Report #15-UT-035

NSTSCE

National Surface Transportation Safety Center for Excellence

Evaluating the Influence of Crashes on Driving Behavior using Naturalistic Driving Study Data

Feng Guo, Ph.D. • Chen Chen

Submitted: July 16, 2015

Technology
Impairment
Mobility
Infrastructure

Housed at the Virginia Tech Transportation Institute
3500 Transportation Research Plaza • Blacksburg, Virginia 24061

ACKNOWLEDGMENTS

The authors of this report would like to acknowledge the support of the stakeholders of the National Surface Transportation Safety Center for Excellence (NSTSCE): Tom Dingus from the Virginia Tech Transportation Institute, John Capp from General Motors Corporation, Lincoln Cobb from the Federal Highway Administration, Chris Hayes from Travelers Insurance, Martin Walker from the Federal Motor Carrier Safety Administration, and Cathy McGhee from the Virginia Department of Transportation and the Virginia Center for Transportation Innovation and Research.

The NSTSCE stakeholders have jointly funded this research for the purpose of developing and disseminating advanced transportation safety techniques and innovations.

ABSTRACT

It is hypothesized that intense events such as crashes could influence driver behavior and driving risk. This study evaluated the influences of crash events on driver behavior and driving risk using data from the 100-Car Naturalistic Driving Study, which included 51 crashes from primary drivers. Two metrics were used to measure driver behavior and risk: the proportion of baselines where the drivers were engaged in complex and moderate secondary tasks and the intensity of the near-crashes (NCs) and safety-critical incidents (SCIs). For the distraction analysis, we sampled 882 6-second baseline epochs within 15-hour windows before and after crashes. Results from a mixed binomial regression model indicated that the percentage of baselines where drivers engaged in complex secondary tasks dropped after crashes (odds ratio = 0.54; 95% CI [0.32, 0.93]). The driving risk analysis used the intensity of SCIs and NCs to measure the driving risk. Since there are typically more than one SCI and NC events before and after a crash, we developed four alternative recurrent event models to evaluate the influence of crashes based on actual driving time. The driving period was divided into several phases based on the relationship to crashes, and the intensities of these periods were compared. Results show a reduction in SCI intensity after the first crash (intensity rate ratio = 0.82; 95% CI [0.693, 0.971]) and the second crash (intensity rate ratio = 0.47; 95% CI [0.377, 0.59]) for male drivers. Females were observed to have a nonsignificant response to the first crash, but SCI intensity decreased after the second crash (intensity rate ratio = 0.43; 95% CI [0.342, 0.547]). This study indicated that crashes do have a positive effect on drivers' behavior in terms of distraction and driving risk.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Driver behavior is a critical contributing factor to traffic safety. It is estimated that more than 90% of crashes are associated with driver errors (Treat, 1980). A study by Curry, Hafetz, Kallan, Winston, and Durbin (2011) indicated that 95.6% of all teen-involved serious crashes were due to driver error.

The current study investigated the influences of crashes on driving behavior and driving risk. The hypothesis was that drivers are more cautious after a crash, which is reflected in the following two aspects: (1) a reduction in the probability of distraction and (2) a reduction in the intensity of safety-critical incidents (SCIs) and/or near-crashes (NCs). We also explored how the observed effects change over time. Furthermore, we evaluated whether the potential reduction differs by demographic factors such as gender and age.

Driving behavior was measured by the probability of distraction among randomly selected baseline samples occurring during a specific time period. Mixed binomial regression models were used to evaluate the factors that affect the probability of distraction. Driving risk was measured by the intensity of an SCI or NC. Since for a given crash, both SCIs and NCs can occur repeatedly before and after the crash, four intensity-based recurrent event models were developed to assess the change of SCI intensity and NC intensity before and after crashes.

EVALUATING THE INFLUENCES OF CRASHES ON DRIVER DISTRACTION

To evaluate driver behavior before and after crashes, a before–after time "window" was defined, e.g., 10 hours of driving time before and 10 hours after a crash. Within this window, random samples of 6-second baselines were identified from two sources. The first source is an existing baseline sample from a previous National Surface Transportation Safety Center for Excellence (NSTSCE) project, which contained 10,952 baseline samples. A detailed discussion of the baseline selection scheme is found in Guo and Hankey (2009). To increase the sample size close to crash time, 514 additional baseline samples were reduced within a 30-hour window around the crashes. Eventually, four samples were randomly selected within a 5-hour window. For the rest of the 30-hour window, two samples were selected in every 5-hour window. For each baseline, a rigorous data reduction protocol was used to extract driver behavior information. Data collection is illustrated in Figure 1.

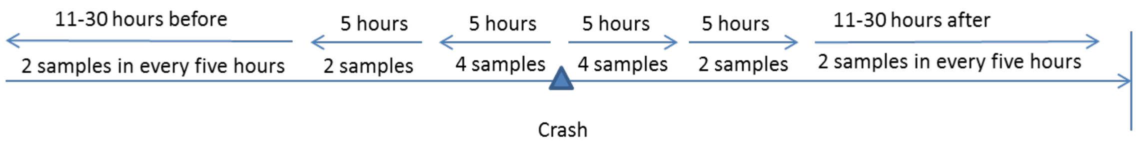


Figure 1. Chart. Before/after baseline sample collection illustration.

Within the before and after windows, we used the percentage of baselines with moderate and complex secondary tasks to evaluate the influences of crashes on driver behavior. According to the visual and manual demand of the secondary task, baselines can be classified into four categories: (1) no distraction, (2) simple secondary task, (3) moderate task, and (4) complex secondary task. We adopted moderate and complex secondary tasks as indicators of high-risk behavior and used the percentage of baselines with moderate and complex secondary tasks to measure the likelihood of the driver to engage in risky behavior.

For each crash, the 15-hour windows before and after were considered as a matched pair, and the number of moderate and complex secondary tasks and total number of baselines were evaluated. Drivers were considered to have driven more cautiously if a larger proportion of moderate and complex secondary tasks occurred before the crash than after. Mixed effect binomial regression was adopted to (1) incorporate the correlation among observations from the same driver, and (2) to adjust for confounding effect (e.g., age and gender) through modeling.

As illustrated in Figure 2, the distraction probability in general is lower during the before-crash period, especially in the initial 15 hours. The difference diminished as the time window increased in size and the percentage is almost equal after about 50 hours. This result suggests that drivers tend to be less engaged by distractions during the initial period after a crash but return to regular behavior after a certain time period. The confidence band is relatively wide, which could be due to the relatively small sample size.

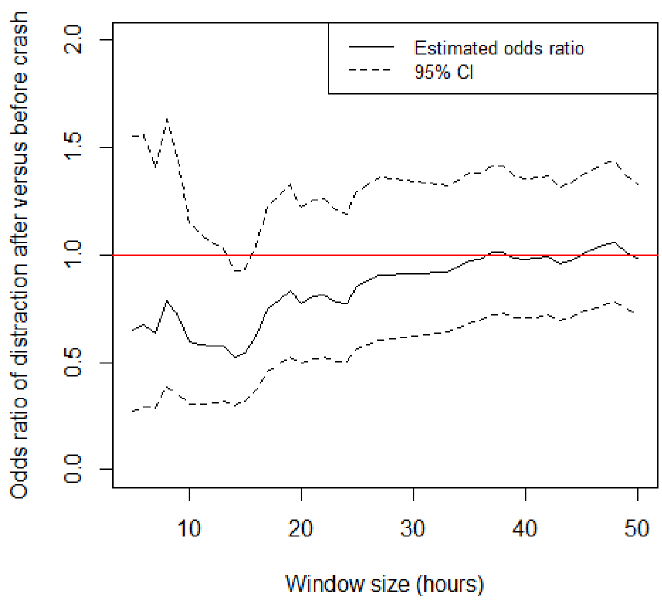


Figure 2. Graph. Crash influence on distraction proportion.

RECURRENT EVENT APPROACH FOR EVALUATING THE INFLUENCES OF CRASH ON DRIVING RISK

The literature suggests that safety-related conflicts, such as high g-force events and SCIs, are related to driving risk (Guo and Fang, 2013; Guo, F., Klauer, S. G., Hankey, J. M., and Dingus, T. A., 2010). In this study, we used the intensity of the SCIs and NCs to measure driving risk. Traditional analysis is based on event frequency, such as Poisson and negative binomial models, which requires arbitrarily defined time intervals and generally lacks statistical power. Therefore, we adopted the intervals between events to measure the intensity. Since both SCIs and NCs occurred repeatedly before and after crashes, recurrent event models were adopted. Recurrent event models focus on the time to multiple events for a subject or cluster (Andersen & Gill, 1982). An intensity-based recurrent event model was used by treating the number of SCIs and NCs over time as a counting process. Data setting is shown in Figure 3, where each horizontal line represents the driving time of one driver. Drivers were subject to different numbers of crashes, NCs, and SCIs at various time points through the entire study. The time to each crash, NC, and SCI was recorded. We focused on the actual driving time. Non-driving times when the vehicle was not in use were excluded. The driving period was divided into several phases based on the relationship to crashes: before first crash (coded as 0), between the first and second crash (coded as 1), and after the second crash (coded as 2).

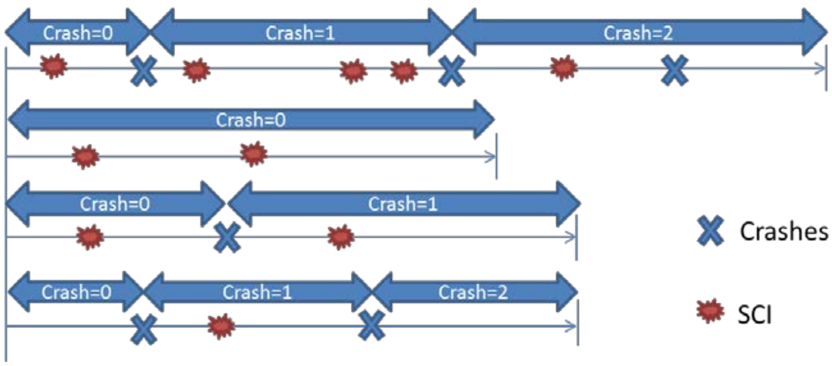


Figure 3. Chart. Data setting for intensity-based recurrent event model.

To account for potential confounding and interacting effects, gender and the age of the driver when first enrolled in the study were also evaluated in the model.

