

**Title:** The effect of fatigue driving on injury severity considering the endogeneity

**Authors:** Yanyan Li<sup>a,\*</sup>, Toshiyuki Yamamoto<sup>b</sup>, Guangnan Zhang<sup>c,d</sup>

<sup>a</sup> Department of Civil Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

<sup>b</sup> Institute of Materials and Systems for Sustainability, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan

<sup>c</sup> Institute of Guangdong, Hong Kong and Macao Development Studies, Sun Yat-Sen University, Xingang Xi Road, Guangzhou, China

<sup>d</sup> Center for Studies of Hong Kong, Macao and Pearl River Delta, Sun Yat-Sen University, Xingang Xi Road, Guangzhou, China

## **Abstract**

*Introduction:* Fatigue driving is one of the most risky driving-related behaviors and represented a significant social and economic cost to the community. Several studies have already examined the relationship between fatigue driving behavior and traffic injury severity from different aspects. However, fatigue driving and injury severity in traffic crash may share some common influential factors. Ignoring the impact of these common factors will lead to endogeneity problem and result in biased parameter estimation. *Method:* Based on 38,564 crash records during 2006-2011 in Guangdong province, China, we apply a bivariate endogenous binary-ordered probit model to examine the relationship between fatigue driving and injury severity considering

---

\* Corresponding author at: Department of Civil Engineering, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8603, Japan.  
E-mail address: [lyy901207@gmail.com](mailto:lyy901207@gmail.com) (Yanyan Li), [yamamoto@civil.nagoya-u.ac.jp](mailto:yamamoto@civil.nagoya-u.ac.jp) (Toshiyuki Yamamoto), [sysuzgn@gmail.com](mailto:sysuzgn@gmail.com) (Guangnan Zhang).

endogeneity of fatigue driving. We also explore the difference of influential factors between commercial and non-commercial vehicle drivers. *Results:* This study identifies several common observed influential factors of fatigue driving propensity and fatal injury propensity and reveals a substantial and significant negative correlation of unobserved factors between them. *Conclusions:* The influence of fatigue driving on injury severity is significantly underestimated if the endogeneity of fatigue driving on fatal injury propensity is ignored. Factors such as vehicle insurance and road types not only affect fatal injury propensity, but also fatigue driving propensity. *Practical applications:* The findings in this study can help better understand how those factors affect fatigue driving and injury severity, and contributes to more efficient policy for preventing the harmfulness of fatigue-related crashes.

**Keywords:** Fatigue driving, Endogeneity, Injury severity, Commercial vehicle driver

# 1 **1. Introduction**

2 Road safety has already become a great threat to human being all around the world.  
3 According to *Global Status Report on Road Safety 2015* by World Health Organization  
4 (World Health Organization, WHO, 2015), more than 1.2 million people die each year,  
5 with millions more sustaining serious injuries and living with long-term adverse health  
6 consequences. In low- and middle-income countries, traffic injuries have become one  
7 of the leading causes of death and cost approximately 3% of their GDP as a result of  
8 traffic crashes (WHO, 2015).

9 Fatigue driving was identified as one of the four most risky driving-related  
10 behaviors, especially in fatal traffic crashes (Fernandes et al., 2010) and represented a  
11 significant social and economic cost to the community. Approximately 20% of fatal  
12 crashes in Canada involved driver fatigue, eliminating the influence of alcohol,  
13 speeding and unsafe passing (Canadian Council of Motor Transport Administrator,  
14 CCMTA, 2010). In Australia, 20-30% of all fatal traffic crashes were found due to  
15 fatigue driving (Australian Transport Council, ATC, 2010). However, this situation  
16 could be worse in developing countries since those countries include most of traffic  
17 crashes worldwide (WHO, 2015). A questionnaire-based research among commercial  
18 bus drivers in Malaysia found that the prevalence of fatigue among commercial bus  
19 drivers was 37.7% (Fadhli et al., 2008). Statistics from China also showed that 1,271  
20 (0.83% of total number of crashes due to any cause) crashes were caused by fatigue  
21 driving in 2013, with 677 (1.16% of total number of people killed in the crashes due to

22 any cause) people killed, 1,600 (0.75% of total number of people injured in crashes due  
23 to any cause) people injured, and over RMB 37 million in property losses (Traffic  
24 Management Bureau, Ministry of Public Security, PRC, 2013). China seems to have  
25 lower fatal fatigue-related crash rate than Canada and Australia. The reason for this  
26 contrast may be related to their criterion for calculating the "crash rate". The fatigue  
27 crash rate in the statistics of Canada and Australia is calculated using the number of  
28 crash which "fatigue is one of the contributing factors". However, the fatigue-related  
29 crash rate for China is calculated by the number of crashes which "fatigue is the major  
30 cause of crash". In this case, China is applying a much narrower concept in calculating  
31 fatigue-related crash rate than Canada and Australia. Applying the similar criteria, UK  
32 estimated the fatigue-related crash rate should be around 2% of all crashes in 2015  
33 (Department for Transport, UK, 2016), which the fatigue-related rate is much closer to  
34 China. Although the reported fatigue-related crash rate of China is not so high, we can  
35 still speculate that the crash rate for "fatigue is one of the contributors of crash" would  
36 be much higher.

37 Despite extensive body of research addressing the harmfulness of fatigue driving  
38 on road safety, it has not attracted enough attention. Drivers were less concerned about  
39 fatigued driving than other traffic safety issues (Vanlaar et al., 2008). Studies from  
40 different countries showed that many people still drove when they felt fatigue (Beirness  
41 et al., 2005; Nordbakke & Sagberg, 2007; Tefft, 2010). Besides drivers, public are also  
42 not fully aware of the potential risk of fatigue driving because it is difficult to evaluate  
43 its effect accurately. For example, fatigue could be resolved after a period of rest (Karrer

44 et al., 2004), this feature made it hard to detect and identify after crashes occurred.  
45 When other risky driving behaviors are involved, it is even harder to tell what the major  
46 contributor is and may lead to misclassification of the cause of crash (Horne & Reyner,  
47 1995; Philip et al., 2005; Armstrong et al., 2008). In addition, police also tended to  
48 assign the cause of crash to current interest (Ogden & Moskowitz, 2004).

49 Several studies have examined the relationship between fatigue driving and traffic  
50 injury severity from different aspects. However, fatigue driving and injury severity in  
51 traffic crashes may share some observed common influential factors (e.g. road types).  
52 There are also some unobserved factors between fatigue driving and injury severity.  
53 The connection between sleep disorder, fatigue and traffic injury severity were  
54 discussed by many researchers (Akerstedt et al., 2001; Horne & Reyner, 2001; Philip  
55 et al., 2003; Stutts et al., 2003). Ignoring the impact of these common factors will lead  
56 to endogeneity problem and incorrect conclusion. This study contributes toward current  
57 fatigue driving research by applying a bivariate endogenous binary-ordered probit  
58 model framework to examine the relationship between fatigue driving propensity and  
59 fatal injury propensity in a crash considering the potential endogeneity of fatigue  
60 driving. Considering the potential systematic differences between commercial and non-  
61 commercial vehicle drivers, this model also identifies the observed common factors of  
62 fatigue driving and injury severity for two groups of drivers and makes a comparison.  
63 This result may help better understand how those factors affect fatigue driving  
64 propensity and injury severity, and contributes to more efficient policy for preventing  
65 the harmfulness of fatigue-related crashes. The analysis includes several types of

66 factors, including driver characteristics, vehicle characteristics, road characteristics,  
67 environmental characteristics, and collision characteristics.

68 The rest of this paper is arranged as follows: In section 2, we review related  
69 literature of factors affecting fatigue driving propensity and injury severity in a crash.  
70 We present the methodology in section 3, and data source and sample descriptive  
71 statistics in section 4. Empirical analysis and discussion of estimation results are  
72 presented in section 5. Conclusions and practical applications are provided in section 6.

## 73 **2. Literature Review**

74 Fatigue is a gradual and cumulative process closely related to deterioration of  
75 performance efficiency like driving performance (Haworth, 1998; Rajaratnam &  
76 Arendt ,2001; Philip et al., 2005), and could be induced by repetitive and monotonous  
77 activities like driving long distances (Stutts et al., 1999). Research pointed out that  
78 fatigue was not a strictly monotone decreased progress (Karrer et al., 2004), but an  
79 interaction between deactivation and compensation processes, resulting in variability  
80 of performance (Dinges & Kribbs, 1991).

