TRR
JOURNAL OF THE TRANSPORTATION RESEARCH BOARD

Article

Transportation Research Record
2018, Vol. 2672(38) 22–34

DOI: 10.1177/0361198118790617
journals.sagepub.com/home/trr
SAGE

Exploration of Contributing Factors Related to Driver Errors on Highway Segments

Mohammad Razaur Rahman Shaon[1], Xiao Qin[1], Zhi Chen[1], and Jian Zhang[2]

Abstract

A significant portion of crashes occurred on highway segments, with more than 90% of crashes associated with driving errors. To avoid a crash, a driver needs to detect a hazard, decide the safest driving maneuvers, and execute them properly. Driver errors at any of these sequential phases may lead to a crash; therefore, it is necessary to identify the contributing factors and assess their influence on driver behavior. To assist this investigation, a multinomial probit model was employed to study driver errors reported in crashes in rural and urban areas. The modeling results identified many highway geometric features, traffic conditions, roadway events, and driver characteristics as statistically correlated to different types of driver error. Following the extensive list, the impacts of error-contributing factors were discussed within each error category. This exercise helps to gain a better understanding of similar or varying effects of explanatory variables across different error categories. The broad and insightful information will help researchers and safety professionals to better understand when, where, and how the driver error may lead to a crash and to develop cost-effective preventive countermeasures.

Highway safety analysis is mostly focused on analyzing crash occurrence or severity, where highway and traffic engineering-related data such as roadway geometric characteristics and traffic conditions are used as explanatory variables. It is well known that human factors probably contribute to over 90% of crashes. According to the National Motor Vehicle Crash Causation Survey (NMVCCS), almost 94% of crashes are caused by driver errors (*1*). Without specifically considering driver factors, crash modeling results may be biased due to the absence of human factors. Thus, understanding what contributes to driver error and how to incorporate driver behavior into crash prediction have become increasingly important topics among safety researchers.

Crashes are complex events, as reflected by the 110 data elements recommended in the Model Minimum Uniform Crash Criteria (*2*). Most information can be obtained directly by reviewing detailed crash reports, including police officers' judgment on driver factors contributing to the occurrence of a crash. Crash information can be augmented by socioeconomic, demographic, land use, and traffic pattern information to substantiate the knowledge of how a driver interacts with—and how his/her behavior is influenced by—roadway design, traffic conditions, and other contextual factors.

For a driver, there is a four-phase process of seeing and reacting to a hazard; that is, perception, intellection, emotion, and volition, or "PIEV." An error can happen during any of these four phases. This study is particularly focused on understanding when, where, and how drivers make mistakes that contribute to a crash, drawing cues from a comprehensive list of variables ranging from highway geometry, traffic, roadway, weather, and lighting conditions, events such as construction zone and debris on roadway, as well as driver information such as age, gender, vehicle types, and so forth. Specifically, the categorization method in the NMVCCS study was followed; this groups driver errors into four categories: recognition error, decision error, performance error, and non-performance error. As each error type is specific and unique, the relating explanatory variables identified through statistical models may be more informative to safety professionals. The new insight into the circumstances leading to driver errors and a better

[1]Department of Civil and Environmental Engineering, University of Wisconsin-Milwaukee, Milwaukee, WI
[2]Research Center for Internet of Mobility, School of Transportation, Southeast University, Nanjing, Jiangsu, PR China

Corresponding Author:
Address correspondence to Xiao Qin: qinx@uwm.edu

understanding of possible causes will support the development of tangible, practical, and, more importantly, targeted and cost-effective enforcement strategies, driver education and training programs, engineering solutions, and vehicle safety technologies.

Literature Review

Crash occurrence may be attributed to errors by drivers or the interaction between driver behavior and roadway design features (*3*). As indicated by police records, driver errors can range from a traffic infraction in which the driver is not paying attention, to an intentional traffic violation such as failure to yield or significantly exceeding the speed limit. However, according to the Human Factors Guidelines for Road Systems: "*Road users cannot be expected to solve either highway design or traffic engineering problems without making mistakes and/or compromising operational efficiency and safety*" (*4*).

Understanding the interaction between driver errors and roadway geometric and contextual features is crucial. It has been well established from crash count models that highway design, traffic conditions, and contextual factors such as weather events are related to crash occurrence and have an effect on driver behavior. The American Automobile Association (AAA) Foundation for Traffic Safety estimated that 56% of the fatal crashes that occurred between 2003 and 2007 involved potentially aggressive driving behavior, in which speeding was the most common aggressive action, making up about 31% of total fatal crashes. Hauer noted that the speed at which people choose to travel is affected by roadway design and vehicle characteristics (*5*). Tate and Turner investigated the relationship between observed travel speed, road geometry, and crashes in New Zealand (*6*) and concluded that drivers' speed choices were more strongly related to curve radius than curve design speed, and that the approach speed environment has a significant impact on speed choice. Liu and Chen documented that "driving too fast for conditions" was more likely to occur on roads with higher speed limits (50+ mph) than other crashes (*7*). The authors also noted that a significant proportion of speeding-related crashes occurred on adverse road surface conditions such as "Snowy/Slushy/Icy/Slippery" and "Wet" road pavement compared with other crashes.

Distracted driving is another major driver error that contributes to crashes. Novice drivers appeared to be prone to distraction while driving (*8*). Naturalistic driving studies showed that talking on a cell phone raises the risk of collision by more than 30%, and drivers who text are at 23 times higher crash risk compared with non-distracted drivers (*9*). Results from the National Occupant Protection Use Survey, conducted annually by the National Center for Statistics and Analysis of the National Highway Traffic Safety Administration (NHTSA), showed that females from all age groups are more prone to use electronic devices while driving (*10*). Electronic device use percentage was found to be similar in age brackets from 16 to 69 years. This suggests that gender might be a more important variable than age cohort in distracted driving error.

Work zone is a roadway event that has been reported to increase crash rate, according to previous literature (*11–13*). Drivers in a construction/work zone encounter a complex array of warning signs, barrels, pylons, construction equipment, and machines, which can create hazards for drivers. The new traffic patterns and challenging roadway configurations in work zones, such as stopping or slow traffic, trucks merging from the ramp, uneven pavement, narrowed lanes, and absence of shoulders, require drivers to operate their vehicles with extra caution and impose considerable stress on their driving tasks.