We developed four alternative models, including an Andersen-Gill (A-G) model, a stratified A-G model, a frailty model, and a stratified frailty model. A simulation study was conducted to examine the performance of the proposed models. Model comparison indicated that the stratified frailty models performed best for the data set and these were adopted for data analysis.

RESULTS

For illustration purposes, Figure 4 presents the ratio of SCI rate between the after-crash and before-crash windows across different window sizes. SCI rate is calculated as the number of SCIs per hour driving for each observed time window. As Figure 4 shows, male drivers show a stronger decrease pattern than female drivers as SCI rate is lower after crashes compared to the before window. The difference between before and after crashes gradually diminishes (ratio is close to 1) as window size increases. For most cases, the observed time period before or after a crash equals the chosen window size. However, the observed time period could also be shorter than the chosen window size because of the issue of overlapping or if not enough driving time records were available before or after that crash.

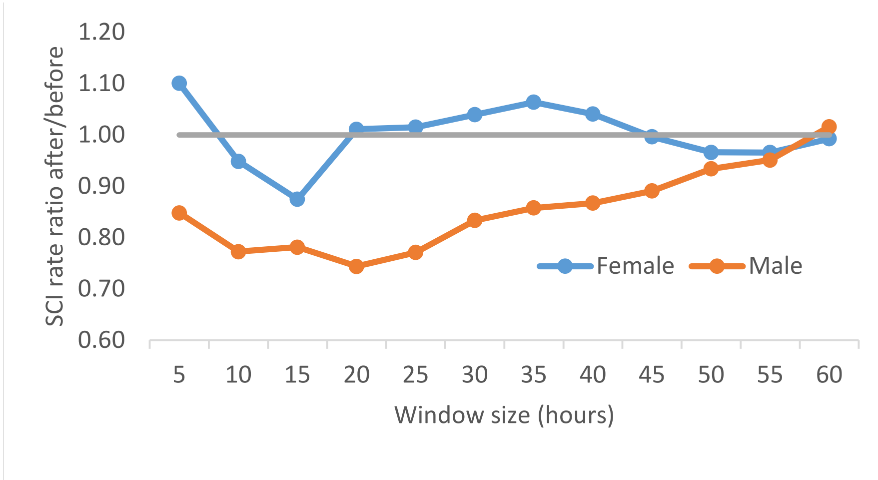


Figure 4. Graph. Ratio of SCI rate between after-crash and before-crash windows across window sizes.

Table 1 provides an estimate of crash effect on SCI from the stratified frailty model. It can be seen that both first and second crashes influence male drivers. The intensity rate after the first crash is 0.82 times (95% CI [0.693, 0.971]) the rate before the first crash. The second crash has a bigger influence, reducing the intensity to 0.472 times (95% CI [0.377, 0.59]) that of before. The first crash does not show a significant influence on female drivers. However, after the second crash, the intensity rate is observed to be 0.432 times (95% CI [0.342, 0.547]). Age is not significant (p = 0.43). Estimation of variation among driver variation (σ) is 1.28 based on restricted maximum likelihood (REML) estimation.

Table 1. Stratified frailty model crash effect estimation on SCI.

Contrast	Regression coefficient	Intensity rate ratio	Lower confidence limit of intensity rate ratio	Upper confidence limit of intensity rate ratio	Pr > ChiSq
1 vs. 0 Female	0.109	1.115	0.96	1.296	0.155
2 vs. 1 Female	-0.839	0.432	0.342	0.547	**<.0001**
1 vs. 0 Male	-0.199	0.820	0.693	0.971	**0.021**
2 vs. 1 Male	-0.751	0.472	0.377	0.59	**<.0001**

NCs were observed much less frequently than SCIs. There were four females who experienced two or more crashes in the study, as shown in Table 4. After careful examination, only one of them had at least one NC recorded after the second crash. The other three have the second crash as their last driving record. Consequently, estimation of the second crash effect for females depends heavily on one particular driver, which may lead to an individual crash influence rather than a population-wise effect. Thus, we decided to use time up to the second crash only and evaluate the first crash influence on NC.

The result is similar to that of SCI, as shown in Table 2. For male drivers, the intensity rate after the first crash is 0.52 times (95% CI [0.314, 0.874]) that before the first crash. Female drivers do not show a significant decreasing trend after the first crash. Estimation of σ is 0.93 based on REML estimation.

Table 2. Stratified frailty model crash effect estimation on NC.

Contrast	Regression coefficient	Intensity rate ratio	Upper confidence limit of intensity rate ratio	Lower confidence limit of intensity rate ratio	Pr > ChiSq
1 vs. 0 Female	-0.117	0.89	0.551	1.438	0.6333
1 vs. 0 Male	-0.646	0.524	0.314	0.874	**0.0134**

SUMMARY AND CONCLUSION

This study evaluated the influences of crashes on driver distraction behavior and the driving risk using 100-Car NDS data. The results indicate that drivers' engagement in moderate and complex secondary tasks tends to be lower after crashes, especially within a 15-hour driving time window. However, this decreasing effect tends to diminish over time and no difference is observed after 50 hours.

As measured by the intensity of SCIs and NCs, female and male drivers showed different responses to crashes. Male drivers responded to both the first crash and the second crash with lower SCI and NC intensity after each crash. Female drivers showed no significant response to the first crash but did show decreased SCI intensity after the second crash.

This study suggests that crashes do have positive effects on drivers' behavior in terms of both distraction and aggressive driving. However, the effect diminished quickly after crashes at about 50 hours. Further study of how to prolong this improvement in safe behavior will benefit both safety education and efforts to develop corresponding safety countermeasures.

The study is limited by the relative small number of crashes (51 crashes) as well as mild crash severity. With a larger NDS data set, such as the Second Strategic Highway Research Program (SHRP 2) NDS data, we hope to find more concrete evidence of the influences of crashes on driver behavior and potentially the influences of crash by severity.

CHAPTER 1. INTRODUCTION AND BACKGROUND

Driver behavior is a major contributing factor to traffic safety. It has been estimated that more than 90% of crashes are associated with driver errors (Treat, 1980). Curry, Hafetz, Kallan, Winston, and Durbin (2011) concluded that 95.6% of all teen-involved serious crashes were due to driver error.

Studies have shown that driving risk decreases with increased driving experience (Waller, Elliot, Shope, Raghunathan, & Little, 2001; Kaneko and Jovanis, 1992; Chipman, 1982). Previous studies on driver experience are typically based on measures from years of driving and/or mileage (e.g., Waller et al., 2001; Chipman, 1982; Levy, 1990). For example, Waller et al. (2001) used time duration since licensure as the measurement of experience, and Kaneko and Jovanis (1992) considered the number of years of experience of drivers from a national less-than-truckload firm as a factor. Experience based on years or mileage of driving includes the effects of many factors. For example, drivers are more mature and engage in fewer risky behaviors, they learn more skills and are able to deal with more complex situations, or they learn lessons from drastic crash events. In addition, using time and the amount of driving as a measurement of experience is commonly associated with age-related changes (af Wahlberg, 2012). Thus, with experience measures it is difficult to isolate the effects of a specific associated factor.

It is hypothesized that crash experience would lead to reduced driving risk. The rationale is that drivers will learn from their collision events (crashes) and change their behavior correspondingly, thus reducing driving risk. From a psychological point of view, Lucas (2003) showed that drivers who had been involved in a motor vehicle accident reported significantly greater worries about driving than did drivers who had not been in an accident.

There are limited studies directly linking crash experience with observed driving risk. Lin, Huang, Hwang, Wu, and Yen (2004) evaluated the association between crash experience and risk-taking behavior among students in Taiwan and found no significant association. Crash experience was measured by self-reported crash history prior to the study, crash frequency, time elapsed since the last crash, and crash severity. af Wahlberg (2012) compared the behavior of bus drivers between drivers with and without crash experience over a three-year period. Repeated measurements of speed-change behaviors were collected and a steady decline in speed change over time was observed within the crash group but was determined not to be due to crashes. A similar pattern was observed for the no-crash group. Rajalin and Summala (1997) studied the effect of fatal accidents on surviving drivers' subsequent driving behavior based on self-reported driving behavior. The study showed that light-vehicle drivers typically returned to their "normal" driving behavior within a few months, while heavy-vehicle drivers tended to be more cautious in terms of driving mileage. The majority of existing studies are based on self-reported driving behavior, including risk-taking score, speed change, and amount of driving. However, there is limited research using objective measures of driving behavior such as driving data collected *in situ*, specifically naturalistic driving data.

1.1 NATURALISTIC DRIVING STUDY AND RISK ANALYSIS

Naturalistic driving study (NDS) provides an innovative way to access traffic safety and driving behavior data (Dingus et al., 2006) and thus makes exploring the relationship between crash experience and driving behavior and risk accessible. Participant vehicles are instrumented with data acquisition systems (DASs) that include cameras and various sensors to continuously monitor the driving process. The video images and kinematic measures can provide not only the exact driving behavior, vehicle kinematic, and driving environmental information, but also the sequence and precise time for each sub-event.

Among many advantages of NDS, the video recordings can be used to assess driver behaviors that were difficult to retrieve before. In the present study, distraction pattern was evaluated as driving behavior. A high frequency of distraction and/or more complex non-driving-related tasks was considered indicative of distracted driving behavior. To be more specific, secondary tasks, such as communications, entertainment, information gathering, and navigation not required to drive, were used to measure distraction. The secondary tasks were categorized into three levels: complex (C), moderate (M), and simple (S), based on whether the task requires multiple steps, multiple eye glances away from the forward roadway, and/or multiple button presses. Detailed categorization of distraction can be found in Table 3 (Guo and Hankey, 2009).

Table 3. Definition of distraction.