81 As for the influential factors related to fatigue driving, prior studies basically  
82 focused on four categories: driver characteristics, road characteristics, environmental  
83 characteristics and vehicle characteristics. Considering driver characteristics, male  
84 drivers were at high risk of fatigue driving for the reason that males were more likely  
85 to drive for a longer time (Fernandes et al., 2010; Amstrong et al., 2011). In Amstrong  
86 et al. (2008)'s study, it was found that drivers aged 17-24 years were more likely to be

87 involved in a fatigue-related crash. However, the influence of age is much more  
88 complicated and there exist different behavior patterns between young drivers and elder  
89 drivers. Young drivers frequently committed their fatigue-related offenses during early  
90 morning and night-time hours (Horne & Reyner, 1995; Pack et al., 1995; Maycock,  
91 1996; Horne & Reyner, 2001) while elder drivers mostly in the afternoon (Summala &  
92 Mikkola, 1994). In addition, the motivation for driving while fatigue for young drivers  
93 might be their overestimation of capabilities (Gregersen & Bjurulf, 1996) and  
94 miscalculation of the cost of consequence (Fernandes et al., 2010).

95 For road characteristics and environmental characteristics, driving on different  
96 types of road can lead to similar consequence. Both high-demand and low-demand road  
97 condition could induce driver fatigue (Oron-Gilad et al., 2008; Zhao & Rong, 2013).  
98 Dyani (2007) divided driver fatigue into two groups: passive fatigue and active fatigue.  
99 Passive fatigue was defined closely related to underload, which has been confirmed by  
100 simulated driving studies in monotonous condition (Desmond & Hancock, 2001;  
101 Thiffault & Bergeron, 2003). Active fatigue was defined related to overload of driver.  
102 For example, poor road condition (Arnold et al., 1997), complex traffic conditions and  
103 road environments (Pilcher & Huffcutt, 1996) required more attention and could easily  
104 induce physical and mental fatigue. Time of day was mentioned by several fatigue-  
105 related studies. Folkard (1997) has reviewed several researches that studied the  
106 relationship between road safety and time of day. It was widely believed that time of  
107 day were closely related to human rhythms, which was identified as an important factor  
108 affecting driver fatigue (Haworth, 1998; Philip et al., 2005). Horne and Reyner (2001)

109 found that 02:00-06:00 and 14:00-16:00 is time period associated with higher  
110 probability of fatigue. Haworth (1998) also pointed out that nighttime is significant  
111 contributor of fatigue-related crashes. Light level (Sullivan & Flannagan, 2002) and  
112 season were also identified to play important role (Radun & Radun, 2009).

113         Nevertheless, fatigue-related crashes are severe among commercial vehicle drivers.  
114 Statistics from Europe pointed out that approximately 20% of commercial vehicle  
115 crashes were related to driver fatigue (European Transport Safety Council, ETSC,  
116 2001). The causes of fatigue varied since fatigue could be developed while on the job  
117 with regular sleep patterns or arrived at work already fatigued with irregular sleep  
118 patterns (Young & Hashemi, 1996). Commercial vehicle drivers suffered from sleep  
119 restriction (Hanowski et al., 2007) and were under great work pressure, which made  
120 them vulnerable to fatigue-related crashes. Specifically, drivers in developing countries  
121 are more likely to drive while fatigue for financial reasons (Mock et al., 1999; Nantulya  
122 & Muli-Musiime, 2001). Surveys conducted among truck and taxi drivers in Beijing,  
123 China, showed that driver fatigue was prevalent and the most important reason was  
124 prolonged driving time (Meng et al., 2015).

125         Even though it is not in agreement, fatigue driving and injury severity in the crash  
126 may share some common influential factors, including observed and unobserved factors.  
127 Radun and Radun (2009) claimed that there was no connection between crash severity  
128 and whether the driver was judged to have been fatigued. However, more studies  
129 believed there existed some kind of connection (Haworth, 1998; Zhang et al., 2016).  
130 Fatigue-related crashes were often severe that drivers could not take evasive action

131 under fatigue (Haworth, 1998). Some factors related to fatigue driving may impair  
132 driver performance, then affect injury severity. For example, some unobserved factors  
133 related to the driver's internal state and circadian cycle can also affect both fatigue  
134 propensity and driving performance (Williamson et al. (2011) has given a detail review  
135 on that). Unfortunately, these information was almost impossible to collect due to  
136 traumatic effects and emotional state change after the crash (Radun & Radun, 2009).  
137 Some drivers might not admit fatigue or falling asleep during driving concerning about  
138 insurance and legal consequences (Corfitsen, 1999). Therefore, those common factors  
139 were often neglected, which may lead to endogeneity problem and biased estimation  
140 when analyzing the relationship between fatigue driving and injury severity.

### 141 **3. Econometric Framework**

#### 142 **3.1. Model structure**

143 In fatigue-related crashes, drivers who are more likely to be involved in fatigue-  
144 related crashes and injury severity can be correlated, which may cause endogenous  
145 problem. In econometrics, endogeneity problem is said to occur if the independent  
146 variable is correlated with the error term. This correlation can be caused by several  
147 reasons: omitted variables, measurement error, and simultaneity in simultaneous  
148 equations models. Endogeneity induces estimation bias in statistical models and may  
149 eventually lead to mistaken conclusions. To take into account the potential endogeneity  
150 of fatigue driving, we apply a bivariate endogenous binary-ordered probit model in the  
151 current paper. Bivariate endogenous binary-ordered probit model is a hierarchical

152 model system of two equations that can be used to model two response variables  
 153 simultaneously, and addresses endogeneity problem. This model addresses endogeneity  
 154 by considering error correlations among two equations that capture the relationships  
 155 among endogenous variable, exogenous variables and error term (for further discussion,  
 156 see Greene, 2007; Fernandez-Antolin et al., 2014).

157 Let  $i$  ( $i=1,2,\dots,N$ ) be an index representing drivers and  $k$  ( $k=1,2,\dots,K$ ) be  
 158 indices representing ordinal categories of injury severity sustained by driver  $i$  in the  
 159 crash. Suppose  $y_i$  is the observed injury severity level and  $y_i^*$  represents latent  
 160 injury severity propensity of driver  $i$  in the crash. Thus, the latent propensity  $y_i^*$  is  
 161 mapped to the actual injury severity level  $y_i$  by threshold  $\psi_k$   
 162 ( $\psi_0 = -\infty$  and  $\psi_K = \infty$ ,  $\psi_0 < \psi_1 < \dots < \psi_K$ ) as the following equations:

$$163 \quad y_i^* = \alpha' x_i + \theta \text{fatig}_i + v_i, \quad (1)$$

$$164 \quad y_i = k, \text{ if } \psi_{k-1} \leq y_i^* < \psi_k \quad (2)$$

165 where  $x_i$  is an  $M \times 1$  column vector of variables that influences  $y_i^*$  (not including a  
 166 constant) and  $\text{fatig}_i$  is a dummy variable indicating whether driver  $i$  is convicted as  
 167 fatigue driving or not.  $\alpha$  represents an  $M \times 1$  coefficient vector of  $x_i$  and  $\theta$  is the  
 168 coefficient of  $\text{fatig}_i$ .  $v_i$  is the error term assumed to be identically and independently  
 169 across driver  $i$ .

170 However,  $\text{fatig}_i$  included in Eq. (1) may be endogenous. Therefore, we specify  
 171 here:

$$172 \quad \text{fatig}_i^* = \beta' z_i + \omega_i, \quad (3)$$

$$173 \quad \text{fatig}_i = \begin{cases} 1, & \text{if } \text{fatig}_i^* \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

174 This equation represents the latent fatigue driving propensity  $fatig_i^*$  of driver  $i$ .  
175  $fatig_i$  is the actual observed fatigue driving behavior by driver  $i$ , and  $z_i$  is an  $L \times 1$   
176 column vector of independent variables (including a constant) influencing fatigue  
177 driving propensity  $fatig_i^*$ .  $\beta$  is an  $L \times 1$  coefficient vector of  $z_i$ .  $x_i$  and  $z_i$  can  
178 share some common variables, representing the common observed influential factors  
179 between fatigue driving propensity and fatal injured propensity.  $\omega_i$  represents the  
180 random components that capture all unobserved factors.