Previously, impaired driving has been identified as a contributing factor to driver error (*14*). Use of alcohol can significantly affect a driver's decision-making process. Blomberg et al. conducted a case-control study to explore the relationship of blood alcohol concentration (BAC) with relative crash risk (*14*). Results showed that elevated relative risk beginning at 0.05–0.06% BAC, with an accelerating increase in risk at BACs greater than 0.10%. In 2015 in the U.S., 41% of drivers killed in roadway speeding-related crashes had a BAC of 0.08 g/dl or higher in their blood (*15*). Besides alcohol impairment, drug-impaired driving has a significant effect on driving behavior (*16–18*). The Governors Highway Safety Association (GHSA) sponsored a study that found that fatalities caused by drugged driving surpass those caused by alcohol-impaired driving in the United States (*19*). In 2015, 43% of motorists who died in a road accident had drugs in their systems, whereas 37% of motorists who died tested positive for alcohol.

Several studies have investigated driver error for segment and intersection-related crashes (*7*, *20–23*), most of which discussed factors contributing to driver error for specific types of crashes, such as speed. An overall discussion of the factors contributing to driver error is rare. Wang and Qin investigated the factors contributing to driver errors at uncontrolled, sign-controlled, and signal-controlled intersections (*23*). The authors categorized driver error based on traffic violations recorded in the crash report, roadway characteristics (presence of curve, visibility, speed limit), driver characteristics (age, gender, driving under the influence), environmental characteristics (weather condition, roadway condition, lighting condition), and vehicle type (passenger car, light truck, heavy truck) to predict different types of driver error. Sign-controlled intersections are found to have the highest percentage of driver error, followed by signalized, and then uncontrolled intersections. Drivers are more prone to serious errors if their vision is obscured. Adverse environmental characteristics, such as snow or ice on the pavement, negatively affect the severity of driver error. Driver age, gender, and alcohol or drug use greatly influence the severity of error outcome. The findings confirmed that driver errors are the outcome not only of a driver's psychological behavior but also of the interaction with external factors during the driver's decision making.

Based on previous research, this study is an attempt to investigate the relationship between driver errors and observable factors on highway segments in both rural and urban areas. More information on these contributing factors would help researchers and safety professionals to develop cost-effective countermeasures.

Methodology

The multinomial probit model (MNP) is a discrete outcome model that considers a response variable with three or more levels without accounting for order between levels. Two popular choices to model multiclass categorical variables are multinomial logit (MNL) and MNP models. The MNL model is built on the independence of irrelevant alternatives assumption, meaning adding or deleting an alternative will not change the ratio between the probability of any pair of existing alternatives. In simple words, MNL cannot account for correlation between any pair of existing alternatives. This may not be always true if the dependent variable categories are correlated. The MNP model relaxes the independence assumption (*24*). The driver error categories defined in this study are not independent as they are categorized based on sequential events. From a practical perspective, it is obvious that performance error depends on the decision of activity the driver tends to execute, and decision error depends on the recognition of hazardous situation perceived by the driver. Considering dependency between driver error categories, the MNP model was an appropriate choice for model development in this study. The utility function of the MNP model that determines the preference or possible value of attaining the outcome i (i = 1, 2, . . ., I) for observation n can be written as (*25*):

$$U_{in} = \beta_i X_{in} + \varepsilon_{in} \tag{1}$$

$$[\varepsilon_{1n}, \varepsilon_{2n}, \varepsilon_{3n}, \ldots\ldots, \varepsilon_{in}] \sim MVN(0, \Sigma)$$

where,

X_{in} = vector of independent variables for nth observation with ith outcome

β_i = vector of corresponding unknown coefficients

ε_{in} = disturbance term that accounts for unobserved effects and random errors

The disturbance term ε_{in} for ith driver error type has a mean of zero and is correlated among different error types. Thus, the disturbance vector is defined by a multivariate normal distribution. The probability of ith driver error can be written as:

$$Prob[Choice_{in}] = Prob[U_{in} > U_{jn}, j = 1, 2, 3, \ldots\ldots, I; i \neq j] \tag{2}$$

Using above formulation, the probability of occurrence of ith driver error can be specified as:

$$\begin{aligned} &Prob[Choice_{in}|X_n] > Prob[(\varepsilon_{in} - \varepsilon_{1n})\rangle X_n'(\beta_1 - \beta_i), \\ &(\varepsilon_{in} - \varepsilon_{2n}) > X_n'(\beta_2 - \beta_i), \ldots., (\varepsilon_{in} - \varepsilon_{(i-1)n}) \\ &> X_n'\left(\beta_{(i-1)} - \beta_i\right), (\varepsilon_{in} - \varepsilon_{(i+1)n}) \\ &> X_n'\left(\beta_{(i+1)} - \beta_i\right), \ldots.(\varepsilon_{in} - \varepsilon_{In}) > X_n'(\beta_1 - \beta_I)] \end{aligned} \tag{3}$$

The estimated coefficient β_i can be interpreted as the marginal effect of X_i on the log odds ratio of ith alternative to the baseline alternative. A "margin" is a statistic computed from predictions from a model while manipulating the values of the covariates. The marginal effect of X_i on the probability of ith alternative can be expressed as follows:

$$\frac{\partial P_r(Y = Y_n|X_{in})}{\partial X_i} = \frac{\partial E(Y_n|X_i)}{\partial X_i} = \emptyset(\beta_i X_{in})\beta_i \tag{4}$$

where, $\emptyset$ represents cumulative normal density function.

Data Description

Data for segment-related crashes that occurred on the Wisconsin state trunk network system between 2013 and 2015 were collected, excluding deer-related crashes (*26*). After removing all crashes without good location information, 48,441 rural crashes and 46,221 urban crashes were available. Specific driver errors were extracted from the Wisconsin Motor Vehicle Accident Reporting Form 4000 (MV4000), in which the investigating police officers documented detailed accident information (*26*, *27*). There is a list of fourteen driver-related factors. When a crash is associated with multiple driver factors, the most severe driver factor is noted based on the police investigation.