Simple secondary tasks	Moderate secondary tasks	Complex secondary tasks
1. Adjusting radio	1. Talking/listening to handheld device	1. Dialing a handheld device
2. Adjusting other devices integral to the vehicle	2. Handheld device-other	2. Locating/reaching/answering handheld device
3. Talking to passenger in adjacent seat	3. Inserting/retrieving CD	3. Operating a personal digital assistant (PDA)
4. Talking/singing: no passenger present	4. Inserting/retrieving cassette	4. Viewing a PDA
5. Drinking	5. Reaching for object (not handheld device)	5. Reading
6. Smoking	6. Combing or fixing hair	6. Animal/object in vehicle
7. Lost in thought	7. Other personal hygiene	7. Reaching for a moving object
8. Other simple tasks	8. Eating	8. Insect in vehicle
	9. Looking at external object	9. Applying makeup

NDS also provides alternative ways to measure driving risk. Specifically, several types of safety-related events can be identified through kinematic signatures of the vehicle and confirmed through visual inspection of video recordings. Crash, near-crash (NC), and safety-critical incident (SCI) are the major categories used. A *crash* is defined as any contact with an object, either moving or fixed, at any speed in which kinetic energy is measurably transferred or dissipated. Crashes include a participant's vehicle making contact with other vehicles, roadside barriers, and objects on or off the roadway, pedestrians, cyclists, or animals. A *near-crash* is defined as any circumstance requiring a rapid, evasive maneuver by the participant (or his/her vehicle) or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. The crashes and NCs were identified through a multiple-step process of automatic trigger identification, followed by visual confirmation by experts as described in Dingus et al. (2006). A *safety-critical incident* is defined as an unexpected event resulting in a close call or requiring fast action (evasive maneuver) on the part of a driver to avoid a crash (Dingus et al., 2006). These safety-related events represent non-desired safety conditions that should be avoided and are widely used in the literature as a surrogate of crash for measuring driving risk (Guo and Fang, 2010) and are adopted in this report.

1.2 STUDY OBJECTIVES AND DATA

The current study sought to investigate the influence of crashes on driving behavior and driving risk. The objective was to evaluate whether drivers are more cautious in terms of frequency of distraction and whether the SCI frequency decreases after a crash. Furthermore, we were also interested in whether male and female drivers respond to crashes differently.

Distraction changes constantly during driving. Although NDS makes the entire driving record available, it is not possible to reduce all video recordings by visual inspection and keep track of distraction. Guo and Hankey (2009) proposed an analysis framework based on a case-cohort approach. Under this analysis framework, a random sampling scheme for baseline reduction led to approximation of odds ratio to event rate ratio. This random sampling scheme is stratified by drivers, and the number of samples for each driver is proportional to the valid moving hours or miles traveled. The random samples also present the general behavior of drivers under normal driving conditions, and thus can be used to evaluate driver distraction.

Data for the analyses were drawn from the 100-Car NDS. The 100-Car NDS was the first instrumented vehicle study undertaken with the primary purpose of collecting large-scale naturalistic driving data. Data were collected from 241 primary and secondary driver participants in northern Virginia. About 2,000,000 vehicle miles and 43,000 hours of driving are recorded in total. This study used data from 107 primary drivers. Three types of safety-related events were identified and used in this study, including 51 crashes, 610 NCs, and 6,659 SCIs. Crash distribution across drivers by their demographic information is given in Table 4.

For this study, 11,466 baseline samples were incorporated from two sources. The first source is an existing baseline sample from a previous National Surface Transportation Safety Center for Excellence (NSTSCE) project (Guo and Hankey, 2009), which contained 10,952 baseline samples. To increase the sample size close to crash time, 514 additional baseline samples were reduced within a 30-hour window around the crashes, which brought the total sample size to 882 within this window. Among all baseline samples, 44% involve various levels of distraction, and 40% are categorized as moderate or complex distractions.

As discussed above, SCIs and NCs were used to measure driving risk. Average rate, average frequency of both SCI and NC by gender, age group, and total crashes are shown in Table 4, Table 5, and Table 6. Event rate is calculated as the ratio between overall frequency and total hours of driving. In general, SCIs occur 8 to 10 times more frequently than NCs across all groups of drivers. Higher SCI and NC rates are associated with drivers who experienced more crashes.

Table 4. SCI/NC rate, frequency, and driving time for 0 crash group.

	Females ≤30 yrs. old (25 subjects)	Males ≤30 yrs. old (24 subjects)	Females 31 to 55 yrs. old (12 subjects)	Males 31 to 55 yrs. old (32 subjects)	Females >55 yrs. old (5 subjects)	Males >55 yrs. old (9 subjects)
Number of drivers	11	15	10	26	4	7
SCI/NC rate^	0.18/0.02	0.15/0.02	0.15/0.02	0.11/0.01	0.04/0.01	0.12/0.02
Average SCI/NC*	57.4/4.6	45.8/4.5	35.8/3.8	49.9/2.4	8.8/1.3	39.1/5.7
Average driving hours#	255	317	236	378	264	332

Table 5. SCI/NC rate, frequency, and driving time for 1 crash group.

	Females ≤30 yrs. old (25 subjects)	Males ≤30 yrs. old (24 subjects)	Females 31 to 55 yrs. old (12 subjects)	Males 31 to 55 yrs. old (32 subjects)	Females >55 yrs. old (5 subjects)	Males >55 yrs. old (9 subjects)
Number of drivers	10	3	2	4	1	2
SCI/NC rate^	0.27/0.03	0.26/0.02	0.41/0.04	0.22/0.03	0.16/0.06	0.14/0.03
Average SCI/NC	71.2/7.2	36.0/2.7	175.5/19	99.5/12.3	5.0/2	53.5/6.5
Average driving hours#	344	345	448	431	32	360

Table 6. SCI/NC rate, frequency, and driving time for ≥2 crash group.

	Females ≤30 yrs. old (25 subjects)	Males ≤30 yrs. old (24 subjects)	Females 31 to 55 yrs. old (12 subjects)	Males 31 to 55 yrs. old (32 subjects)	Females >55 yrs. old (5 subjects)	Males >55 yrs. old (9 subjects)
Number of drivers	4	6	—	2	—	—
SCI/NC rate^	0.52/0.05	0.33/0.03		0.27/0.02		
Average SCI/NC	214.5/20.8	95.5/11		132/7.5		
Average driving hours#	415	359		489		

^: number of events per driving hour
*: average number of events per driver
#: average driving hours per subject
—: no data

CHAPTER 2. EVALUATING THE INFLUENCES OF CRASHES ON DRIVER DISTRACTION

In this study, driving behavior is measured by the probability of distraction in randomly selected baseline samples occurring over a specific time period. According to the visual and manual demand of the secondary task, baselines can be classified into four categories: (1) no distraction, (2) simple secondary task, (3) moderate task, and (4) complex secondary tasks. We adopted moderate and complex secondary tasks as indicators of high-risk behavior and used the percentage of baselines with moderate and complex secondary tasks to measure the likelihood of a driver to engage in high-risk behavior. Drivers were considered to have driven more cautiously if a larger proportion of moderate and complex secondary tasks occurred before a crash than after.

2.1 COUNT-BASED APPROACH

The count-based approach requires a predefined "window," e.g., 10 hours of driving time. For each crash, the predefined before-and-after windows were considered as a matched pair, and the number of moderate and complex secondary tasks as well as the total number of baselines were evaluated. Data collection for this approach is illustrated in Figure 5.

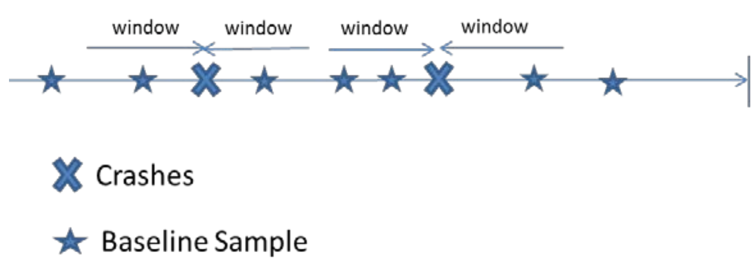


Figure 5. Chart. Before and after baseline sample collection.

An appropriate length of the time for the window needs to be carefully selected. If crash experience acts as a short stimulation for drivers and only affects driving behavior temporarily, a large window size will mask the effect of the crash by including non-influenced data. On the other hand, a small window size will not be able to capture enough event data and thus lose power to evaluate crash influence.

Window size selection is constrained by the overlapping problem, which refers to a situation where the time interval between two crashes is less than the window size. Ideally, window size is chosen to be smaller than the shortest time interval between two consecutive crashes for one driver. However, in the 100-Car NDS, one driver experienced two crashes within 1 hour of driving. Figure 6 shows a histogram of the time gaps between two consecutive crashes. Over 75% of two successive crashes for one driver were at least 30 hours apart, so the current study begins with a window size of 15 hours. Crash effect based on other window sizes is investigated later.

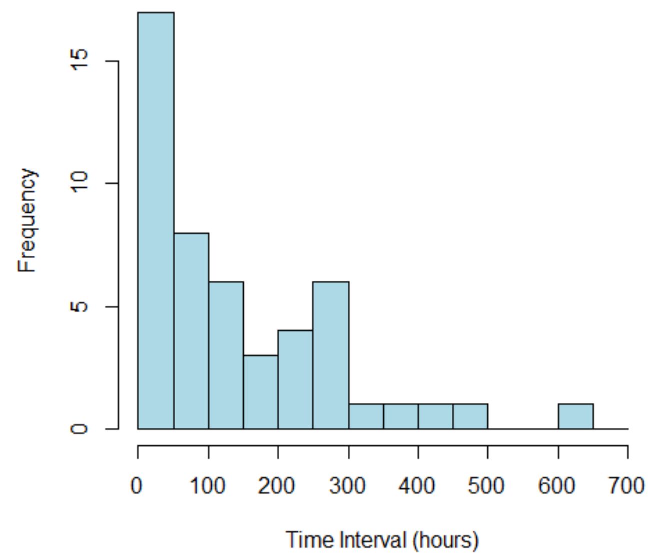


Figure 6. Graph. Histogram of crash time intervals.

An existing sample of 10,952 baselines was randomly sampled from the 100-Car data (Guo and Hankey, 2009). For each baseline, a rigorous data reduction protocol was used to extract driver behavior information. Among these baselines samples, only 952 fall in the 30-hour window around observed crashes. For the purpose of comparing distraction before and after a crash, we reduced an additional 514 baselines within a 30-hour driving window around crashes. With these additional data, the final sampling scheme, as illustrated in Figure 7, was as follows. Two samples were randomly selected within a 2-hour window before and after a crash, and two were randomly selected in the 2–5 hour window. For the rest of the 30 hours, two samples were randomly selected in every 5-hour window.

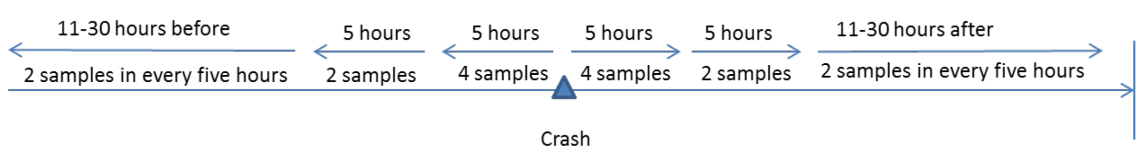


Figure 7. Chart. Final baseline sampling scheme.