181 Still, there could be unobserved correlation between injury severity and fatigue  
182 driving. To capture the unobserved correlation, without losing generality we assume  
183 that  $v_i$  and  $\omega_i$  form a bivariate normal distribution. In particular, the probability is  
184 given as:

$$\begin{aligned}
& \text{Prob}(y_i = k, fatig_i = j | x_i, z_i) \\
185 &= \text{Prob}(\psi_{k-1} \leq y_i^* < \psi_k, \tau_{j-1} \leq fatig_i^* < \tau_j) \\
&= \left( \begin{array}{l} \Phi_2(\psi_k - (\alpha'x_i + \theta fatig_i), \tau_j - \beta'z_i; \rho) \\ - \Phi_2(\psi_{k-1} - (\alpha'x_i + \theta fatig_i), \tau_j - \beta'z_i; \rho) \\ - \Phi_2(\psi_k - (\alpha'x_i + \theta fatig_i), \tau_{j-1} - \beta'z_i; \rho) \\ + \Phi_2(\psi_{k-1} - (\alpha'x_i + \theta fatig_i), \tau_{j-1} - \beta'z_i; \rho) \end{array} \right) \quad (5)
\end{aligned}$$

186 where  $\Phi_2(\cdot)$  is the standard bivariate normal cumulative distribution function.  $\tau_j$   
187 and  $\tau_{j-1}$  ( $j=0,1$ ) represent thresholds for mapping the latent variable  $fatig_i^*$  to  
188 the observed variable  $fatig_i$  in Eq. (4). Specifically, in the binary probit model we set  
189  $\tau_{-1} = -\infty$ ,  $\tau_0 = 0$ , and  $\tau_1 = \infty$ .  $\rho$  measures the correlation between disturbances in  
190 the equations, which measures correlation between injury severity and fatigue driving  
191 propensity after the influence of fatigue is accounted in injury severity function. If this  
192 correlation between fatigue driving propensity and fatal injury propensity is ignored

193 when actually exists, it could lead to inconsistent estimation of the effect of fatigue on  
 194 injury severity. We also introduce a univariate endogenous binary-ordered probit model  
 195 in which we assume  $\rho=0$ , neglecting the correlation between  $v_i$  and  $w_i$  for  
 196 comparison purpose.

### 197 3.2. Model estimation

198 The log-likelihood function is given by:

$$199 \quad LL = \sum_{i=1}^N \ln \left( \begin{array}{l} \Phi_2(\psi_k - (\alpha'x_i + \theta fatig_i), \tau_j - (\beta'z_i); \rho) \\ - \Phi_2(\psi_{k-1} - (\alpha'x_i + \theta fatig_i), \tau_j - (\beta'z_i); \rho) \\ - \Phi_2(\psi_k - (\alpha'x_i + \theta fatig_i), \tau_{j-1} - (\beta'z_i); \rho) \\ + \Phi_2(\psi_{k-1} - (\alpha'x_i + \theta fatig_i), \tau_{j-1} - (\beta'z_i); \rho) \end{array} \right) \quad (6)$$

200 The corresponding parameters  $\alpha$ ,  $\beta$ ,  $\theta$ ,  $\psi_k$ , and  $\rho$  are estimated  
 201 simultaneously by maximizing the log-likelihood function of Equation (6). R software  
 202 (version 3.3.1) is used for estimation in this study.

## 203 4. Data

### 204 4.1. Data source

205 The Guangdong Traffic Accident Dataset (GTAD) is sourced from the Traffic  
 206 Management Sector Specific Incident Case Data Report, the Road Traffic Accident  
 207 Database of China's Public Security Department. A total of 38,564 crash records during  
 208 2006-2011 are applied in this study. The data we used in this study were drawn from  
 209 police-reported crashes in 21 cities across Guangdong Province, and compiled from a

210 sample of crashes that involve at least one motor vehicle and resulting in property  
211 damage, injury, or death.

212 Several crash-related attributes are collected for each record in GTAD, including  
213 driver characteristics, vehicle characteristics, road characteristics, environmental  
214 characteristics, and crash characteristics. The injury severity of each individual  
215 involved in the crash is categorized into four ordinal levels: (1) No injury, (2) Minor  
216 injury, (3) Serious injury, and (4) Fatal injury.

217 This study mainly focuses on drivers who were chiefly responsible for the  
218 occurrence of crash that was convicted to be fatigue-related. The reason is that crash  
219 and personal information is better recorded. The definition for fatigue driving in GTAD  
220 is defined as fulfilling one of the following conditions: (a) Driving cars more than eight  
221 hours a day, (b) Engaging in other work with excessive physical exertion, (c) Lacking  
222 of sleep which results in sleepy or weakness of limbs, so that the driver is having  
223 difficulty in assessing traffic conditions immediately and reacting accurately. Normally,  
224 the police officer would interview the involved parties and witnesses, and check the  
225 driving records to identify the cause of crash. Technical reconstruction is also helpful  
226 for determining the cause of crash by studying testimony of witnesses and physical  
227 evidence, especially in serious crashes. Fatigue-related crashes defined by this  
228 definition constitute 6.5% of all crashes in GTAD dataset. The distribution of fatigue  
229 driving and injury severity across observations is presented in Table 1. Overall, the  
230 descriptive statistics in Table 1 indicate a substantially higher percentage of fatal  
231 fatigue-related crashes (13.2%) than non-fatigue-related crashes (6.5%).

233 **Table 1: Number of Fatigue Related Crashes by Injury Level**

Injury Severity	Fatigue Driving		All (%)
	No (%)	Yes (%)	
<b>All</b>			
No Injury	25070 (65.5)	173 (56.9)	25243 (65.5)
Minor Injury	8156 (21.3)	64 (21.0)	8220 (21.3)
Serious Injury	2559 (6.7)	27 (8.9)	2586 (6.7)
Fatal Injury	2475 (6.5)	40 (13.2)	2515 (6.5)
Total	38260 (100)	304 (100)	38564 (100)
<b>Commercial</b>			
No Injury	8297 (86.0)	96 (57.5)	8393 (85.5)
Minor Injury	725 (7.5)	34 (20.3)	759 (7.7)
Serious Injury	227 (2.4)	15 (9.0)	242 (2.5)
Fatal Injury	400 (4.1)	22 (13.2)	422 (4.3)
Total	9650 (100)	167 (100)	9817 (100)
<b>Non-commercial</b>			
No Injury	16773 (58.6)	77 (56.2)	16850 (58.6)
Minor Injury	7431 (26.0)	30 (21.9)	7461 (25.9)
Serious Injury	2332 (8.1)	12 (8.8)	2344 (8.2)
Fatal Injury	2075 (7.3)	18 (13.1)	2093 (7.3)
Total	28610 (100)	137 (100)	28747 (100)

## 235 4.2. Variables

236 Five types of variables were considered in the empirical analysis. *Driver*  
237 *characteristics* include: driver's age ( $\leq 25$ , 26-35, 36-45, 46-55, 56-65, and  $\geq 66$  years  
238 old), driver's gender, driving experience ( $\leq 2$  years), and whether the driver has a valid  
239 driving license. *Vehicle characteristics* include: whether the vehicle has insurance.  
240 Other vehicle characteristics, such as vehicle speed just before collision, could not be  
241 included because of the absence of data in the GTDA. *Road characteristics* include:  
242 road type (whether the crash occurred on express way or urban roads), isolated lanes  
243 (whether the road has separated lanes for motorized and non-motorized vehicles), and  
244 terrain (mountain area). *Environmental characteristics* include: time of day  
245 represented in three categories (early morning (00:00-06:59), morning peak hours  
246 (07:00-08:59), and afternoon peak hours (17:00-19:59)) and lighting conditions (dark  
247 with street lights and dark without street lights). *Crash characteristics* include:  
248 collision type (head-on collision, rear-end collision, sideways collision). Variable  
249 description is presented in Table 2.