Modeling fourteen choices may not be effective because of the sample size, strong correlation between some error types, and difficulties of interpretation. Based on the similarities in driver errors, the NMVCCS study classified driver-related critical reasons into recognition errors, decision errors, performance errors, and non-performance errors (*1*). Recognition error includes driver inattention, internal and external distraction, inadequate surveillance, and so forth; aggressive driving behavior, driving too fast, and so on, are categorized as decision error; overcompensation and poor directional controls are categorized as performance error; sleep and physical impairment are considered as non-performance error. This categorization combines narrative errors with similar traits. The driver factors in Wisconsin crash data were grouped into four NMVCCS driver error categories based on the NMVCCS criteria and definition of each category. Table 1 shows the NMVCCS driver error types, and corresponding Wisconsin driver factors along with summary statistics for each category.

The broad categorization of driver errors follows a sequence of information processing. When driving, a driver needs to detect and identify a hazard, decide what to do, and react accordingly. Driver errors leading to a crash are also categorized following the driving tasks. A driver's recognition efficiency can be affected by any internal or external distraction or by any form of inattention. Recognition error refers to all the driver factors that may lead to lack of awareness or failure in recognition of hazardous situations. In Wisconsin, 18% and 20% of total crashes that occurred between 2013 and 2015 were attributable to inattentive driving in rural and urban areas, respectively.

A driver's decision on what to do directly leads to what happens next, whether it is a decision after detecting a hazard or a decision while driving. A bad maneuver decision after recognizing a hazardous situation may cause a crash. A reckless decision such as "exceeding the speed limit" may go wrong even without an imminent hazard. In Wisconsin, 35% and 38% of crashes occurred as a result of decision error in rural and urban areas, respectively.

If a maneuver is not properly performed, a crash may happen. Poorly performed driving tasks are categorized as performance error, which is dependent on the driver's experience and skills. Although non-performance error is not related to driver behavior, it represents a driver's health conditions, fatigue, level of impairment, or other non-performance issues.

The crash dataset does not contain detailed roadway geometric information at the crash location. Roadway geometry, pavement characteristics, mobility, safety, and other roadway-related data tables stored in Meta-Manager at the Wisconsin Department of Transportation (WisDOT) were linked with crash data using spatial join in ArcGIS. The joined dataset contains the data collected by the crash investigating police officer, roadway geometry, and traffic information for each crash. Table 2 provides the summary statistics.

Analysis of Results

The coefficient estimates of the final MNP models for rural and urban crashes are presented in Tables 3 and 4, respectively. The STATA command "mprobit" was used to estimate the coefficient of the MNL model (*28*). In

Table 1. Categorization and Distribution of Driver Error

Error type	NMVCCS criteria	Wisconsin criteria	Rural	Urban
Recognition error	• Inadequate surveillance • Internal distraction • External distraction • Inattention	• Inattentive driving	8,659 (17.88%)	9,044 (19.57%)
Decision error	• Too fast for conditions • Too fast for curve • False assumption of other's action • Illegal maneuver • Misjudgment of gap or other's action • Following too closely • Aggressive driving behavior	• Too fast for condition • Exceed speed limit • Disregard traffic control • Following too close • Improper overtake • Improper turn	17,139 (35.38%)	17,662 (38.21%)
Performance error	• Overcompensation • Poor directional control • Panic/Freezing • Other performance error	• Failure to keep vehicle under control • Left of center • Unsafe backing • Failure to yield	10,288 (21.24%)	9,867 (21.35%)
Non-performance error	• Sleep • Heart attack • Other non-performance error	• Disability • Driver condition • Others	2,402 (4.96%)	3,030 (6.55%)
No error			9,953 (20.55%)	6,620 (14.32%)

Table 2. Summary Statistics of Explanatory Variables

			Rural		Urban	
Variable	Description	Type	Mean	Std. dev.	Mean	Std. dev.
AADT	Annual average daily traffic	Continuous	21,610.37	26,554.21	55,450.19	48,566.43
Truck	Truck percentage (%)	Continuous	11.44	4.59	7.735	2.86
Speed	Posted speed limit (MPH)	Continuous	57.49	11.31	46.95	14.17
Lane	Number of lanes (Count)	Continuous	2.13	0.43	2.59	0.715
LW	Lane width (feet)	Continuous	12.10	0.83	12.34	1.08
SW	Shoulder width (feet)	Continuous	8.60	3.87	5.58	5.49
Rut	Pavement rutting (inch)	Continuous	0.088	0.08	0.07	0.07
Percent passing	Passing percentage (%)	Continuous	26	31.90	3.25	15.85
Highway type	Interstate	Categorical with 3 levels	9,840 (20.31%)		12,012 (25.99%)	
	State highway		37,377 (77.16%)		30,062 (65%)	
	Other state roadway		1,224 (2.53%)		4,147 (9%)	
Roadway type	Undivided	Categorical with 3 levels	24,177 (49.91%)		8,225 (17.79%)	
	Divided		23,889 (49.32%)		36,324 (78.59%)	
	One-way		375 (0.77%)		1,672 (3.62%)	
Presence of median	No	Categorical with 2 levels	31,530 (65.09%)		20,170 (43.64%)	
	Yes		16,911 (34.91%)		26,051 (56.36%)	
Roadway condition	Dry	Categorical with 4 levels	27,830 (57.45%)		31,740 (68.67%)	
	Wet		5,255 (10.85%)		7,517 (16.26%)	
	Snow		10,281 (21.22%)		5,307 (11.48%)	
	Ice		5,075 (10.48%)		1,657 (3.58%)	
Weather condition	Clear	Categorical with 5 levels	20,591 (42.51%)		22,378 (48.42%)	
	Fog/cloudy		13,619 (28.11%)		14,671 (31.74%)	
	Wind		1,041 (2.15%)		140 (0.3%)	
	Rain		3,057 (6.31%)		4,157 (8.99%)	
	Snow/sleet		10,133 (20.92%)		4,875 (10.55%)	
Lighting condition	Day	Categorical with 3 levels	33,065 (68.26%)		30,046 (73.66%)	
	Night-unlit		13,477 (27.82%)		3,026 (6.55%)	
	Night-lit		1,899 (3.92%)		9,149 (19.79%)	
Horizontal curve	No	Categorical with 2 levels	39,390 (81.32%)		41,750 (90.33%)	
	Yes		9,051 (18.68%)		4,471 (9.67%)	
Vertical curve	No	Categorical with 2 levels	38,865 (80.23%)		40,542 (87.71%)	
	Yes		9,576 (19.77%)		5,679 (12.29%)	
Age group	Adolescent (<18 years)	Categorical with 5 levels	2,363 (4.88%)		1,789 (3.87%)	
	Young adults (18–25 years)		11,206 (23.13%)		11,271 (24.39%)	
	Adults (26–35 years)		10,309 (21.28%)		11,406 (24.68%)	
	Middle age (36–65 years)		20,223 (41.47%)		18,294 (39.58%)	
	Old (>65 years)		4,340 (8.96%)		3,461 (7.49%)	
Gender	Male	Categorical with 2 levels	30,090 (62.12%)		26,989 (58.39%)	
	Female		18,351 (37.88%)		19,232 (41.61%)	
Vehicle	Passenger car	Categorical with 4 levels	35,498 (73.28%)		38,051 (82.32%)	
	Motorcycle		888 (1.83%)		545 (1.18%)	
	Light truck		8,096 (16.71%)		4,898 (10.6%)	
	Heavy truck		3,959 (8.17%)		1,727 (5.9%)	
Alcohol	No	Categorical with 2 levels	45,725 (94.39%)		44,470 (96.21%)	
	Yes		2,716 (5.61%)		1,751 (3.79%)	
Drug	No	Categorical with 2 levels	47,920 (98.92%)		45,881 (99.26%)	
	Yes		521 (1.08%)		340 (0.74%)	
Visibility obscured	No	Categorical with 2 levels	48,078 (99.25%)		46,013 (99.55%)	
	Yes		363 (0.75%)		208 (0.45%)	
Work zone	No	Categorical with 2 levels	47,625 (98.32%)		45,339 (98.09%)	
	Yes		816 (1.68%)		882 (1.91%)	
Debris on road	No	Categorical with 2 levels	47,695 (98.46%)		45,860 (99.22%)	
	Yes		746 (1.54%)		361 (0.78%)	