Table 7 shows the total number of baseline samples from existing data and new data reduction across window sizes. There is a small proportion of double-counted samples in the table, due to overlapping windows between crashes. Baseline samples could be included in an after-crash window while being included in the before window of the next crash. Thus, the total number of baselines in Table 7 is larger than the total number of unique baselines that are identified from data reduction. Sampling rate, defined as the number of samples per hour, shows consistency between the before and after windows.

Table 7. Baseline sample distribution.

Window size (hours)	Guo and Hankey 2009 sample	New baseline reduction	Total samples before	Total sample after	Sampling rate before	Sampling rate after
5	179	235	218	196	44	39
10	326	320	336	310	34	31
15	502	380	450	432	30	29
20	669	425	559	535	28	27
25	796	498	660	634	26	25
30	952	551	773	730	26	24
35	1097	552	864	785	25	22
40	1246	558	949	855	24	21
45	1345	559	1003	901	22	20
50	1484	562	1076	970	22	19
55	1604	563	1136	1031	21	19
60	1706	564	1184	1086	20	18

Table 8 compares the proportion of moderate and complex distractions among the baseline samples. It can be observed that female drivers had a higher distraction rate than male drivers. Figure 8 provides the ratio between the moderate and complex distraction proportions before and after a crash. The results indicate that drivers' engagement in moderate and complex secondary tasks tends to be lower after crashes, especially within a 15-hour driving time window. However, this decreasing effect tends to diminish over time.

Table 8. Proportion of moderate and complex distraction comparison.

Window size (hours)	Proportion distraction before – Female	Proportion distraction before – Male	Proportion of distraction after – Female	Proportion of distraction after – Male	Ratio of distraction percentage – Female	Ratio of distraction percentage – Male
5	26% (26/102)	14% (16/116)	25% (22/93)	15% (14/103)	0.93	1.09
10	27% (41/158)	17% (28/178)	22% (32/153)	14% (21/157)	0.80	0.84
15	27% (52/210)	18% (41/240)	19% (45/220)	13% (28/212)	0.72	0.76
20	26% (60/259)	16% (47/300)	22% (61/268)	13% (35/267)	0.84	0.81
25	26% (69/303)	15% (51/357)	23% (74/317)	12% (38/317)	0.87	0.81
30	25% (79/353)	15% (61/420)	22% (82/360)	13% (48/370)	0.87	0.83
35	25% (84/390)	15% (67/474)	21% (85/390)	13% (52/395)	0.85	0.84
40	25% (95/434)	15% (73/515)	21% (91/418)	13% (59/437)	0.86	0.88
45	25% (98/460)	15% (78/543)	21% (94/431)	13% (63/470)	0.87	0.87
50	24% (104/500)	15% (81/576)	21% (96/462)	13% (68/508)	0.85	0.87
55	24% (108/528)	14% (82/608)	20% (100/493)	13% (71/538)	0.86	0.89
60	23% (113/555)	15% (88/629)	20% (103/518)	13% (76/568)	0.85	0.89

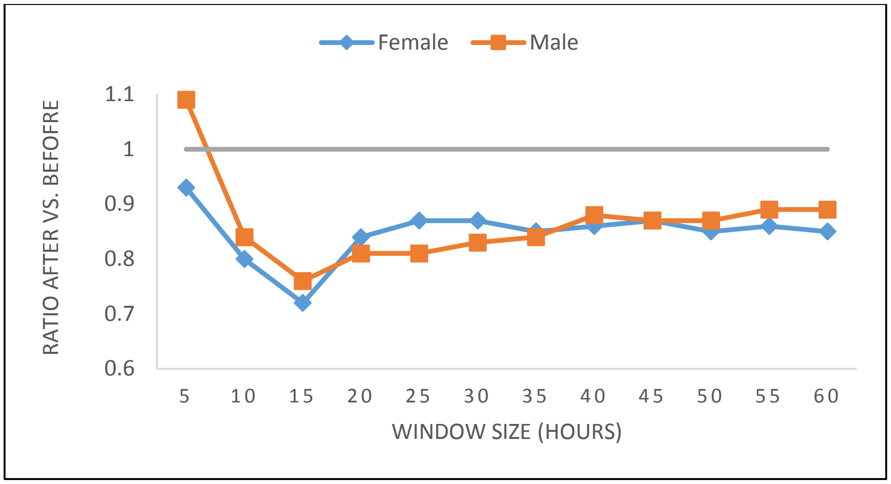


Figure 8. Graph. Ratio of distraction rate after vs. before by gender.

2.2 MODELING BASELINE DISTRACTION USING MIXED BINOMIAL REGRESSION

A formal statistical inference was conducted to investigate the influence of crashes on driver distraction. For each crash, the number of baselines in which drivers engaged in moderate and complex secondary tasks during the before and after period are considered as a matched pair. A crash effect is observed if a larger probability of moderate and complex secondary tasks occurs before a crash than after. Mixed binomial regression models are used to evaluate the factors that affect the probability of distraction. A mixed binomial regression model is adopted to (1) incorporate the correlation among observations from the same driver, and (2) to adjust for confounding effects, e.g. gender, through modeling. Gender effect allows distraction behavior between male and female to be discerned. The model is given as follows:

$$Y_{ijk} = binomial\left(n_{ijk}, p_{ijk}\right)$$

$$log\left(\frac{p_{ijk}}{1-p_{ijk}}\right) = \beta_0 + \beta_1 * x_{ijk} + \beta_2 * gender_i + \epsilon_i \text{ }, \epsilon_i \sim N(0, \sigma^2)$$

where

- Y_{ijk} is the total number of moderate and complex distractions for subject i (where $i = 1, \dots, 107$) in the before–after window of number j crash. Frequency follows a binomial distribution.
- n_{ijk} is the total number of baseline samples in the associated window.
- p_{ijk} is the probability of a moderate or complex distraction happening.
- x_{ijk} indicates whether the distraction happens before or after a crash: $x_{ijk} = 0$ if it happens before, otherwise $x_{ijk} = 1$.
- β_1 is crash effect and β_2 is gender effect.
- ϵ_i is a normally distributed random effect of mean 0 associated with each individual.

Results indicate that the percentage of baselines where drivers engaged in complex secondary tasks dropped after crashes. The maximum decrease occurred in the 15-hour window, with odds ratio = 0.54; 95% CI [0.32, 0.93]. Crash influence on distraction was also explored with varying window sizes, as shown in

Figure 9. Distraction probability decreased after a crash, especially in the initial 15 hours. The difference diminished as window size increased and became negligible after 50 hours. This result suggests that drivers tend to engage less in distractions during the initial period after a crash but return to regular behavior after a certain time period. The confidence band is relatively wide, which is primarily due to the relatively small sample size.

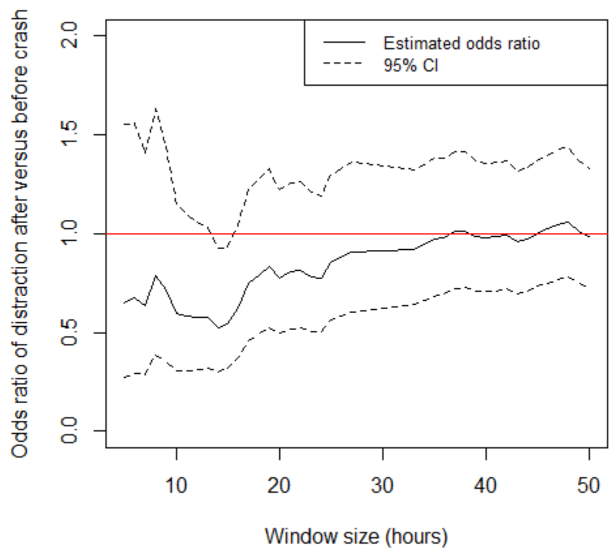


Figure 9. Graph. Crash influence on distraction proportion.

CHAPTER 3. INTENSITY-BASED RECURRENT EVENT APPROACH

Compared to the event frequency–based approach, time to event, i.e., SCI and NC, does not require an arbitrarily defined time interval and could be more informative in estimating the intensity of the event. Cox's proportional hazards model (Cox, 1972) is a commonly used method for time-to-event data and provides reliable estimates of survival times, as well as relative risk associated with risk factors. The majority of survival models are based on the assumption that there can be only one event for a subject. However, both SCIs and NCs were observed repeatedly for a given driver in the 100-Car NDS. As a result, a recurrent event approach, which focuses on the time to multiple events for any subject or cluster (Andersen and Gill, 1982), is more appropriate. Intensity-based recurrent event models treat the number of SCIs and NCs over time as a counting process. For the purpose of both modeling and statistical analysis, the concepts of intensity functions and counting process are illustrated with data setting as below. Without loss of generality, we show the setting for the SCI process first. A similar setting is expected for the NC process.

Data setting is shown in Figure 10, where each horizontal line represents the driving record of one driver. Drivers were subject to different numbers of crashes, NCs, and SCIs at different time points throughout study. Thus, it was important to record all timestamps. We focused on the actual driving time. Non-driving time when the vehicle was not in use was excluded. As illustrated in Figure 10, the driving period was divided into several phases based on the relationship with crashes: before the first crash (coded as 0), between the first and second crash (coded as 1), and after the second crash (coded as 2). Driving period was taken into account as a covariate, working as an external and independent factor on SCI intensity. To account for potential confounding and interacting effects, gender and the age of the driver when first enrolled in the study were also evaluated.

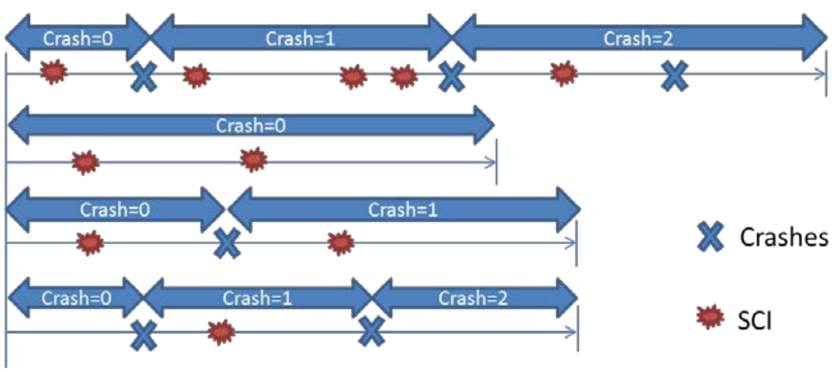


Figure 10. Chart. Data setting for intensity-based recurrent event model.