250 Firstly, a general model including all the variables suggested by prior studies and  
251 intuitiveness considerations are applied. Then, variables are chosen based on a  
252 systematic process of removing statistically insignificant variables and combining  
253 variables when their effects were not significantly different. Furthermore, continuous  
254 variables, such as driver' age and time of day, were converted into dummy variables  
255 and different ranges are also tested.

256

**Table 2: Variable Description**

<b>Variables</b>	<b>Description</b>	<b>Mean</b>
<i><b>Driver characteristics</b></i>		
Driver's gender		
Male	Male=1 ; Female=0	0.947
Driver's age		
≤25	≤ 25=1; Others=0	0.209
26-35	26-35=1; Others=0	0.355
36-45	36-45=1; Others=0	0.296
46-55	46-55=1; Others=0	0.106
56-65	56-65=1; Others=0	0.028
≥ 66	≥ 66=1; Others=0	0.006
Driving experience		
≤2 years	≤2 years=1; Others=0	0.135
Driving license		
Not valid	Not valid=1; Others=0	0.280
<i><b>Vehicle characteristics</b></i>		
Insurance		
Yes	Insurance=1; Others=0	0.779
<i><b>Road characteristics</b></i>		
Road type		
Express way	Express way=1; Others=0	0.046
Urban road	Urban road=1; Others=0	0.409
Isolated lanes		
Yes	Isolated lanes=1; Others=0	0.397
Terrain		
Mountain	Mountain=1; Others=0	0.600
<i><b>Environmental characteristics</b></i>		
Lighting condition		
Dark with street light	Dark with street light=1; Others=0	0.261
Dark without street light	Dark without street light=1; Others=0	0.173
Time of day		
00:00-06:59	00:00-06:59=1; Others=0	0.155
07:00-08:59	07:00-08:59=1; Others=0	0.087
17:00-19:59	17:00-19:59=1; Others=0	0.171
<i><b>Crash characteristics</b></i>		
Head-on	Head-on collision=1; Others=0	0.224
Sideway	Sideway collision=1; Others=0	0.421
Rear-end	Rear-end collision=1; Others=0	0.119

258 **5. Estimation Results and Discussion**

259 In this study, two different models were estimated: (1) Bivariate Endogenous  
260 Binary-Ordered Probit model, and (2) Univariate Endogenous Binary-Ordered Probit  
261 model (by assuming  $\rho = 0$  as noted earlier). There could be potential systematic  
262 differences between commercial and non-commercial vehicle drivers (i.e. driving skill,  
263 driving time), we compare the factors associated with fatigue driving propensity and  
264 fatal injury propensity between them. Variables considered in the models at the very  
265 beginning of fatigue and injury severity function are listed in Table 3, and all the  
266 variables that were included in fatigue function also being included in injury severity  
267 function. In Table 4 and Table 5, we present the results of both models for commercial  
268 and non-commercial vehicle drivers, and only significant variables (at 95% significant  
269 level) will be listed and discussed in the following parts.  
270

271 **Table 3: Variable Selection for Fatigue Model and Injury Severity Model**

Variable	Fatigue		Injury severity	
	Commercial	Non-commercial	Commercial	Non-commercial
<b><i>Driver characteristics</i></b>				
Driver's gender	√	√	√	√
Driver's age	√	√	√	√
Driving experience	√	√	√	√
Driving license	√	√	√	√
<b><i>Vehicle characteristics</i></b>				
Vehicle type	√	√	√	√
Insurance	√	√	√	√
<b><i>Road characteristics</i></b>				
Road type	√	√	√	√
Isolated lanes	√	√	√	√
Terrain	√	√	√	√
<b><i>Environmental characteristics</i></b>				
Lighting condition	√	√	√	√
Time of day	√	√	√	√
<b><i>Crash characteristics</i></b>				
Collision type			√	√

272

273 **5.1. Measures of fit**

274 Before discussing the estimation results, likelihood ratio test is conducted to  
 275 compare bivariate and univariate model. The test statistic is given as

276 
$$LR = -2 \times (llk_{nc} - llk_c) \quad (7)$$

277 where  $llk_{nc}$  is the log-likelihood at convergence of bivariate model, and  $llk_c$  is the  
 278 log-likelihood at convergence of the models estimated on univariate model. The  $LR$   
 279 statistic for commercial and non-commercial vehicle drivers is 4.38 and 3.66, which  
 280 reject the null hypothesis of  $\rho = 0$  at  $p < 0.05$  and  $p < 0.1$ , respectively. It should  
 281 be noted that, in this case,  $\rho$  is conservatively retained in non-commercial vehicle

282 driver sample since the correlation does significantly change the coefficient of fatigue  
283 in the model. This result indicates that correlation due to unobserved factors between  
284 injury severity and fatigue driving propensity is significant, and model estimation  
285 without considering these correlation may result in inefficient parameter estimates  
286 (Yamamoto & Shankar, 2004).

287 We also conduct information criteria to compare model performance. Both of the  
288 value of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC)  
289 decline for commercial and non-commercial vehicle drivers by including the  
290 correlation, which also suggests that the proposed model is more efficient.

## 291 **5.2. Estimation results**

292 The results in Table 4 and Table 5 indicate that fatigue has significant impacts on  
293 fatal injury propensity for both commercial and non-commercial vehicle drivers. The  
294 coefficient of fatigue in the bivariate endogenous model is larger than in univariate  
295 model. The estimated coefficient of fatigue driving among commercial drivers is 0.984  
296 in bivariate model and 0.291 in univariate model. The impact of fatigue driving on fatal  
297 injury propensity is underestimated by 0.693 in the latter. For non-commercial vehicle  
298 drivers, the coefficient of fatigue driving in bivariate model is 0.895 while in univariate  
299 model is 0.234 and the gap is 0.661. The impact of fatigue driving on injury severity in  
300 a crash on both groups are underestimated and the gaps between these two groups are  
301 similar. Larger coefficient of fatigue driving indicates higher risk of involving in fatal  
302 injured crash. Results suggest that commercial vehicle drivers are somewhat more risky

303 when driving under fatigue than non-commercial vehicle drivers. Commercial vehicle  
304 drivers often drove a high number of miles (National Sleep Foundation, 2009), and  
305 some of them tend to break the rules about duty and rest hours for pursuing more profit  
306 (Radun & Radun, 2009). Thus, they are more likely to lose focus or even fall asleep at  
307 the wheel, which may lead to severe crashes.

308 This study also identifies the observed common factors of fatigue driving  
309 propensity and fatal injury propensity. In summary, the observed common factors for  
310 commercial vehicle drivers are: insurance, road types, and terrain. For non-commercial  
311 vehicle drivers, the observed common factors are: insurance and road types. More detail  
312 discussions of estimation results by groups are as following:

313 *Driver characteristics:* although gender, age and driving experience do not show  
314 significant impact on fatigue driving propensity, they do influence the driver's  
315 propensity of fatal injury in the crash. Non-commercial vehicle drivers who is over 45  
316 years old are more likely to be fatal injured in the crash while male, young non-  
317 commercial vehicle drivers are found less likely. This result is consistent with previous  
318 findings (Hatfield et al., 2005; McConnell, 2003). Less experienced drivers ( $\leq 2$  years)  
319 are found more likely to be more severe injured than those with more driving  
320 experienced. Less experienced drivers with dynamic driving style are with more risk in  
321 the monotonous setting than experienced and calm drivers (Stutts et al., 2003; Karrer  
322 et al., 2004). However, these effects are not significant for commercial vehicle drivers  
323 due to small variation in commercial driver group. Drivers without a valid driving  
324 license are significantly more likely to be severe injured for both commercial and non-

325 commercial vehicle drivers.

326 *Vehicle characteristics:* the impact of factors associated with vehicle itself shows  
327 contrast effects on fatigue driving propensity and fatal injury propensity. For both  
328 groups of drivers, driving vehicles with insurance are less likely to be fatally injured in  
329 the crash. Insurance lowers the monetary loss of crashes. Nevertheless, monetary  
330 compensation can never compensate for losing one's life. Therefore, drivers will pay  
331 enough attention to preventing themselves from fatal crashes. On the other hand, they  
332 might let their defenses down under the circumstances which they thought to be not  
333 serious. Thus, light-injured crashes are more likely to happen. In addition, non-  
334 commercial vehicle with insurance presents higher risk of fatigue driving while the  
335 impact for commercial vehicle is not significant. This finding may be related to  
336 different penalties for commercial drivers and non-commercial vehicle drivers when  
337 conducting fatigue driving. According to the Road Traffic Safety Law of the People's  
338 Republic of China and local traffic regulations, commercial vehicle drivers will have  
339 their driving licenses endorsed with at least six penalty points even lose their driving  
340 licenses once caught fatigue driving. However, for non-commercial vehicle drivers,  
341 fatigue driving will only incur traffic tickets without losing any points on their driving  
342 license.