Table 3. Coefficient Estimates for MNP Model for Driver Errors in Rural Crashes

Variable	Recognition error		Decision error		Performance error	
	Coef.	Std. err.	Coef.	Std. err.	Coef.	Std. err.
AADT (in thousands)	**0.002**	**0.000**	**0.010**	**0.000**	**0.004**	**0.000**
Truck	**–0.005**	**0.003**	**–0.006**	**0.002**	**–0.011**	**0.002**
Speed	**–0.02**	**0.002**	**–0.019**	**0.002**	**–0.01**	**0.002**
Lanes	–0.046	0.032	**–0.084**	**0.03**	**–0.12**	**0.032**
Shoulder width	**0.012**	**0.004**	0.001	0.004	**–0.014**	**0.004**
Pavement rutting	**–0.462**	**0.166**	**–0.515**	**0.154**	**–0.518**	**0.158**
Highway type						
Interstate	Base condition					
State highway	–0.033	0.037	**–0.102**	**0.031**	**0.08**	**0.034**
Other state roadway	–0.114	0.077	**–0.186**	**0.072**	**0.338**	**0.072**
Roadway type						
Undivided	0.07	0.047	**–0.146**	**0.043**	4E-5	0.044
Divided	Base condition					
One-way	**–0.36**	**0.143**	0.036	0.123	–0.026	0.13
Horizontal curve						
No	Base condition					
Yes	**0.153**	**0.031**	**0.255**	**0.027**	**0.356**	**0.028**
Vertical curve						
No	Base condition					
Yes	-0.029	0.029	**0.065**	**0.025**	**0.054**	**0.026**
Roadway condition						
Dry	Base condition					
Wet	–0.026	0.048	**0.413**	**0.045**	**0.208**	**0.046**
Snow	**–0.892**	**0.057**	**1.072**	**0.039**	**0.284**	**0.042**
Ice	**–1.56**	**0.076**	**0.947**	**0.038**	**0.18**	**0.041**
Weather condition						
Clear	Base condition					
Fog/cloudy	**0.157**	**0.026**	**0.16**	**0.025**	**0.19**	**0.026**
Wind	**–0.895**	**0.169**	0.054	0.070	–0.068	0.077
Rain	**–0.163**	**0.065**	**0.269**	**0.057**	0.019	0.06
Snow/sleet	**–0.335**	**0.063**	**0.17**	**0.041**	0.085	0.044
Lighting condition						
Day	Base condition					
Night-unlit	**–0.207**	**0.026**	**–0.341**	**0.023**	**–0.116**	**0.024**
Night-lit	–0.020	0.059	**–0.281**	**0.056**	**–0.194**	**0.058**
Visibility						
No	Base condition					
Yes	**–0.364**	**0.132**	–0.046	0.108	–0.105	0.114
Work zone						
No	Base condition					
Yes	**0.210**	**0.082**	**0.565**	**0.076**	–0.074	0.089
Debris on road						
No	Base condition					
Yes	**–2.136**	**0.117**	**–1.863**	**0.094**	**–1.762**	**0.101**
Age group						
Adolescent	Base condition					
Young adults	**–0.181**	**0.056**	**–0.174**	**0.052**	**–0.178**	**0.054**
Adults	**–0.420**	**0.056**	**–0.341**	**0.052**	**–0.278**	**0.054**
Middle age	**–0.561**	**0.053**	**–0.527**	**0.050**	**–0.44**	**0.051**
Old	**–0.339**	**0.061**	**–0.582**	**0.058**	**–0.211**	**0.059**
Gender						
Male	0.022	0.024	0.006	0.021	**–0.056**	**0.022**
Female	Base condition					
Vehicle						
Passenger car	**0.194**	**0.040**	**0.332**	**0.037**	**0.216**	**0.039**
Motorcycle	**–0.564**	**0.091**	**0.225**	**0.081**	**0.488**	**0.078**
Light truck	**0.193**	**0.046**	**0.325**	**0.041**	**0.237**	**0.044**
Heavy truck	Base condition					

(continued)

Table 3. (continued)

Variable	Recognition error Coef.	Recognition error Std. err.	Decision error Coef.	Decision error Std. err.	Performance error Coef.	Performance error Std. err.
Alcohol						
No	Base condition					
Yes	**1.120**	**0.067**	**1.282**	**0.066**	**1.459**	**0.065**
Drug						
No	Base condition					
Yes	**0.754**	**0.129**	**0.740**	**0.129**	**0.875**	**0.128**
Intercept	**1.277**	**0.255**	**1.324**	**0.244**	**1.138**	**0.252**

Note: Variables that are statistically significant at 90% confidence interval are presented in bold font.