Assume an individual driver i, who is under observation in $(0, S]$, experienced n_i SCIs and c_i crashes in the entire study. Let $0 < t_1 < t_2 < \cdots < t_{n_i}$ and $0 < t_{c_1} < \cdots < t_{c_i}$ denote the event times and crash times respectively, where t_k is time to the k^{th} SCI and t_{c_k} is time to the k^{th} crash. In the 100-Car Study, only three drivers experienced more than two crashes. Thus, up to two crash effects are considered and k takes values from 0, 1, and 2. In this data setting, crash is an external time-varying covariate that has influence on SCI rate, denoted as $\mathrm{crash}(t)$, with the following piece-wise constant function:

$$\text{crash}(t) = \begin{cases} 0, & 0 \le t < t_{c_1} \\ 1, & t_{c_1} \le t < t_{c_2} \\ 2, & t_{c_2} \le t < S \end{cases}.$$

The associated counting process of SCI $\{N(t), 0 \le t\}$ records the cumulative number of SCIs generated by the process. To be specific, $N(t) = \sum_{k=1}^{\infty} I(t_k \le t)$ is the number of events occurring over the time interval $(0, t]$. The intensity function of the process gives the instantaneous probability of an event occurring at *t* and is mathematically defined as:

$$\lambda(t|H(t)) = \lim_{\Delta t \to 0} \frac{Pr(\Delta N(t) = 1|H(t))}{\Delta t}$$

where $\Delta N(t) = N(t + \Delta t^-) - N(t^-)$ is the number of events in the interval $[t, t + \Delta t]$ and $H(t) = \{N(s), 0 \le s < t\}$ represents the history of the process at time t. $H(t)$ can be considered as the mean of the cumulative number of events from $(0, t)$. This intensity function is very general and accommodates various data structure, models, and dependences within cluster structures.

In this study, we considered four alternative models, including an Andersen-Gill (A-G) model, a stratified A-G model, a frailty model, and a stratified frailty model. A simulation study was conducted to examine the performance of the proposed models. Model comparison indicated that the stratified frailty model performed best for the data set and thus it was adopted for data analysis.

3.1 ANDERSEN-GILL (A-G) MODEL

A commonly used and fundamental model in the recurrent event literature is the Poisson process model or Andersen-Gill (A-G) model (Andersen and Gill, 1982). It describes situations where events occur randomly in such a way that the number of events in non-overlapping time intervals are statistically independent. The overall intensity function of the Poisson process is:

$$\lambda_i\big(t\big|\boldsymbol{z_i}(t)\big) = \lambda_0(t)\exp(\boldsymbol{z}_{\boldsymbol{i}}'(t)\boldsymbol{\beta}) \ ,$$

where:

- $\lambda_0(t)$ is an arbitrary nonnegative integrable function.
- $\boldsymbol{z_i}(t)$ is a vector of fixed or time-varying external covariates and acts multiplicatively on the baseline.
- $\boldsymbol{\beta}$ is a vector of regression parameters of the same length as $\boldsymbol{z_i}(t)$.

The Poisson process model assumes that the probability of an event in $[t, t + \Delta t)$ depends on t but not on $H(t)$. All drivers share the same unstructured baseline function and are differentiated by covariates. Cumulative intensity, denoted as $\mu_i(t) = \int_0^t \lambda_i(s)ds$ is continuous and finite for all $t > 0$ and can be explained as the average number of events occurring in time period $(0, t]$.

The estimation of coefficients ($\hat{\beta}$) can be derived from maximizing log partial likelihood (PL), which is given as follows:

$$\log(PL) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left\{ \boldsymbol{z}_{\boldsymbol{i}}'(t)\boldsymbol{\beta} - \sum_{k \in R_{ij}} \exp(\boldsymbol{z}_{\boldsymbol{k}}'(t)\boldsymbol{\beta}) \right\},$$

where R_{ij} contains all subjects who are at risk at given time t_{ij}. Then, the baseline function can be estimated by inserting $\hat{\beta}$ into log partial likelihood and is given in

$$\widehat{\lambda_0}(t) = \frac{\sum_{i=1}^{m} Y_i(t)\, dN_i(t)}{\sum_{i=1}^{m} Y_i(t) \exp(\boldsymbol{z}_{\boldsymbol{i}}'(t)\widehat{\boldsymbol{\beta}})},$$

where $Y_i(t)$ indicates whether subject i is under study at time t, and $\sum_{i=1}^{m} Y_i(t)\, dN_i(t)$ is the total number of SCIs observed at t.

3.2 STRATIFIED A-G MODEL

It is likely that subjects are sampled from subgroups of individuals with varying intensity functions. For example, drivers who experience more crashes in the same time period are more risky than others and may behave differently in terms of SCI rate. An effective way to accommodate this situation is to stratify the baseline function. Three levels (0 crash drivers, 1 crash drivers, 2 crash drivers) were defined as strata. It is assumed that baseline functions vary among strata while coefficients are consistent. The stratification model is written as below:

$$\lambda_{ri}\left(t \middle| \boldsymbol{z}_{\boldsymbol{ri}}(t)\right) = \lambda_{r0}(t) \exp(\boldsymbol{z}_{\boldsymbol{ri}}'(t)\boldsymbol{\beta}),$$

where:

- λ_{ri}(t) is the intensity function for driver i in stratum r.
- $\lambda_{r0}(t)$ is the baseline function of stratum r.

The estimation of coefficients ($\hat{\beta}$) can be derived from maximizing log partial likelihood, which is given as follows:

$$\log(PL_{str}) = \sum_{r=1}^{R} \sum_{i=1}^{m_r} \sum_{j=1}^{n_{ri}} \left\{ \boldsymbol{z}_{\boldsymbol{ri}}'(t)\boldsymbol{\beta} - \sum_{k \in R_{rij}} \exp(\boldsymbol{z}_{\boldsymbol{rk}}'(t)\boldsymbol{\beta}) \right\},$$

where:

- r is an indicator of stratum level, $r = 1, \ldots,$ R.
- m_r is the total number of drivers in stratum r.
- n_{ri} is the number of SCIs for driver i in stratum r.
- R_{rij} contains all subjects who are at risk at given time t_{rij} in stratum r.

Then, the baseline function can be estimated by inserting $\widehat{\boldsymbol{\beta}}$ into log partial likelihood and is given in

$$\widehat{\lambda_{r0}}(t) = \frac{\sum_{i=1}^{m_r} Y_{ri} dN_{ri}(t)}{\sum_{i=1}^{m_r} Y_{ri} \exp(\boldsymbol{z}'_{\boldsymbol{ri}}(t)\hat{\beta})},$$

where Y_{ri} indicates whether subject i in strata r is under study at time t.

3.3 SHARED FRAILTY MODEL

In applications involving multiple subjects, heterogeneity is often apparent and requires consideration. Heterogeneity describes, conditioning on covariates, the variation among individual intensity rate functions. In other words, there is more within-individual variation in event occurrence than is accounted for by a Poisson process. To capture the relation of the correlated observations, it has been considered that the event times of one subject share an unobserved effect (McGilchrist & Aisbett, 1991; Nielsen, Gill, Andersen, & Sorensen, 1992). This shared, individual random effect accounts for the variation beyond conditioning on covariates.

The shared frailty model assigns a random effect, μ_i, to each subject acting multiplicatively on the Poisson intensity model. Then, the general intensity function can be written as below:

$$\lambda_i(t|\boldsymbol{z_i}(t), \mu_i) = \mu_i \lambda_0(t) \exp(\boldsymbol{z}'_{\boldsymbol{i}}(t)\boldsymbol{\beta}),$$

where the random terms, also known as frailty, $\mu_i, i = 1, \dots, m$ are taken to be independent and identically distributed (i.i.d) with mean and distribution function $G(\mu)$. Frailty gives the interpretation that individuals with $\mu_i > 1$ tend to occur at a faster rate.

There are many choices for distribution $G(\mu)$, including gamma, inverse Gaussian, lognormal (Wienke, 2010). In this report, lognormal distribution is primarily used, which means if one specifies $\gamma_i = \log(\mu_i)$, then $\gamma_i \sim Normal(0, \sigma)$. Then the intensity can be written in the format:

$$\lambda_i(t|\boldsymbol{z_i}(t), \gamma_i) = \lambda_0(t) \exp(\boldsymbol{z}'_{\boldsymbol{i}}(t)\boldsymbol{\beta} + \gamma_i).$$

For the shared frailty model, there are two commonly used methods to obtain $\hat{\beta}$: one is the expectation–maximization (EM) algorithm and the other is maximizing penalized partial log-likelihood. Therneau, Grambsch, and Pankratz (2003) have proved that the E-M algorithm solution for lognormal shared frailty models is closely linked to penalized estimation. The logarithm penalized partial likelihood (PPL) is given below:

$$Log(PPL) = \sum_{i=1}^{m} \sum_{j=1}^{n_i} \left\{ \boldsymbol{z}'_{\boldsymbol{i}}(t)\boldsymbol{\beta} - \sum_{k \in R_{ij}} \exp(\boldsymbol{z}'_{\boldsymbol{k}}(t)\boldsymbol{\beta}) \right\} - \frac{1}{2\sigma^2} \boldsymbol{\gamma}'\boldsymbol{\gamma},$$

where $\boldsymbol{\gamma} = (\gamma_1, \dots, \gamma_m)$. The maximization of this approximate likelihood is a doubly iterative process that alternates between the following two steps:

Step 1: For a fixed value of σ^2, find the best covariates estimation by maximizing the penalized partial log likelihood, $Log(PPL)$.

Step 2: For fixed values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, calculate the REML estimation of $\hat{\sigma}^2 = \frac{\widehat{\gamma'}\hat{\gamma}+\boldsymbol{trace}(\boldsymbol{H_{22}^{-1}})}{m}$, in which $\boldsymbol{H_{22}^{-1}}$ is the inverse of the second derivative matrix associated with the frailty terms.