343 *Road characteristics:* driving on express way is at high risk of fatigue driving and  
344 fatal injury for commercial and non-commercial vehicle drivers. Express ways are  
345 mostly monotonous and of high speed. Driving on them can be regarded as a repetitive  
346 activity which requires sustained attention and can easily lead to fatigue (McCartt et al.,

347 2000; Thiffault & Bergeron, 2003). On the contrast, driving on urban road or mountain  
348 area is less likely to fatigue driving as well as sustain fatal injured. The lower propensity  
349 of fatigue driving may be the result of high rate of environmental stimulation and  
350 continuous changes in the driving scenery (Mavjee & Horne, 1994; Horne & Reyner,  
351 1999), which help to maintain driver's attention persistently. The impact of driving in  
352 mountain area on fatigue driving propensity is not significant for non-commercial  
353 vehicle drivers. Isolated lanes show no significant impact on fatigue driving propensity,  
354 however, its impact on injury severity differs between commercial and non-commercial  
355 vehicle drivers. For commercial vehicle drivers, driving on isolated lanes is less likely  
356 to sustain fatal injury, but is more likely for non-commercial vehicle drivers.

357 *Environmental characteristics:* The effect of time period on fatigue driving  
358 propensity also shows different patterns. During midnight to early morning (00:00-  
359 06:59), both commercial and non-commercial vehicle drivers are more likely to fatigue  
360 driving compared to other time period in a day. For example, 75% of fatigue-related  
361 crashes occurred between 02:00 and 08:00 in 107 heavy truck crashes reviewed  
362 (National Transportation Safety Board, U.S., 1995). Morning peak hour (07:00-08:59)  
363 only affects the fatigue driving propensity of commercial drivers. It is still not clear  
364 whether this result is due to sleep loss or other reasons. Driving at night significantly  
365 contributes to more severe injury crashes for both commercial and non-commercial  
366 vehicle drivers, but the propensity of fatal injured crashes declines following the  
367 installation of street lights. This finding is also consistent with several previous studies  
368 (Elvik, 1995; Owens & Sivak, 1996; Plainis et al., 2006).

369 *Crash characteristics:* since collision type does not affect fatigue driving behavior,  
370 this variable is only considered in injury severity function. The result indicates that  
371 commercial vehicle drivers are more likely to fatal injured when involved in rear-end  
372 collision while sideway collision and head-on collision is less likely to fatal injured.  
373 Some commercial vehicles have larger size and are heavier than the other passenger  
374 vehicles with which they share the roads, and the stopping distance for them is much  
375 longer. Thus, large and heavy commercial vehicles involving in rear-end crashes may  
376 be due to their inability to stop immediately, and cause severe injuries. For sideway and  
377 head-on collision, commercial vehicle drivers can reduce the harmfulness of collision  
378 by taking sudden turns or other protecting behaviors based on their experience.  
379 However, for non-commercial vehicle drivers, both head-on and rear-end collision have  
380 higher propensity of fatal injury, which may be related to lacking experience in  
381 handling emergencies on road compared to commercial vehicle drivers. The impact of  
382 side collision is not significant for non-commercial vehicle drivers.

383  
384

**Table 4: Estimation Result of Commercial Vehicle Driver Sample**

Variables	Correlated		Uncorrelated	
	Coef.	SE	Coef.	SE
<i>Fatigue Driving Propensity</i>				
Road type				
Express way	0.368***	0.080	0.400***	0.079
Urban road	-0.718***	0.152	-0.702***	0.151
Terrain				
Mountain	-0.179**	0.072	-0.178**	0.072
Time of day				
00:00-06:59	0.771***	0.079	0.740***	0.079
07:00-08:59	0.466***	0.114	0.457***	0.114
Intercept	-2.379***	0.068	-2.381***	0.068
<i>Injury Severity Propensity</i>				

Fatigue	0.984 <sup>***</sup>	0.312	0.291 <sup>***</sup>	0.097
Driving license				
Not valid	0.291 <sup>***</sup>	0.064	0.293 <sup>***</sup>	0.065
Insurance				
Yes	-0.192 <sup>***</sup>	0.064	-0.193 <sup>***</sup>	0.064
Road type				
Express way	0.850 <sup>***</sup>	0.056	0.883 <sup>***</sup>	0.054
Urban road	-0.129 <sup>***</sup>	0.044	-0.133 <sup>***</sup>	0.044
Isolated lanes				
Yes	-0.085 <sup>**</sup>	0.040	-0.084 <sup>**</sup>	0.040
Terrain				
Mountain	-0.217 <sup>***</sup>	0.034	-0.223 <sup>***</sup>	0.034
Lighting condition				
Dark with street light	0.131 <sup>***</sup>	0.048	0.136 <sup>***</sup>	0.048
Dark without street light	0.243 <sup>***</sup>	0.040	0.252 <sup>***</sup>	0.040
Collision type				
Head-on	-0.142 <sup>***</sup>	0.050	-0.140 <sup>***</sup>	0.051
Side	-0.377 <sup>***</sup>	0.047	-0.376 <sup>***</sup>	0.047
Rear-end	0.276 <sup>***</sup>	0.044	0.280 <sup>***</sup>	0.044
$\rho$	-0.313 <sup>**</sup>	0.132		
<i>Cut1</i>	0.925 <sup>***</sup>	0.071	0.921 <sup>***</sup>	0.071
<i>Cut2</i>	1.450 <sup>***</sup>	0.072	1.449 <sup>***</sup>	0.072
<i>Cut3</i>	1.719 <sup>***</sup>	0.074	1.720 <sup>***</sup>	0.074
<i>log-likelihood</i>	-5491		-5493	
<i>AIC</i>	11025		11028	
<i>BIC</i>	11057		11059	
<i>N</i>	9816		9816	

385 \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

386

387 **Table 5: Estimation Result of Non-commercial Vehicle Driver Sample**

Variables	Correlated		Uncorrelated	
	Coef.	SE	Coef.	SE
<i>Fatigue Driving Propensity</i>				
Insurance				
Yes	0.351 <sup>***</sup>	0.090	0.351 <sup>***</sup>	0.089
Road type				
Express way	0.595 <sup>***</sup>	0.108	0.604 <sup>***</sup>	0.108
Urban road	-0.267 <sup>***</sup>	0.068	-0.268 <sup>***</sup>	0.068
Time of day				
00:00-06:59	0.419 <sup>***</sup>	0.071	0.384 <sup>***</sup>	0.070