Table 4. Coefficient Estimates for MNP Model for Driver Errors in Urban Crashes

Variable	Recognition error Coef.	Recognition error Std. err.	Decision error Coef.	Decision error Std. err.	Performance error Coef.	Performance error Std. err.
AADT (in thousands)	–3E-05	0.000	**0.002**	**0.000**	–0.0002	0.000
Truck	**–0.015**	**0.004**	**–0.015**	**0.004**	**–0.01**	**0.004**
Speed	**–0.006**	**0.001**	–0.001	0.001	**–0.011**	**0.001**
Lanes	**0.057**	**0.019**	**0.049**	**0.018**	**0.095**	**0.019**
Lane width	**0.047**	**0.011**	**0.045**	**0.011**	**0.036**	**0.011**
Shoulder width	0.003	0.003	**0.006**	**0.003**	**0.017**	**0.003**
Pavement rutting	**–0.482**	**0.162**	**–0.393**	**0.152**	**0.387**	**0.159**
Percent passing	**0.005**	**0.001**	**0.003**	**0.001**	**0.002**	**0.001**
Highway type						
Interstate	Base condition					
State highway	0.001	0.033	–0.017	0.029	**–0.244**	**0.032**
Other state roadway	**–0.12**	**0.049**	**–0.335**	**0.045**	**–0.146**	**0.047**
Roadway type						
Undivided	0.043	0.043	**–0.105**	**0.04**	**–0.069**	**0.042**
Divided	Base condition					
One-way	–0.056	0.063	**–0.171**	**0.06**	**–0.248**	**0.064**
Horizontal curve						
No	Base condition					
Yes	**–0.196**	**0.043**	0.053	0.037	**0.283**	**0.038**
Vertical curve						
No	Base condition					
Yes	**–0.082**	**0.036**	–0.029	0.032	**–0.12**	**0.034**
Roadway condition						
Dry	Base condition					
Wet	–0.069	0.045	**0.157**	**0.042**	**0.203**	**0.044**
Snow	**–0.802**	**0.063**	**0.323**	**0.049**	**0.104**	**0.052**
Ice	**–1.991**	**0.123**	0.061	0.055	**–0.408**	**0.063**
Weather condition						
Clear	Base condition					
Fog/cloudy	**0.145**	**0.026**	**0.168**	**0.025**	**0.198**	**0.026**
Wind	–0.456	0.303	0.119	0.173	–0.074	0.2
Rain	**–0.153**	**0.06**	**0.219**	**0.054**	0.077	0.057
Snow/sleet	**–0.296**	**0.07**	**0.181**	**0.052**	**0.12**	**0.056**
Lighting condition						
Day	Base condition					
Night-unlit	**–0.19**	**0.05**	**–0.321**	**0.043**	**–0.159**	**0.047**
Night-lit	**–0.243**	**0.029**	**–0.37**	**0.027**	**–0.225**	**0.028**
Visibility						
No	Base condition					
Yes	**–0.588**	**0.166**	**–0.682**	**0.153**	0.22	0.142

(continued)

Table 4. (continued)

Variable	Recognition error Coef.	Recognition error Std. err.	Decision error Coef.	Decision error Std. err.	Performance error Coef.	Performance error Std. err.
Construction zone						
No	Base condition					
Yes	–0.043	0.083	**0.289**	**0.075**	–0.052	0.084
Debris on road						
No	Base condition					
Yes	**–2.04**	**0.149**	**–1.895**	**0.113**	**–1.925**	**0.14**
Age group						
Adolescent	Base condition					
Young adults	**–0.202**	**0.065**	**–0.213**	**0.061**	**–0.188**	**0.065**
Adults	**–0.376**	**0.065**	**–0.431**	**0.061**	**–0.296**	**0.065**
Middle age	**–0.393**	**0.064**	**–0.5**	**0.059**	**–0.307**	**0.063**
Old	**–0.271**	**0.073**	**–0.531**	**0.069**	**0.001**	**0.072**
Gender						
Male	**–0.069**	**0.024**	–0.011	0.022	–0.035	0.023
Female	Base condition					
Vehicle						
Passenger car	**0.413**	**0.046**	**0.512**	**0.042**	**0.44**	**0.046**
Motorcycle	**–0.536**	**0.118**	0.02	0.099	**0.436**	**0.099**
Light truck	**0.44**	**0.056**	**0.512**	**0.051**	**0.434**	**0.054**
Heavy truck	Base condition					
Alcohol						
No	Base condition					
Yes	**0.917**	**0.086**	**0.849**	**0.083**	**1.226**	**0.083**
Drug						
No	Base condition					
Yes	**0.667**	**0.18**	**0.527**	**0.176**	**0.783**	**0.176**
Intercept	0.035	0.195	–0.011	0.184	0.036	0.194

Note: Variables that are statistically significant at 90% confidence interval are presented in bold font.

both tables, the coefficient estimates represent the log odds ratio between the probability of a defined driver error type and no error type, with a positive sign for increase and a negative sign for decrease. "No error" category was considered as the base outcome in the MNP model.

The modeling results for "Non-performance error" were excluded as this error category does not include behavioral driver factors. For a quick summary, middle age and old age groups are more prone to non-performance error. Alcohol and drug consumption also increases the probability of non-performance error compared with no error.

In Table 3, it can be noted that both traffic variables annual average daily traffic (AADT) and truck percentage are significantly related to all driver error categories. A thousand-unit change in AADT results in increased probability 1.002 ($e^{0.002}$) times, 1.01 ($e^{0.010}$) times and, 1.004 ($e^{0.004}$) times in recognition error, decision error, and performance error compared with no error, respectively. The signs of estimated coefficients of truck percentage, speed, number of lanes, shoulder width, and pavement rutting suggest the reduction in the probability of an error compared with no error.

Interesting results found in roadway classification show that the highway type is significantly related to both decision and performance error types, but not recognition error. This suggests that a driver's recognition/inattentive driving error may not depend on highway type. Decision error mostly occurs on Interstate highway, whereas performance error occurs least on the Interstate. The change in the probability of performance error is the highest in other highways which include rural city or town roads. One-ways reduce recognition error, and undivided highways lead to fewer decision errors. Horizontal and vertical curves significantly increase the probability of all error categories, with a maximum increase in performance error for horizontal and decision error for vertical curves.