3.4 STRATIFIED SHARED FRAILTY MODEL

The stratified shared frailty model incorporates both varying baseline functions and among-individual variation as a combination of the stratified A-G model and shared frailty model as follows:

$$\lambda_{ri}(t|\boldsymbol{z_{ri}}(t), \gamma_{ri}) = \lambda_{r0}(t)\exp(\boldsymbol{z'_{ri}}(t)\boldsymbol{\beta} + \gamma_{ri}),$$

where:

- $\boldsymbol{r}$ is an indicator of stratum level, $\boldsymbol{r} = \mathbf{1}, \dots, \mathbf{R}$.
- $\boldsymbol{\lambda_{r0}(t)}$ is the baseline function for stratum $\boldsymbol{r}$.
- $\boldsymbol{\gamma_{ri}} \sim \boldsymbol{N(0, \sigma)}$ $\boldsymbol{r} = \mathbf{1}, \dots, \boldsymbol{R}$; $\boldsymbol{i} = \mathbf{1}, \dots, \boldsymbol{m_i}$ are independent frailty terms which explain possible correlation among events for an individual.

$\hat{\beta}$ is estimated by a similar algorithm as discussed in Section 3.3. The penalized partial log-likelihood for the stratified shared frailty model is:

$$\log(PPL_{str}) = \sum_{r=1}^{R}\sum_{i=1}^{m_r}\sum_{j=1}^{n_{ri}}\left\{\boldsymbol{z'_{ri}}(t)\boldsymbol{\beta} - \sum_{k\in R_{rij}}\exp(\boldsymbol{z'_{rk}}(t)\boldsymbol{\beta})\right\} - \frac{1}{2\sigma^2}\boldsymbol{\gamma'}\boldsymbol{\gamma}.$$

For a fixed value of σ^2, the best covariates estimation is found by maximizing the penalized partial log likelihood $Log(PPL_{str})$, and then for fixed values of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, calculating the REML estimation of $\hat{\sigma}^2 = \frac{\widehat{\gamma'}\hat{\gamma}+\boldsymbol{trace}(\boldsymbol{H_{22}^{-1}})}{m}$, where $\boldsymbol{H_{22}^{-1}}$ is the inverse of the second derivative matrix associated with the frailty terms. These two steps are performed iteratively until they converge.

Then, R baseline functions are estimated from the following equation:

$$\widehat{\lambda_{r0}}(t) = \frac{\sum_{i=1}^{m_r} Y_{ri}dN_{ri}(t)}{\sum_{i=1}^{m_r} Y_{ri}\exp(\boldsymbol{z'_{ri}}(t)\hat{\beta}+\widehat{\gamma_{ri}})}.$$

3.5 MODEL FITTING EVALUATION

Cox-Snell residuals are useful for checking the overall fitting of an intensity-based model. For the case of several counting processes with intensity $\lambda_i(t), i = 1, \dots, m$, Cox-Snell residuals can be defined as:

$$\mathrm{r}_{ij} = \int_{t_{i,j-1}}^{t_{i,j}} \widehat{\lambda_i}(s)ds,\ i = 1, \dots, m;\ j = 1, \dots n_i + 1,$$

where:

- $t_{i,0}$ and t_{i,n_i+1} are the start and stop times for subject i.
- $\widehat{\lambda_i}(s)$ is the estimated intensity function.

If the model is correct, then r_{ij} should behave like a censored sample from a unit exponential distribution. Thus, a plot of the estimated cumulative intensity rate of the residuals versus the residuals should be a straight line through the origin with a slope of 1.

CHAPTER 4. SIMULATION STUDY

We conducted a simulation study to evaluate the performance of the alternative approaches. Simulation setup was analogous to the real situation and described as below:

Step 1: Driving time for 50 drivers is generated from a normal distribution with a mean of 335 and a standard deviation of 160, estimated from the 100-Car study.

Step 2: For each driver, up to two crash times are generated based on the intensity function:

$$\lambda_i(t) = \frac{1}{150}$$

The rate $\frac{1}{150}$ was selected based on the crash rate estimated from 100-Car data, which means we observed one crash in every 150 hours of driving history on average. Other baseline values were also examined, and the results were robust to the change in baseline rate. Due to space constraints, all of the results from other values are not included here. The generation of crash times assumed drivers performed similarly. Gender was not considered, nor any other external factor. Crash intensity was restricted to be constant over time. Crashes were considered to occur independently. If one driver had a crash time that was greater than the driver's study time, then the crash time was censored.

Step 3: After getting the censor time (total driving time) and crash time, the intensity function for each driver is assumed as follows:

$$\lambda_{ri}(t) = \frac{1}{c_r} * \mathrm{t}^{k_r - 1} * \exp(\beta_1(gender_{ri} = M) + \beta_2(crash_{ri}(t) = 1) + \beta_3(crash_{ri}(t) = 2) + \beta_4(crash_{ri}(t) = 1)(gender_{ri} = M) + \beta_5(crash_{ri}(t) = 2)(gender_{ri} = M) + \gamma_{ri})$$

where:

- $\boldsymbol{r = 1, 2, 3}$, which indicates stratum level. Drivers in level 1 do not experience any crashes, drivers in level 2 have only one crash, and drivers in level 3 have two crashes.
- Baseline functions vary based on two parameters: scale parameter c and shape parameter *k*. They may differ from stratum to stratum, as denoted by $\boldsymbol{c_r}$ and $\boldsymbol{k_r}$. $\boldsymbol{k_r > 1}$ indicates that the SCI rate increases over time; $\boldsymbol{k_r = 1}$ corresponds to a constant rate; $\boldsymbol{k_r < 1}$ represents decreasing trend.
- $crash_{ri}(t)$ is a time-varying crash impact indicator. It takes a value of 1 when t is between the first and second crash and 2 when t is larger than the second crash time.
- β_1 is gender effect; β_2 is the first crash effect for female drivers; β_3 is the second crash effect for female drivers;
- $\beta_2 + \beta_4$ is the first crash effect for male drivers; $\beta_3 + \beta_5$ is the second crash effect for male drivers.
- $\boldsymbol{\gamma_{ri} \sim N(0, \sigma)}$ $\boldsymbol{r = 1, \ldots, 3; i = 1, \ldots, 50}$ are independent frailty terms.

In order to cover a wide range of parameter space, 24 settings of different baseline parameters (c and k) combinations, as well as different combinations of gender, and crash effects ($\beta_1, \beta_2, \beta_3$) were tested in total. In each setting, 500 realizations were generated and two models were implemented: a stratified frailty model and a frailty model. Because of space limits, only figures and tables for selected scenarios are provided.

Figure 11 shows a coverage probability (CP) comparison between two models, where three strata share the same shape parameter, k (set as 1), but different scale parameters, c. Seven setting results with assorted combinations of c are presented. Both models perform well, with an average CP around 95% and small bias (1% to 3% difference). The stratified frailty model does not show a great benefit over the frailty model because the variation among strata is proportional, and thus can be explained through frailty terms. Figure 12 shows results from another seven settings where three strata share the same scale parameter, c (set as 0.2), but different shape parameters, k. The stratified frailty model retains a CP of around 95% while the frailty model performs poorly, with the CP being as low as 20%.

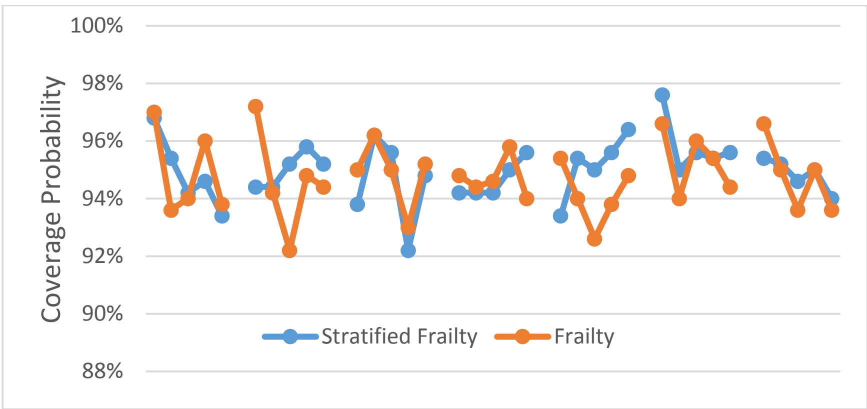


Figure 11. Graph. Coverage probability comparison (I).

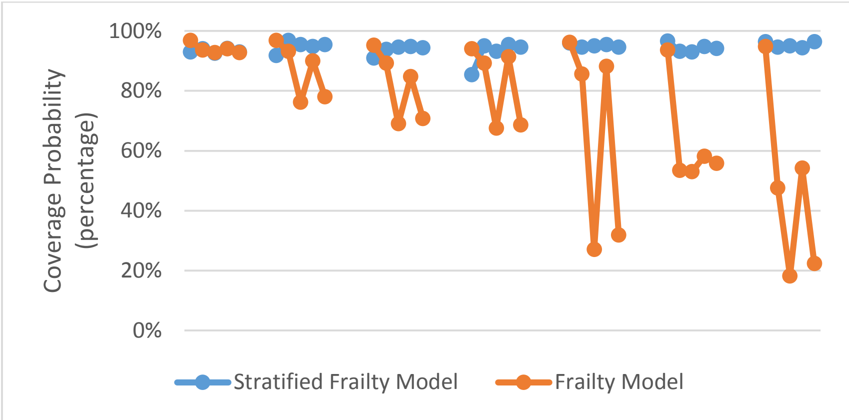


Figure 12. Graph. Coverage probability comparison (II).

To evaluate model performance at different levels of standard deviation for frailty terms, two settings are shown in Table 9: one has frailty terms follow $N(0,0.5)$, the other has frailty terms follow $N(0,1)$. Baseline functions are set to be the same. This variance of frailty terms represents a degree of heterogeneity among subjects. With higher levels of heterogeneity, we observed larger bias and empirical standard error for fixed effect estimation (β_1). Bias and empirical standard error for other effects remain similar. Compared to the low CP of the frailty model, the stratified frailty model has around a 95% CP for all effects. In Table 10, we test the performance of the stratified frailty model when there is no stratification. Three baseline functions are set with the same *c* and *k*. Although the stratified frailty model is more complicated than the situation requires, the 95% CP on average indicates a credible and stable estimation.

Table 9. Simulation result (I), *k* = (.95, 1.1, 1.22), *c* = (0.2, 0.17, 0.15).