Intercept	-2.896 <sup>***</sup>	0.086	-2.891 <sup>***</sup>	0.086
<i>Injury Severity Propensity</i>				
Fatigue	0.895 <sup>**</sup>	0.364	0.234 <sup>**</sup>	0.101
Driver's gender				
Male	-0.200 <sup>***</sup>	0.028	-0.199 <sup>***</sup>	0.028
Driver's age				
26-35	-0.170 <sup>***</sup>	0.016	-0.170 <sup>***</sup>	0.016
46-55	0.238 <sup>***</sup>	0.023	0.238 <sup>***</sup>	0.023
56-65	0.500 <sup>***</sup>	0.037	0.500 <sup>***</sup>	0.037
$\geq 65$	0.606 <sup>***</sup>	0.075	0.606 <sup>***</sup>	0.075
Driving experience				
$\leq 2$ years	0.153 <sup>***</sup>	0.022	0.153 <sup>***</sup>	0.022
Driving license				
Not valid	0.703 <sup>***</sup>	0.017	0.703 <sup>***</sup>	0.017
Insurance				
Yes	-0.281 <sup>***</sup>	0.017	-0.280 <sup>***</sup>	0.017
Road type				
Express way	0.370 <sup>***</sup>	0.050	0.388 <sup>***</sup>	0.050
Urban road	-0.250 <sup>***</sup>	0.015	-0.251 <sup>***</sup>	0.015
Isolated lanes				
Yes	0.060 <sup>***</sup>	0.016	0.060 <sup>***</sup>	0.016
Terrain				
Mountain	-0.135 <sup>***</sup>	0.015	-0.135 <sup>***</sup>	0.015
Lighting condition				
Dark with street light	0.076 <sup>***</sup>	0.017	0.077 <sup>***</sup>	0.017
Dark without street light	0.181 <sup>***</sup>	0.020	0.182 <sup>***</sup>	0.020
Collision type				
Head-on	0.140 <sup>***</sup>	0.017	0.140 <sup>***</sup>	0.017
Rear-end	0.180 <sup>***</sup>	0.024	0.180 <sup>***</sup>	0.024
$\rho$	-0.234 <sup>*</sup>	0.122		
<i>Cut1</i>	0.053	0.036	0.052	0.036
<i>Cut2</i>	0.960 <sup>***</sup>	0.036	0.960 <sup>***</sup>	0.036
<i>Cut3</i>	1.442 <sup>***</sup>	0.037	1.442 <sup>***</sup>	0.037
<i>log-likelihood</i>	-28726		-28728	
<i>AIC</i>	57504		57506	
<i>BIC</i>	57541		57541	
<i>N</i>	28748		28748	

388 \*  $p < 0.1$ , \*\*  $p < 0.05$ , \*\*\*  $p < 0.01$

### 389 5.3. Marginal effects

390 The coefficient estimates do not provide the magnitude of impacts on probability  
391 in each injury level. Thus, we calculate the marginal effect of variables, which directly  
392 influence fatal injury propensity, on each injury level for both commercial and non-  
393 commercial vehicle drivers. Considering all variables in this model are dummy  
394 variables, we compute probabilities by setting the variable to one and then zero, and  
395 take the difference. That is,

$$396 \quad ME = \text{Prob}(y_i = k | m_i = 1) - \text{Prob}(y_i = k | m_i = 0) \quad (7)$$

397 where  $ME$  is the marginal effect of dummy  $m_i$  on injury level  $k$ . The marginal  
398 effect can be interpreted as the change of probability due to the change in variable from  
399 zero to one. The results of bivariate and univariate models for commercial and non-  
400 commercial vehicle drivers are presented in Table 6 and Table 7. (These tables only  
401 present the marginal effect for significant variables identified in earlier discussion).

402 Some important features should be addressed here. First, this study shows that the  
403 marginal effect of fatigue driving on fatal injury for commercial vehicle driver and non-  
404 commercial driver is 12.9% and 17.9%. That is, the occurrence of fatigue driving will  
405 increase the probability of fatal injury in a crash by 12.9% for commercial vehicle driver  
406 and 17.9% for non-commercial vehicle drivers. Moreover, comparing marginal effect  
407 in bivariate and univariate model, we find that the estimated impact of fatigue are much  
408 lower without considering correlation. Ignoring the correlation of unobserved factors  
409 may lead to underestimation of the harmfulness of fatigue driving behavior. In our study,  
410 the harmfulness of fatigue is underestimated by 10.5% and 14.6% for commercial and

411 non-commercial vehicle driver, respectively. Second, for commercial vehicle drivers,  
412 other major risk factors of fatal injury include express way, not valid driving license,  
413 and rear-end collision. For non-commercial vehicle, elder driver (aged  $\geq 66$  years old  
414 and 56-65 years old), not valid driving license, and express way are the most significant  
415 contributors. Third, side collision, driving in mountain area, and insurance are  
416 recognized as the three most influencing factors for commercial vehicle drivers to  
417 survive in a crash while for non-commercial vehicle drivers are insurance, urban road,  
418 and male.

419 **Table 6: Marginal Effect for Commercial Vehicle Driver Sample**

	<b>Bivariate Binary-Ordered Probit</b>				<b>Univariate Binary-Ordered Probit</b>			
	<b>Y = 1</b>	<b>Y = 2</b>	<b>Y = 3</b>	<b>Y = 4</b>	<b>Y = 1</b>	<b>Y = 2</b>	<b>Y = 3</b>	<b>Y = 4</b>
Fatigue Driving	-0.256	0.086	0.041	0.129	-0.060	0.025	0.010	0.024
Driving License								
Not valid	-0.058	0.025	0.010	0.023	-0.060	0.025	0.010	0.024
Insurance								
Yes	0.037	-0.016	-0.006	-0.014	0.038	-0.016	-0.006	-0.015
Road type								
Express way	-0.225	0.084	0.037	0.087	-0.237	0.087	0.039	0.094
Urban road	0.037	-0.009	-0.004	-0.008	0.036	-0.009	-0.003	-0.007
Isolated lanes								
Yes	0.015	-0.007	-0.003	-0.006	0.015	-0.007	-0.003	-0.006
Terrain								
Mountain	0.054	-0.022	-0.008	-0.018	0.046	-0.018	-0.007	-0.015
Lighting condition								
Dark with street light	-0.024	0.011	0.004	0.009	-0.026	0.011	0.004	0.010
Dark without street light	-0.046	0.021	0.008	0.017	-0.049	0.022	0.009	0.019
Collision type								
Head-on	0.024	-0.011	-0.004	-0.009	0.024	-0.011	-0.004	-0.009
Sideway	0.063	-0.031	-0.011	-0.022	0.065	-0.031	-0.011	-0.022
Rear-end	-0.054	0.024	0.009	0.020	-0.056	0.025	0.010	0.021

420

421 **Table 7: Marginal Effect for Non-commercial Vehicle Driver Sample**

	<b>Bivariate Binary-Ordered Probit</b>				<b>Univariate Binary-Ordered Probit</b>			
	<b>Y = 1</b>	<b>Y = 2</b>	<b>Y = 3</b>	<b>Y = 4</b>	<b>Y = 1</b>	<b>Y = 2</b>	<b>Y = 3</b>	<b>Y = 4</b>
Fatigue Driving	-0.303	0.051	0.073	0.179	-0.081	0.028	0.020	0.033
Driver's gender								
Male	0.069	-0.025	-0.017	-0.027	0.069	-0.025	-0.017	-0.027
Driver's age								
26-35	0.058	-0.024	-0.014	-0.020	0.058	-0.024	-0.014	-0.020
46-55	-0.083	0.030	0.021	0.032	-0.083	0.030	0.021	0.032
56-65	-0.175	0.051	0.044	0.080	-0.175	0.051	0.044	0.080
≥65	-0.211	0.054	0.053	0.105	-0.211	0.054	0.053	0.105
Driving experience								
Less than 2 years	-0.053	0.020	0.013	0.020	-0.052	0.020	0.013	0.020
Driving License								
Not valid	-0.259	0.098	0.067	0.093	-0.258	0.098	0.067	0.093
Insurance								
Yes	0.097	-0.039	-0.025	-0.036	0.097	-0.039	-0.025	-0.037
Road type								
Express way	-0.140	0.038	0.031	0.055	-0.143	0.038	0.032	0.057
Urban road	0.089	-0.035	-0.021	-0.029	0.089	-0.035	-0.021	-0.029
Isolated lanes								
Yes	-0.021	0.008	0.005	0.007	-0.020	0.008	0.005	0.007
Terrain								
Mountain	0.046	-0.018	-0.011	-0.017	0.046	-0.018	-0.011	-0.017
Lighting condition								

Dark with street light	-0.026	0.010	0.006	0.009	-0.026	0.010	0.006	0.010
Dark without street light	-0.063	0.024	0.016	0.024	-0.063	0.024	0.016	0.024
Collision type								
Head-on	-0.048	0.019	0.012	0.018	-0.049	0.019	0.012	0.018
Rear-end	-0.062	0.023	0.015	0.024	-0.063	0.023	0.016	0.024

---

422

## 423 **6. Conclusions and practical applications**

424 Several studies have examined the relationship between driver fatigue and traffic  
425 injury severity from different aspects. However, some of factors that affect driver's  
426 fatigue propensity also have influence on driver's injury severity in a crash, including  
427 observed and unobserved factors. Ignoring the impact of these common factors will  
428 lead to endogeneity problem and incorrect conclusion. Based on 38,564 crash records  
429 during 2006-2011, we conduct an empirical analysis to examine the relationship  
430 between fatigue driving propensity and fatal injury severity by comparing bivariate and  
431 univariate endogenous binary-ordered probit model. Five types of factors are included.  
432 It is essential to quantify the impact of these characteristics on injury severity by  
433 calculating marginal effect, so that measures to prevent or reduce harmfulness of fatigue  
434 driving can be identified and implemented.