Roadway events have a significant effect on driver errors. A comparison between roadway and weather condition variables illustrates a few important observations. For example, snowy pavement increases decision error from no error by 4.13 times, whereas snow precipitation only increases this by 1.24 times. Another important observation is that snowy pavement has a higher increase in probability than icy pavement. Drivers tend to be more cautious during adverse weather events because of the negative impact of recognition error. A construction zone increases the probability of decision and recognition error but is not statistically significant for performance error. The negative impact of roadway debris on all types of errors suggests drivers may be more vigilant toward unusual objects on the roadway.

Driver age, gender, vehicle type, alcohol, and drug impairment were found to be statistically significant in predicting all driver error categories. Adolescents are more prone to driver errors compared with all other age groups. For decision error, the probability gradually reduces with the increase in age. But for performance and recognition error types, old drivers are more prone to error compared with young and middle-aged drivers. Decision and recognition error types do not vary by gender, whereas female drivers were found to have a higher probability of performance error. Motorcycle drivers are least likely to have a recognition error, but they are most likely to commit a performance error. Alcohol or drug impairment increases the probability of all error categories, with a maximum increase in performance error.

Table 4 provides the coefficient estimates for the MNP model with urban crash data. Except for median variable, all explanatory variables were found statistically significant at 10% significance level to predict driver error categories.

There are dissimilarities found in the urban crash analysis compared with rural crashes. AADT is only significant in predicting decision error. This means the probability of making a performance or recognition error in an urban setting does not vary by AADT. The posted speed limit does not affect decision error, which is counterintuitive because one of the major driver errors in this category is "Exceeding Speed Limit." Plausibly, speed violation-related crashes may occur at any posted speed limit. The number of lanes, lane width, shoulder width, and passing percent have a positive effect on driver errors.

For the highway type variable, both decision and performance errors mostly occur on Interstate highways in urban areas. For recognition error, other highway types have the highest increase in probability compared with no error. The roadway type variable is not significant in predicting recognition error but is significant for both decision and performance error types at all levels. Divided highways increase the probability of both decision and performance error types compared with no error. In urban areas, drivers are least likely to make performance mistakes with ice on the roadway. For other explanatory variables, similar trends as discussed for rural crashes are observed.

Discussion of Factors Contributing to Driver Errors

With numerous factors contributing to a variety of driver error types, it is challenging to summarize their individual effects. Thus, a review of contributing factors by error type is necessary. The marginal effect of a variable in the driver error model for rural crashes can be seen in Table 5. The marginal effect has varying definitions based upon the variable type. For a continuous independent variable, the marginal effect is the difference in the probability at each level following a one-unit change; for a categorical independent variable, the marginal effect is calculated as the changes in the probability for each level caused by a change in the value from its base condition.

Traffic and roadway variables significantly affect the probability of decision error. For example, a unit increase in AADT increases the probability of decision error, whereas an increase in truck percentage decreases the probability. In the other error types, the marginal effect of truck percentage and AADT are not statistically significant. However, the posted speed limit decreases the probability of recognition error but increases the probability of performance error.

Compared with no errors, the higher probability of recognition error is likely to happen on undivided highways and/or at places where vertical and horizontal curves are present. On foggy/cloudy and/or windy days drivers are more likely to make recognition mistakes. Furthermore, nighttime with (street) light has a positive impact on recognition error. On the other hand, drivers are less likely to commit a recognition error on other highways and one-way streets than on the Interstate and state highways. In addition, recognition error is low when the pavement is either wet or covered in snow, or weather type is snow/sleet/rain, or nighttime without light. This suggests that drivers may exercise caution when traveling in adverse weather or dark conditions. Similarly, when visibility is low or roadway debris is present, the probability of making a recognition error is low. Another source of low recognition error is people that are older than 18, with middle-aged drivers having the lowest probability of recognition error. For different vehicle types, motorcyclists have the lowest probability of recognition error.

Despite the fact that recognition error shares many similar circumstances leading to decision error, the latter is more likely to take place on Interstate highways or one-way streets, but less likely to take place on undivided highways, whereas the opposite pattern is observed for recognition error. A deterrent for decision error seems to be poor pavement condition (i.e., large rutting value). However, the probability of decision error is higher under adverse weather (e.g., fog/cloudy, snow/sleet, rain) and/or on slippery pavement (snow, ice, and wet) as well as the night condition irrespective of the availability of street lighting. Heavy trucks have the lowest probability of decision error among all vehicle types, perhaps because of the safety regulations imposed on drivers. Work zones may

Table 5. Review of Marginal Effects for Rural Crashes

Variable	Recognition error	Decision error	Performance error
Traffic variables	–	Truck (–0.0001)	Truck (–0.0001)
	AADT (0.0002)	AADT (0.0001)	–
Roadway geometry	–	–	Lanes (-0.015)
	Speed (–0.002)	Speed (–0.003)	Speed (0.001)
Highway type (base: interstate)	Other highways (–0.028)	Other highways (–0.076)	Other highways (0.014)
	–	State highways (–0.032)	State highways (0.006)
Roadway type (base: divided)	Undivided (0.026)	Undivided (–0.042)	Undivided (0.015)
	One-way (–0.043)	One-way (0.041)	–
Alignment	Ver.Curve = Yes (0.010)	Hor.Curve = Yes (0.0177)	Hor.Curve = Yes (0.048)
	Hor.Curve = Yes (0.009)	Ver.Curve = Yes (0.0129)	Ver.Curve = Yes (0.010)
Pavement	–	Rutting (–0.047)	–
Roadway condition (base: dry)	Snow (-0.197)	Snow (0.342)	
	Wet (-0.048)	Wet (0.094)	Wet (0.015)
		Ice (0.340)	
Weather condition (base: clear)	Fog/cloudy (0.007)	Fog/cloudy (0.011)	Fog/cloudy (0.017)
	Snow/sleet (–0.085)	Snow/sleet (0.061)	Snow/sleet (0.025)
	Rain (–0.045)	Rain (0.082)	–
	Wind (0.137)	–	–
Lighting condition (base: day)	Night-unlit (–0.012)	Night-unlit (0.061)	Night-unlit (0.026)
	Night-lit (0.021)	Night-lit (0.056)	Night-lit (0.020)
Events	Debris = Yes (–0.145)	Debris = Yes (–0.226)	Debris = Yes (–0.131)
	–	Work zone = Yes (0.145)	Work zone = Yes (0.084)
	Visibility: Yes (–0.058)	–	–
Impairment	Alcohol (–0.028)	Alcohol (0.022)	Alcohol (0.073)
	–	Drug (0.021)	Drug (0.060)
Age (base: adolescent)	Old (–0.022)	Old (–0.112)	Old (0.023)
	Adult (–0.040)	Adult (–0.033)	–
	Middle age (–0.046)	Middle age (–0.058)	Middle age (–0.016)
Gender (base: female)	–	–	Male (–0.003)
Vehicle type (base: heavy truck)	Motorcycle (–0.104)	Motorcycle (0.032)	Motorcycle (0.133)
	–	Passenger car (0.053)	Passenger car (0.014)
	–	Light truck (0.047)	–

Note: Marginal effect presented with "–" is not significant at 90% confidence interval.

see a higher probability of decision error. Finally, decision error is increased by the use of alcohol and/or drugs.