		Stratified frailty model					Frailty model				
Parameters	**True value**	**Mean**	**Bias**	**SE%**	**SEM^**	**CP***	**Mean**	**Bias**	**SE**	**SEM**	**CP**
β_1	-0.2	-0.198	0.002	0.159	0.173	95.8	-0.196	0.004	0.198	0.205	95.6
β_2	0	-0.001	-0.001	0.059	0.062	94.8	0.04	0.04	0.06	0.059	89.4
β_3	-0.6	-0.597	0.003	0.078	0.078	95.6	-0.485	0.115	0.085	0.074	66.6
$\beta_2 + \beta_4$	-0.2	-0.198	0.002	0.069	0.068	94.2	-0.154	0.046	0.071	0.066	87
$\beta_3 + \beta_5$	-0.7	-0.693	0.007	0.082	0.083	94	-0.576	0.124	0.089	0.079	66.4
σ	0.5	0.464	-0.036	NA	NA	NA	0.621	0.121	NA	NA	NA
Sample size		6397									
β_1	-0.2	-0.161	0.039	0.323	0.301	91.6	-0.211	-0.011	0.698	0.372	92.77108
β_2	0	-0.003	-0.003	0.056	0.055	96	0.025	0.025	0.06	0.052	88.75502
β_3	-0.6	-0.595	0.005	0.083	0.069	94.2	-0.505	0.095	0.079	0.065	67.67068
$\beta_2 + \beta_4$	-0.2	-0.202	-0.002	0.061	0.06	95.2	-0.173	0.027	0.062	0.057	89.55823
$\beta_3 + \beta_5$	-0.7	-0.691	0.009	0.087	0.073	95	-0.599	0.101	0.083	0.069	65.06024
σ	1	0.937	-0.063	NA	NA	NA	1	0	NA	NA	NA
Sample size		9112									

%: Empirical standard error
^: Mean of standard error from Hessian matrix
*: Coverage probability

Table 10. Simulation result (II), *k* = (1, 1, 1), *c* = (0.2, 0.2, 0.2).

		Stratified frailty model					Frailty model				
Parameters	**True value**	**Mean**	**Bias**	**SE%**	**SEM^**	**CP***	**Mean**	**Bias**	**SE**	**SEM**	**CP**
β_1	-0.2	-0.206	-0.006	0.228	0.265	96.8	-0.207	-0.007	0.218	0.237	97
β_2	0	-0.003	-0.003	0.086	0.086	95.4	-0.004	-0.004	0.082	0.08	93.6
β_3	-0.6	-0.605	-0.005	0.121	0.116	94.2	-0.607	-0.007	0.106	0.104	94
$\beta_2 + \beta_4$	-0.2	-0.206	-0.006	0.1	0.097	94.6	-0.205	-0.005	0.092	0.091	96
$\beta_3 + \beta_5$	-0.7	-0.697	0.003	0.129	0.125	93.4	-0.698	0.002	0.115	0.113	93.8
σ	0.75	0.714	-0.036	NA	NA	NA	0.729	-0.021	NA	NA	NA
Sample size		3454									

%: Empirical standard error
^: Mean of standard error from Hessian matrix
*: Coverage probability

In summary, the stratified frailty model is capable of accommodating possible variation among groups without losing power to test for effects of interest. If drivers behave differently with various risk levels, aggregating them together will mask the effect of crashes at the individual level.

CHAPTER 5. INTENSITY-BASED RECURRENT EVENT APPROACH RESULT

In this chapter, we examine the influence of crashes on driving risk, which is measured by SCI and NC intensity (Guo and Fang, 2010). Crashes are considered an external factor that may influence driving risk. We explored the influence by comparing SCI rate in a specified window before and after a crash, analogous to the idea of comparing baseline data in Chapter 2. A higher SCI rate, in terms of number of events per hour driving, implies a larger driving risk. We applied four intensity-based models, discussed in Chapter 3, to the 100-Car data. An overall model fitting evaluation is provided for model comparison at the end of the chapter.

5.1 EXPLORATORY DATA ANALYSIS

Figure 13 illustrates the defined evaluation window before and after a crash that was applied to SCIs for exploratory data analysis. The number of SCIs and NCs in various window sizes before and after the occurrence of crashes were compared. For most cases, the observed time period before or after a crash was the same as the chosen window size. However, the observed time period may be shorter because of the issue of overlapping or if not enough driving data were available before or after that crash.

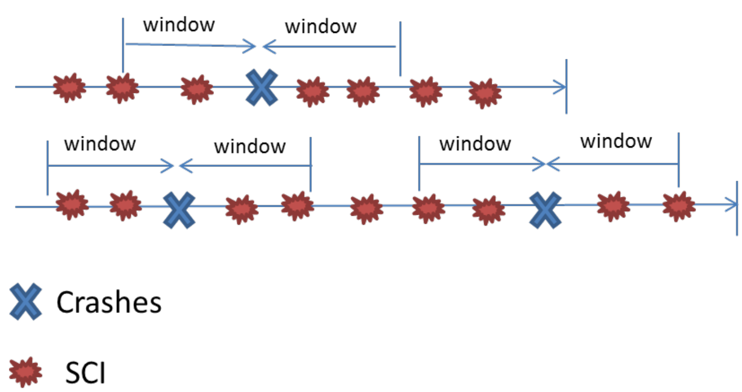


Figure 13. Chart. Before/after crash SCI collection.

Table 11 lists the ratios between the SCI rates of the before- and after-crash windows, varied across different window sizes. By comparing column 2 (or 3) to column 4 (or 5), we find that female drivers are associated with a higher SCI rate than male drivers. Figure 14 presents the SCI rate ratio after and before crashes, where SCI rate shows a tendency to be lower after crashes for male drivers within a 60-hour window. The amount of decrease diminishes as window size increases, which agrees with their distraction behavior. The influence of crashes on the SCI rate for female drivers varies from one window size to another. The decline in driving risk is not persistent. A similar analysis was also performed using calendar time, as shown in Figure 15. Within a window of between 6 days and 38 days, male drivers show a decreasing pattern comparable to the analysis using driving time. Female drivers incline to driving safer within a 38-day window.

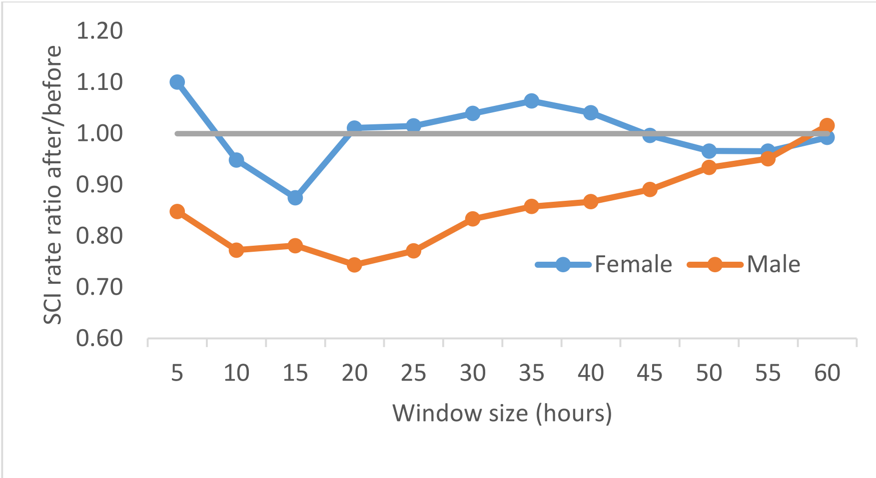


Figure 14. Graph. Ratio of SCI rate between after and before aggregated by driving hour.

Table 11. SCI rate and SCI rate ratio between after and before window.

Window size (hours)	Before female	After female	Before male	After male	Rate ratio female	Rate ratio male
5	0.33	0.36	0.34	0.29	1.10	0.85
10	0.42	0.40	0.32	0.24	0.95	0.77
15	0.42	0.37	0.32	0.25	0.87	0.78
20	0.38	0.38	0.34	0.25	1.01	0.74
25	0.37	0.37	0.35	0.27	1.01	0.77
30	0.38	0.39	0.34	0.29	1.04	0.83
35	0.38	0.41	0.34	0.29	1.06	0.86
40	0.40	0.41	0.33	0.29	1.04	0.87
45	0.41	0.40	0.32	0.29	1.00	0.89
50	0.41	0.40	0.31	0.29	0.97	0.93
55	0.40	0.39	0.31	0.29	0.97	0.95
60	0.40	0.40	0.30	0.30	0.99	1.02

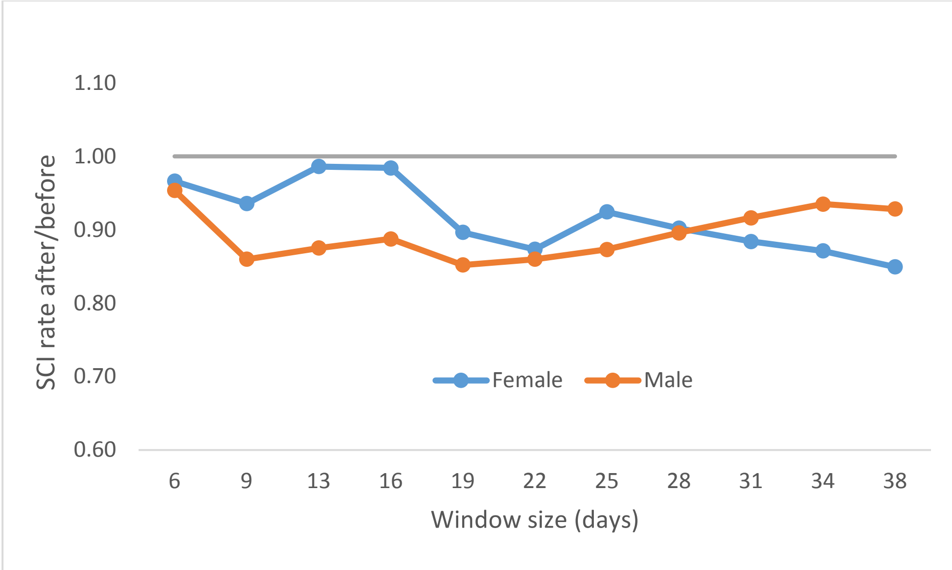


Figure 15. Graph. Ratio of SCI rate between after and before window aggregated by calendar time.

5.2 INTENSITY-BASED RECURRENT EVENT MODELING RESULT

Three covariates are incorporated into the model: gender (G), age when the driver first enrolled in the study, and crash effect based on relationship with crashes (0 for before first crash, 1 for between first and second crash, and 2 for after second crash). In order to test and estimate each crash effect, the time period related to crash time was considered as a categorical variable. Since the study lasted for one year, age was considered to be constant.