435 The result reveals a substantial and significant negative error correlation between  
436 fatigue driving propensity and fatal injury propensity, which lends strong support for  
437 endogeneity of fatigue driving propensity. The influence of fatigue driving on injury  
438 severity is significantly underestimated if ignoring the unobserved correlation between  
439 fatigue driving behavior and crash injury severity propensity. This study also compares  
440 the difference in risk factors of fatigue driving behavior and crash-related injury  
441 between commercial vehicle drivers and non-commercial drivers. Some common  
442 observed influential factors are identified. For instance, driving on express way not only  
443 contribute to higher fatal injury propensity but also high fatigue driving propensity.

444 Measures aiming at preventing driver fatigue such as light signals or signs may also  
445 help to reduce injury severity in the crash. It is also found in the paper that factors show  
446 different impacts on them. Driver's gender and age has significant influence on fatigue  
447 driving propensity of non-commercial vehicle driver, but this influence is not  
448 significant on commercial vehicle driver.

449 It should arouse the attention of researchers that the harmfulness of driver fatigue  
450 on traffic crash injury severity is larger than expected due to neglecting of the  
451 endogeneity of fatigue. Furthermore, correctly understanding the impact of fatigue-  
452 related crash is considered to be essential to the development and design of  
453 countermeasures aimed at reducing the hazard of fatigue crash. Different impact factors  
454 identified between commercial and non-commercial vehicle drivers in this study should  
455 be addressed. Some factors have similar impacts for both commercial and non-  
456 commercial vehicle drivers (e.g. road types and lighting conditions), but some factors  
457 have not (e.g. collision types). Thus, developing effective measures to reduce fatigue-  
458 related crash occurrence and its injury severity should take into account those  
459 differences. Moreover, according to our findings, police makers should also consider  
460 installing driver fatigue prevention devices (e.g. deceleration strip or warning signs) on  
461 express ways since those devices help reducing driver fatigue as well as injury severity.

462 With respect to fatigue driving behavior, our results suggest that fatigue is  
463 important in reducing the likelihood of fatal injury. However, one of the major  
464 limitations of this research is the sparseness of fatigue-related crash in this dataset for  
465 the reason that small sample size may influence model estimation. It is essential to

466 address endogenous in the model since endogeneity would cause inconsistent  
467 estimation. And this model can apply for any crash type when there is potential  
468 endogenous dependent variable. Therefore, it is reasonable to introduce this model in  
469 our analysis given the data limitation. In addition, the number of observations of some  
470 variables are small that may also limit the ability of determining effects precisely. And  
471 the vague and broad definition of fatigue may also cause misclassification problem and  
472 reduce the accuracy of our data analysis. This paper also does not consider the potential  
473 confounding effects of driving mileage, driver's health condition, drug use, which could  
474 affect both fatigue and risk of crash (Connor et al., 2001), due to the limitation of data.  
475 To deal with this problem, more detail and complete data are needed. Interaction effects  
476 or non-linear effects of variables and heterogeneity of drivers, which may also have  
477 significant impacts on injury severity and fatigue driving propensity, are not considered  
478 in this study for the first attempt since the focus is on the endogeneity of fatigue driving.  
479 Those problems will be discussed in our future studies.

480

## 481 **Acknowledgements**

482 This research was supported in part by the National Natural Science Foundation of  
483 China [grant number 71573286]. The first author would like to thank the China  
484 Scholarship Council (CSC) for financial support.

485 **Reference**

- 486 Akerstedt, T. & Kecklund, G., 2001. Age, gender and early morning highway accidents.  
487 Journal of Sleep Research, 10(2), 105-110.
- 488 Armstrong, K. A., Smith, S. S., Steinhardt, D. A. & Haworth, N. L., 2008. Fatigue  
489 crashes happen in urban areas too: characteristics of crashes in low speed urban  
490 areas. Road Safety Conference.
- 491 Armstrong, K. A., Obst, P., Livingstone, K. & Haworth, N., 2011. Investigation of  
492 differences in crash characteristics between males and females involved in  
493 fatigue-related crashes or close-call events. Women's Issues in Transportation, 26.
- 494 Arnold, P. K., Hartley, L. R., Corry, A., Hochstadt, D., Penna, F. & Feyer, A. M., 1997.  
495 Hours of work, and perceptions of fatigue among truck drivers. Accident Analysis  
496 & Prevention, 29(4), 471-477.
- 497 Australian Transport Council, 2011. National Road Safety Strategy 2011-2020.
- 498 Beirness, D. J., Simpson, H. M. & Desmond, K., 2005. The road safety monitor 2004:  
499 Drowsy driving.
- 500 Canadian Council of Motor Transport Administrator, 2010. Canadian fatigue related  
501 collisions: Fatality estimates 2000-2005.
- 502 Connor, J., Norton, R., Ameratunga, S., Robinson, E., Wigmore, B. & Jackson, R., 2001.  
503 Prevalence of driver sleepiness in a random population-based sample of car  
504 driving. Sleep, 24(6), 688-694.
- 505 Corfitsen, M. T., 1999. 'Fatigue' among young male night-time car drivers: is there a  
506 risk-taking group?. Safety Science, 33(1), 47-57.
- 507 Department for Transport, 2016. Reported Road Casualties Great Britain (2015)  
508 Annual Report. Department for Transport, London.
- 509 Desmond, P. A. & Hancock, P. A., 2001. Active and passive fatigue states. Stress,  
510 workload and fatigue.
- 511 Dinges, D. F. & Kribbs, N. B., 1991. Performing while sleepy: Effects of  
512 experimentally-induced sleepiness.
- 513 Dyani, J. S., 2007. Active versus passive fatigue in simulated driving. University of  
514 Cincinnati.
- 515 European Transport Safety Council, 2001. The role of driver fatigue in commercial road  
516 transport crashes.
- 517 Elvik R., 1995. Meta-analysis of evaluations of public lighting as accident

- 518 countermeasure. *Transportation Research Record*, 1485, 112-123.
- 519 Fadhli, M. M., Mohamed, N., Othman, H., Sarani, R. & Voon, W. S., 2008. Prevalence  
520 of fatigue among commercial bus drivers in Malaysia.
- 521 Fernandez-Antolin, A., Stathopoulos, A., & Bierlaire, M., 2014. Exploratory analysis  
522 of endogeneity in discrete choice models. In 14th Swiss Transport Research  
523 Conference (No. EPFL-CONF-202581).
- 524 Fernandes, R., Hatfield, J. & Job, R. S., 2010. A systematic investigation of the  
525 differential predictors for speeding, drink-driving, driving while fatigued, and not  
526 wearing a seat belt, among young drivers. *Transportation Research Part F: Traffic  
527 Psychology and Behaviour*, 13(3), 179-196.
- 528 Folkard, S., 1997. Black times: temporal determinants of transport safety. *Accident  
529 Analysis & Prevention*, 29(4), 417-430.
- 530 Greene, W.H., 2012. *Econometric Analysis*. 7th Edition, Pearson Education, Inc.,  
531 Boston.
- 532 Gregersen, N. P. & Bjurulf, P., 1996. Young novice drivers: Towards a model of their  
533 accident involvement. *Accident Analysis & Prevention*, 28(2), 229-241.
- 534 Hatfield, J., Murphy, S., Kasparian, N. & Soames Job, R. F., 2005. Risk perceptions,  
535 attitudes and behaviours regarding driver fatigue in NSW youth. Injury Risk  
536 Management Research Centre (IRMRC), University of New South Wales.
- 537 Hanowski, R. J., Hickman, J., Fumero, M. C., Olson, R. L. & Dingus, T. A., 2007. The  
538 sleep of commercial vehicle drivers under the 2003 hours-of-service regulations.  
539 *Accident Analysis & Prevention*, 39(6), 1140-1145.
- 540 Haworth, N., 1998. *Fatigue and fatigue research: the Australian experience*. Accident  
541 Research Centre, Monash University.
- 542 Horne, J. A. & Reyner, L. A., 1995. Sleep related vehicle accidents. *British Medical  
543 Journal*, 310(6979), 565-567.
- 544 Horne, J. A. & Reyner, L. A., 1999. Vehicle accidents related to sleep: a review.  
545 *Occupational and Environmental Medicine*, 56(5), 289-294.
- 546 Horne, J. A. & Reyner, L. A., 2001. Sleep-related vehicle accidents: some guides for  
547 road safety policies. *Transportation Research Part F: Traffic Psychology and  
548 Behaviour*, 4(1), 63-74.
- 549 Karrer, K., Vöhringer-Kuhnt, T., Baumgarten, T. & Briest, S., 2004. The role of  
550 individual differences in driver fatigue prediction. The third International  
551 Conference on Traffic and Transportation Psychology.
- 552 Mavjee, V. & Home, J. A., 1994. Boredom effects on sleepiness/alertness in the early

