] Ackaah W, Aidoo E N. Modelling risk factors for red light violation in the Kumasi Metropolis, Ghana [J]. *International Journal of Injury Control and Safety Promotion*, 2020, **27** (4): 432 – 437. DOI: 10.1080/17457300.2020.1792936.
- [35] Li Y X, Abdel-Aty M, Yuan J H, et al. Analyzing traffic violation behavior at urban intersections: A spatio-temporal kernel density estimation approach using automated enforcement system data [J]. *Accident Analysis & Prevention*, 2020, **141**: 105509. DOI:10.1016/j.aap.2020.105509.
- [36] Cao Y, Yang Z Z, Zuo Z Y. The effect of curb parking on road capacity and traffic safety [J]. *European Transport Research Review*, 2016, **9** (1): 1 – 10. DOI:10.1007/s12544-016-0219-3.

基于电子执法数据的交通违法行为影响因素研究 ——以上虞区为例

范浩轩 任刚 李豪杰 马景峰

(东南大学交通学院, 南京 211189)

摘要:为了研究交通违法行为的影响因素,以上虞市电子执法数据和历史天气数据为基础,研究了车辆属性、违法时间段、违法时间是否为工作日、违法地点和天气对交通违法行为的影响.从原始数据中确定了10类交通违法行为,运用卡方检验方法分析交通违法行为与潜在影响因素之间的关系.为了进一步估计不同影响因素对交通违法行为发生的可能性的影响,进行了多元 logistic 回归分析.通过对 SPSS 卡方检验结果的分析,上述5个因素均被确定为与交通违法行为显著相关的因素.多元 logistic 回归结果显示,这5个因素对相应的交通违法行为发生的可能性有显著影响.这些结论对制定有效的交通干预措施,减少交通违法行为,提高道路交通安全水平具有重要意义.

关键词:交通违法行为;道路交通安全;电子执法数据;多元 logistic 回归;影响因素

中图分类号:U491.3