Compared with the other error types, performance error is the most probable with the change in roadway geometry and traffic configuration. When making comparisons with no error, the probability of performance error is high for all highway types, horizontal or vertical alignments, adverse weather, wet pavement surfaces, and during the night. Similar to decision error, performance error is more likely to occur when a work zone is present and is aggravated by the use of alcohol and drugs. Finally, both motorcycles and passenger cars are associated with a higher probability of performance error than truck drivers. On the other hand, middle-aged male drivers have the lowest probability of committing performance error.

Table 6 provides the estimates of marginal effects of covariates for urban crashes. Only the statistically significant variables at 90% confidence interval are shown in the table.

Similar to rural crashes, traffic variables significantly affect the probability of decision error. An increase in AADT increases the probability of decision error, and an increase in truck percentage decreases the probability of decision error. In urban crashes, changes in roadway geometric configuration affect performance error more compared with other error categories. The effect of highway type, roadway type, and existence of horizontal and vertical curve on recognition error in urban crashes is similar to rural crashes. The weather and roadway conditions are more likely to affect decision error, whereas only wet or snowy roadway surface is responsible for performance error.

The gender variable is only statistically significant in predicting the probability of recognition error. Males are less likely to commit recognition error in urban crashes. Having alcohol in the blood while driving contributes to the probability of all driver error, but the presence of drug only affects the probability of decision error. This coincides with practical knowledge that having taken a

Table 6. Review of Marginal Effects for Urban Crashes

Variable	Recognition error	Decision error	Performance error
Traffic variables	–	Truck (–0.002)	Truck (0.0001)
	AADT (–1.45E–7)	AADT (4.89E–7)	–
Roadway geometry	–	Speed (0.002)	Speed (–0.002)
	–	–	Lanes (0.013)
	Lane width (0.004)	Lane width (0.006)	–
	–	–	Shoulder width (0.003)
	Percent passing (0.001)	–	–
Highway type (base: interstate)	Other highways (0.011)	Other highways (–0.070)	–
	State highways (–0.015)	State highways (–0.019)	State highways (-0.057)
Roadway type (base: divided)	Undivided (0.022)	Undivided (–0.029)	–
	–	One-way (-0.026)	One-way (–0.039)
Alignment	Hor.Curve = Yes (–0.060)	–	Hor.Curve = Yes (0.075)
	Ver.Curve = Yes (–0.006)	Ver.Curve = Yes (0.014)	Ver.Curve = Yes (–0.017)
Pavement	Rutting (–0.084)	Rutting (–0.105)	Rutting (0.164)
Roadway condition (base: dry)	Snow (–0.159)	Snow (0.157)	Snow (0.031)
	Wet (–0.040)	Wet (0.037)	Wet (0.041)
	Ice (–0.210)	Ice (0.200)	–
Weather condition (base: clear)	–	Fog/cloudy (0.011)	Fog/cloudy (0.016)
	Snow/sleet (–0.074)	Snow/sleet (0.073)	Snow/sleet (0.030)
	Rain (–0.056)	Rain (0.067)	–
	–	Wind (0.091)	–
Lighting condition (base: day)	Night-unlit (–0.004)	Night-unlit (0.061)	Night-unlit (0.008)
	Night-lit (–0.002)	Night-lit (0.059)	Night-lit (0.002)
Events	Debris = Yes (–0.142)	Debris = Yes (–0.230)	Debris = Yes (–0.142)
	Work zone = Yes (–0.003)	Work zone = Yes (0.096)	Work zone = Yes (–0.040)
	Visibility: Yes (–0.075)	Visibility: Yes (–0.162)	Visibility: Yes (0.183)
Impairment	Alcohol (–0.033)	Alcohol (–0.074)	Alcohol (0.064)
		Drug (–0.051)	–
Age (base: adolescent)	Young Adult (–0.006)	–	–
	Adult (–0.052)	Adult (–0.053)	–
	Middle age (–0.015)	Middle age (–0.075)	Middle age (–0.011)
	–	Old (–0.134)	Old (0.075)
Gender (base: female)	Male (–0.013)	–	–
Vehicle type (base: heavy truck)	Passenger car (0.023)	Passenger car (0.079)	Passenger car (0.031)
	Motorcycle (–0.095)	–	Motorcycle (0.151)
	Light truck (0.028)	Light truck (0.074)	Light truck (0.027)

Note: Marginal effect presented with "–" is not significant at 90% confidence interval.

drug will decrease the attentiveness of the driver and eventually increase the probability of making a decision error. The interpretation of the marginal effect for all other variables can be expressed in a similar way as discussed for rural crashes.

Conclusion

More than 90% of crashes that occurred on a roadway segment involve driver error. Driver error can be categorized as recognition, decision, performance, and nonperformance, based on the definition of each error category investigated in NMVCCS. The reasons behind these errors can be complicated, including highway and traffic characteristics, environmental factors, roadway events, driver characteristics, and vehicle types.

This study established a statistical relationship between driver errors with a series of factors including roadway, traffic, and crash data elements. MNP models were applied to quantify the effect of each explanatory variable. The model results suggest that many of the roadway geometry, highway classification, traffic characteristics, roadway event, and driver-related variables are statistically correlated with different driver error categories in both rural and urban areas. Dissimilarities were found by comparing results between rural and urban crashes, which suggests a possible influence of safety culture.