553           afternoon vs. early evening and interactions with warm ambient temperature.  
554           British Journal of Psychology, 85(3), 317-333.

555   Maycock, G., 1996. Sleepiness and driving: the experience of UK car drivers. *Journal*  
556           of Sleep Research, 5(4), 229-231.

557   McCartt, A. T., Rohrbaugh, J. W., Hammer, M. C. & Fuller, S. Z., 2000. Factors  
558           associated with falling asleep at the wheel among long-distance truck drivers.  
559           Accident Analysis & Prevention, 32(4), 493-504.

560   McConnell, C., Bretz, K. & Dwyer, W., 2003. Falling asleep at the wheel: A close look  
561           at 1,269 fatal and serious injury-producing crashes. *Behavioral Sleep Medicine*,  
562           1(3), 171–83.

563   Meng, F., Li, S., Cao, L., Li, M., Peng, Q., Wang, C. & Zhang, W., 2015. Driving  
564           fatigue in professional drivers: a survey of truck and taxi drivers. *Traffic Injury*  
565           Prevention, 16(5), 474-483.

566   Mock, C., Amegashie, J. & Darteh, K., 1999. Role of commercial drivers in motor  
567           vehicle related injuries in Ghana. *Injury Prevention*, 5(4), 268-271.

568   Nantulya, V. M. & Muli-Musiime, F., 2001. Kenya: Uncovering the social determinants  
569           of road traffic accidents. *Challenging inequities in health: from ethics to action*,  
570           211-225.

571   National Sleep Foundation, 2009. *Facts About Drowsy Driving*.

572   National Transportation Safety Board, 1995. Factors that affect fatigue in heavy truck  
573           accidents. *Safety Study 95/01*. National Transportation Board, Washington, DC.

574   Nordbakke, S. & Sagberg, F., 2007. Sleepy at the wheel: Knowledge, symptoms and  
575           behaviour among car drivers. *Transportation Research Part F: Traffic Psychology*  
576           and Behaviour, 10(1), 1-10.

577   Ogden, E. J. & Moskowitz, H., 2004. Effects of alcohol and other drugs on driver  
578           performance. *Traffic Injury Prevention*, 5(3), 185-198.

579   Oron-Gilad, T., Ronen, A. & Shinar, D., 2008. Alertness maintaining tasks (AMTs)  
580           while driving. *Accident Analysis & Prevention*, 40(3), 851-860.

581   Owens, D. A. & Sivak, M., 1996. Differentiation of visibility and alcohol as  
582           contributors to twilight road fatalities. *Human Factors*, 38(4), 680-689.

583   Pack, A. I., Pack, A. M., Rodgman, E., Cucchiara, A., Dinges, D. F., Schwab, C. W.,  
584           1995. Characteristics of crashes attributed to the driver having fallen asleep.  
585           Accident Analysis & Prevention, 27(6), 769-775.

586   Philip, P., Sagaspe, P., Moore, N., Taillard, J., Charles, A. E., Guilleminault, C. &  
587           Bioulac, B., 2005. Fatigue, sleep restriction and driving performance. *Accident*

588           Analysis & Prevention, 37(3), 473-478.

589 Philip, P., Sagaspe, P., Taillard, J., Moore, N., Guilleminault, C., Sanchez-Ortuno, M.,  
590           Akerstedt, T. O. R. & Bioulac, B., 2003. Fatigue, sleep restriction, and  
591           performance in automobile drivers: a controlled study in a natural environment.  
592           *Sleep*, 26(3), 277-280.

593 Pilcher, J. J. & Huffcutt, A. J., 1996. Effects of sleep deprivation on performance: a  
594           meta-analysis. *Sleep*, 19(4), 318-26.

595 Plainis, S., Murray, I. J. & Pallikaris, I. G., 2006. Road traffic casualties: understanding  
596           the night-time death toll. *Injury Prevention*, 12(2), 125-138.

597 Radun, I. & Radun, J. E., 2009. Convicted of fatigued driving: Who, why and how?  
598           *Accident Analysis & Prevention*, 41(4), 869-875.

599 Rajaratnam, S. M. & Arendt, J., 2001. Health in a 24-h society. *The Lancet*, 358(9286),  
600           999-1005.

601 Stutts, J. C., Wilkins, J. W., Osberg, J. S. & Vaughn, B. V., 2003. Driver risk factors  
602           for sleep-related crashes. *Accident Analysis & Prevention*, 35(3), 321-331.

603 Stutts, J. C., Wilkins, J. W. & Vaughn, B. V., 1999. Why do people have drowsy driving  
604           crashes? Input from drivers who just did. AAA Foundation for Traffic Safety,  
605           Washington D.C..

606 Sullivan, J. M. & Flannagan, M. J., 2002. The role of ambient light level in fatal crashes:  
607           inferences from daylight saving time transitions. *Accident Analysis & Prevention*,  
608           34(4), 487-498.

609 Summala, H. & Mikkola, T., 1994. Fatal accidents among car and truck drivers: effects  
610           of fatigue, age, and alcohol consumption. *Human Factors: The Journal of the*  
611           *Human Factors and Ergonomics Society*, 36(2), 315-326.

612 Tefft, B. C., 2010. Asleep at the wheel: the prevalence and impact of drowsy driving.

613 Thiffault, P. & Bergeron, J., 2003. Monotony of road environment and driver fatigue:  
614           a simulator study. *Accident Analysis & Prevention*, 35(3), 381-391.

615 Traffic Management Bureau, Ministry of Public Security, PRC, 2013. China Road  
616           Traffic Accidents Annual Statistical Report.

617 Vanlaar, W., Simpson, H., Mayhew, D. & Robertson, R., 2008. Fatigued and drowsy  
618           driving: A survey of attitudes, opinions and behaviors. *Journal of Safety Research*,  
619           39(3), 303-309.

620 Williamson, A., Lombardi, D. A., Folkard, S., Stutts, J., Courtney, T. K. & Connor, J.  
621           L., 2011. The link between fatigue and safety. *Accident Analysis & Prevention*,  
622           43(2), 498-515.

- 623 World Health Organization. Global status report on road safety 2015. Geneva: World  
624 Health Organization, Department of Violence and Injury Prevention and  
625 Disability.
- 626 Yamamoto, T. & Shankar, V. N., 2004. Bivariate ordered-response probit model of  
627 driver's and passenger's injury severities in collisions with fixed objects. *Accident  
628 Analysis & Prevention*, 36(5), 869-876.
- 629 Young, S. L. & Hashemi, L., 1996. Fatigue and trucking accidents: two modes of  
630 accident causation. *Proceedings of the Human Factors and Ergonomics Society  
631 Annual Meeting*, 40(18), 952-956.
- 632 Zhang, G., Yau, K. K., Zhang, X. & Li, Y., 2016. Traffic accidents involving fatigue  
633 driving and their extent of casualties. *Accident Analysis & Prevention*, 87, 34-42.
- 634 Zhao, X. & Rong, J., 2013. The relationship between driver fatigue and monotonous  
635 road environment. *Computational Intelligence for Traffic and Mobility*, 8, 19-36.