To better understand the impact of various factors contributing to driver error, a review was conducted using marginal effects from the MNP models. The marginal effect of each explanatory variable represents the quantity of increase or decrease in the probability of a specific driver error type. Thus, each error category can be characterized by a combination of unique variables that help to differentiate future safety treatments. These findings provide evidence-based information to support safety professionals in the development of cost-effective engineering countermeasures, safety enforcement, or driver training programs focused on specific driver errors.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: Qin and Shaon; data collection: Shaon; literature review: Shaon, Qin and Chen; Analysis and interpretation of results: Shaon, Qin and Chen; draft manuscript preparation: Shaon, Qin, Chen and Zhang. All authors reviewed the results and approved the final version of the manuscript.

Reference

1. National Highway Traffic Safety Administration. National Motor Vehicle Crash Causation Survey: Report to Congress. *National Highway Traffic Safety Administration Technical Report DOT HS*, Vol. 811, 2008, p. 59.
2. USDOT. *MMUCC Guideline: Model Minimum Uniform Crash Criteria*. US Department of Transportation, Washington, D.C., 2014.
3. Hauer, E. Traffic Safety Toolbox: A Primer on Traffic Safety. *Institute of Transportation Engineers*, Washington, D.C., 1999.
4. Campbell, J. L. *NCHRP Report 600: Human Factors Guidelines for Road Systems*, 2nd ed. Transportation Research Board of the National Academies, Washington, D.C., 2012.
5. Hauer, E. Speed and Safety. *Transportation Research Record: Journal of the Transportation Research Board*, 2009. 2103: 10–17.
6. Tate, F., and S. Turner. Road Geometry and Drivers' Speed Choice. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*, Vol. 16, No. 4, 2007, p. 53.
7. Liu, C., and C.-L. Chen. *An Analysis of Speeding-Related Crashes: Definitions and the Effects of Road Environments*. Report No. HS-811 090. NHTSA, Washington, D.C., 2009.
8. NHTSA. Traffic Safety Facts: Driver Electronic Device Use Observation Protocol. *DOT HS*, Vol. 811, 2010, p. 361.
9. Box, S. New Data from VTTI Provides Insight Into Cell Phone Use and Driving Distraction. *Virginia Tech Transportation Institute*, Blacksburg, VA, 2009. Available at: https://vtnews.vt.edu/articles/2009/07/2009-571.html. Accessed July 15, 2017.
10. Pickrell, T. M., and C. Liu. *Occupant Restraint Use in 2014: Results from the NOPUS Controlled Intersection Study*. Report No. DOT HS 812 244. NHTSA, Washington, D.C., 2016.
11. Juergens, W. *Construction Zone, Detour and Temporary Connection Accidents*. Report No. 1263-38103. Bridge Department, California Division of Highways, Sacramento, CA, 1962.
12. Graham, J. L., R. J. Paulsen, and J. C. Glennon. *Accident and Speed Studies in Construction Zones*. Report No. FHWA-RD-77-80. FHWA, U.S. Department of Transportation, Washington, D.C., 1977.
13. Rouphail, N. M., Z. S. Yang, and J. Fazio. *Comparative Study of Short-And Long-Term Urban Freeway Work Zones. Transportation Research Record: Journal of the Transportation Research Board*, 1988. 1163: 4–14.
14. Blomberg, R. D., R. C. Peck, H. Moskowitz, M. Burns, and D. Fiorentino. Fiorentino. *Crash Risk of Alcohol Involved Driving: A Case-Control Study*. Final Report to the National Highway Traffic Safety Administration, Dunlap and Associates, Inc., Stamford, CT, 2005.
15. NHTSA. 2015 Motor Vehicle Crashes: Overview. *Traffic Safety Facts Research Note*, Vol. 2016, 2016, pp. 1–9.
16. Compton, R., M. Vegega, and D. Smither. *Drug-Impaired Driving: Understanding the Problem and Ways to Reduce It: A Report to Congress*. Report No. DOT HS 811 268. NHTSA, U.S. Department of Transportation, Washington, D.C., 2009.
17. Asbridge, M., J. A. Hayden, and J. L. Cartwright. Acute Cannabis Consumption and Motor Vehicle Collision Risk: Systematic Review of Observational Studies and Meta-Analysis. *Bmj*, Vol. 344, 2012, p. e536.
18. Walsh, J. M., A. G. Verstraete, M. A. Huestis, and J. Mørland. Guidelines for Research on Drugged Driving. *Addiction*, Vol. 103, No. 8, 2008, pp. 1258–1268.
19. Hedlund, J. *Drug Impaired Driving: A Guide for States*. Governors Highway Safety Association (GHSA), Washington, D.C., 2017.
20. Devlin, A., N. Candappa, B. Corben, and D. Logan. Designing Safer Roads to Accommodate Driver Error. *Psychopharmacology*, Vol. 10, 2011, pp. 193–212.
21. Bao, S., and L. N. Boyle. Age-Related Differences in Visual Scanning at Median-Divided Highway Intersections in Rural Areas. *Accident Analysis & Prevention*, Vol. 41, No. 1, 2009, pp. 146–152.
22. Dingus, T. A., F. Guo, S. Lee, J. F. Antin, M. Perez, M. Buchanan-King, and J. Hankey. Driver Crash Risk Factors and Prevalence Evaluation Using Naturalistic Driving Data. *Proceedings of the National Academy of Sciences*, Vol. 113, No. 10, 2016, pp. 2636–2641.
23. Wang, K., and X. Qin. Exploring Driver Error at Intersections: Key Contributors and Solutions. *Transportation Research Record: Journal of the Transportation Research Board*, 2015. 2514: 1–9.
24. Borooah, V. K. *Logit and Probit: Ordered and Multinomial Models*. Thousand Oaks, CA: Sage, 2002.
25. Greene, W. H. *Econometric Analysis*. Prentice Hall, Upper Saddle River, NJ, 2000.
26. Parker, S. T., and Y. Tao. WisTransPortal: A Wisconsin Traffic Operations Data Hub. In K. C. P. Wang, B. L.

Smith, D. R. Uzarski, and S. C. Wong (eds.), *Applications of Advanced Technology in Transportation*. American Society of Civil Engineers, Reston, VA, 2006, pp. 611–616.
27. WisDOT. *Wiscosnin Crash Data User Guide*. http://transportal.cee.wisc.edu/documents/database/crash/crash-data-user-guide.pdf. Accessed June 20, 2017.
28. Long, J. S., and J. Freese. *Regression Models for Categorical Dependent Variables Using Stata*, 2nd ed. Stata press, College Station, TX, 2006.

The Safety Section (ANB00) peer-reviewed this paper (18-03131).