Figure 16 shows the empirical cumulative intensity plot of SCI and NC respectively by drivers with different numbers of crashes. The cumulative intensity function of *t* approximates the number of events in time interval [0, *t*]. The patterns of SCI and NC intensity functions are consistent in that drivers who experienced more crashes tended to have more events over time. It supports the idea that SCI, NC, and crash are consistent measurements of driving risk. In addition, empirical cumulative intensity functions based on crash levels are not completely proportional to each other, which reinforces the idea of stratifying drivers by number of crashes and allowing distinct baseline functions. The empirical cumulative intensity plot for SCI and NC by gender is given in Figure 17. Compared to males, females are observed with a higher number of SCIs and NCs consistently over time, analogous to the results from exploratory analysis.

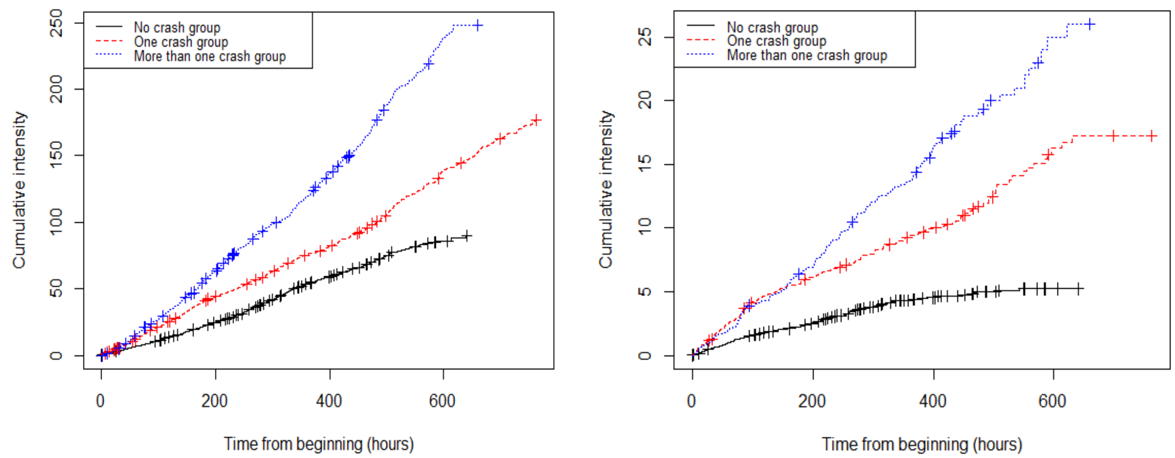


Figure 16. Graph. Cumulative intensity function of SCI (left) and NC (right) stratified by number of crashes.

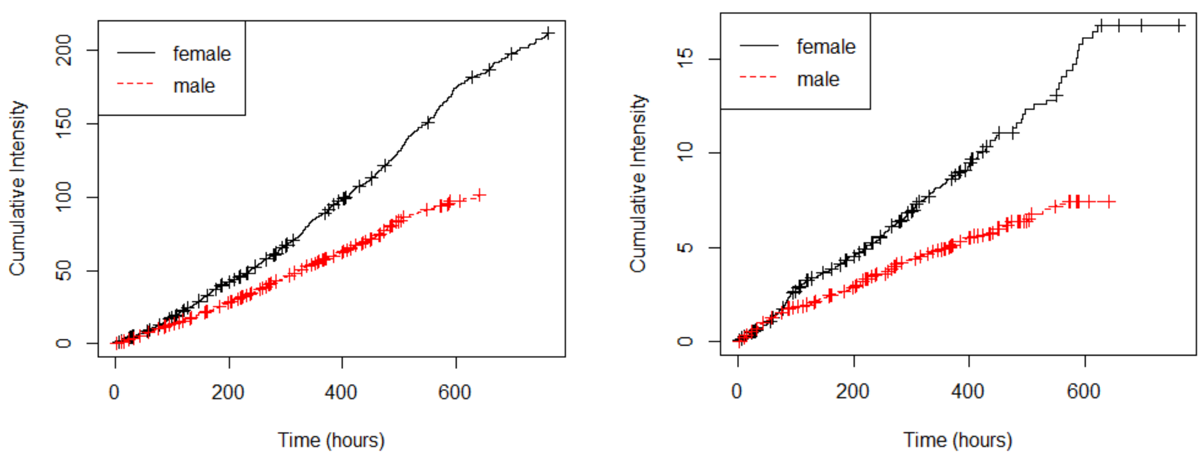


Figure 17. Graph. Cumulative intensity function of SCI (left) and NC (right) by gender.

A crash effects estimation on SCI using the A-G model, stratified A-G model, shared frailty model, and stratified shared frailty model is elaborated in Table 12. The stratified frailty model yields an intensity rate ratio of 0.8 (95% CI [0.693, 0.971]) between after the time of the first crash and before the first crash for male drivers. There is no significant first crash influence on female drivers, with an intensity rate ratio of 1.115 (95% CI [0.96, 1.296]). Unlike first crash effect, SCI risk drops sharply after the second crash for both female and male drivers, with a corresponding intensity rate ratio of 0.432 (95% CI [0.342, 0.547]) and 0.472 (95% CI [0.377, 0.59]), respectively. The detailed stratified frailty model is as follows:

$$\lambda_{ri}(t|x_{ri}) = \lambda_{0r}(t)exp\{-.008age_{ri} + 0.18(G_{ri} = F) - 0.2(crash_{ri}(t) = 1) - 0.95(crash_{ri}(t) = 2) + 0.3(crash_{ri}(t) = 1)(G_{ri} = F) + 0.32(crash_{ri}(t) = 2)(G_{ri} = F) + \gamma_{ri}\},$$

where:

- r = 0, 1, 2 indicates strata level based on the number of crashes.
- i denotes each individual in strata r.
- $\mathrm{crash}(t)$ is an indicator function defined in Chapter 3, revealing the relationship between time t and crash time.
- $\gamma_{ri} \sim N(0, \sigma)$ is a shared random term associated with driver i in strata r, accounting for correlation among events within a driver. The estimation of σ is 1.28 based on REML estimation.

The shared frailty model shows a second crash effect comparable to the stratified frailty model. Post-crash intensity is significantly lower. However, in terms of first crash effect, it is not significant for male drivers. Other than that, intensity rate increases substantially. Models without frailty terms have larger standard errors, and thus wider confidence bands on the estimation, which leads to non-significant results.

Table 12. Crash effect estimation on SCI.

Model	Contrast	Estimate	Intensity rate ratio	Lower confidence limit of intensity rate ratio	Upper confidence limit of intensity rate ratio	Pr > ChiSq
A-G	1 vs. 0 Female	0.228	1.256	0.745	2.119	0.393
A-G	2 vs. 1 Female	0.526	1.693	1.083	2.645	**0.021**
A-G	1 vs. 0 Male	0.162	1.176	0.712	1.943	0.528
A-G	2 vs. 1 Male	0.269	1.309	0.778	2.201	0.310
Stratified A-G	1 vs. 0 Female	-0.053	0.948	0.552	1.628	0.847
Stratified A-G	2 vs. 1 Female	0.056	1.058	0.691	1.619	0.795
Stratified A-G	1 vs. 0 Male	-0.382	0.683	0.378	1.233	0.206
Stratified A-G	2 vs. 1 Male	-0.065	0.937	0.572	1.533	0.795
Shared frailty	1 vs. 0 Female	0.145	1.156	1.017	1.314	**0.027**
Shared frailty	2 vs. 1 Female	-0.385	0.681	0.57	0.812	**<.0001**
Shared frailty	1 vs. 0 Male	-0.030	0.970	0.836	1.126	0.691
Shared frailty	2 vs. 1 Male	-0.337	0.714	0.598	0.852	**0.000**
Stratified frailty	1 vs. 0 Female	0.109	1.115	0.960	1.296	0.155
Stratified frailty	2 vs. 1 Female	-0.839	0.432	0.342	0.547	**<.0001**
Stratified frailty	1 vs. 0 Male	-0.199	0.820	0.693	0.971	**0.021**
Stratified frailty	2 vs. 1 Male	-0.751	0.472	0.377	0.590	**<.0001**

Table 13 lists estimations of first crash effect on NC based on four models. NCs were observed much less frequently compared to SCIs. Four female drivers experienced two or more crashes in the study, as shown in Table 4. After careful examination, only one of them had more than one NC recorded after the second crash. The rest have the second crash as their last driving record. Consequently, the estimation of the second crash effect for female drivers depends heavily on one single driver, which may lead to an individual crash influence rather than a population-wise effect. Thus, we decided to use time up to the second crash only and evaluate the first crash influence on NC.

As indicated by the stratified frailty model, the intensity rate after first crash is 0.52 times (95% CI [0.314, 0.874]) the before-crash intensity rate for male drivers. Female drivers do not show a significant decreasing trend after a crash. A similar crash influence was found for SCI data. Estimation of σ is 0.93 based on REML estimation. The detailed stratified frailty model is as follows:

$$\boldsymbol{\lambda_{ri}(t|x_{ri}) = \lambda_{0r}(t) exp\{-.008 age_{ri} + 0.20(G_{ri} = F) - 0.65(crash_{ri}(t) = 1) + 0.53(crash_{ri}(t) = 1)(G_{ri} = F) + \gamma_{ri}\}}\ .$$

The shared frailty model proposes a different crash influence compared to the stratified frailty model. Neither male nor female drivers reveal a remarkably lower post-crash intensity. Models without frailty terms have larger standard errors, and thus wider confidence bands on the estimation, which leads to non-significant results.

Table 13. Crash effect estimation on NC.

Model	Contrast	Estimate	Intensity rate ratio	Upper confidence limit of intensity rate ratio	Lower confidence limit of intensity rate ratio	Pr > ChiSq
A-G	1 vs. 0 Female	0.206	1.228	0.712	2.120	0.460
A-G	1 vs. 0 Male	0.284	1.329	0.816	2.163	0.253
Stratified A-G	1 vs. 0 Female	-0.359	0.698	0.378	1.289	0.251
Stratified A-G	1 vs. 0 Male	-0.545	0.58	0.329	1.024	0.060
Shared frailty	1 vs. 0 Female	0.222	1.249	0.815	1.913	0.308
Shared frailty	1 vs. 0 Male	-0.140	0.870	0.558	1.355	0.537
Stratified frailty	1 vs. 0 Female	-0.117	0.890	0.551	1.438	0.633
Stratified frailty	1 vs. 0 Male	-0.646	0.524	0.314	0.874	**0.013**

Figure 18 shows baseline intensity rate functions for SCI and NC estimated by the stratified frailty model. It can be concluded that intensity rate behavior among different strata is not identical, which supports the idea of stratification.

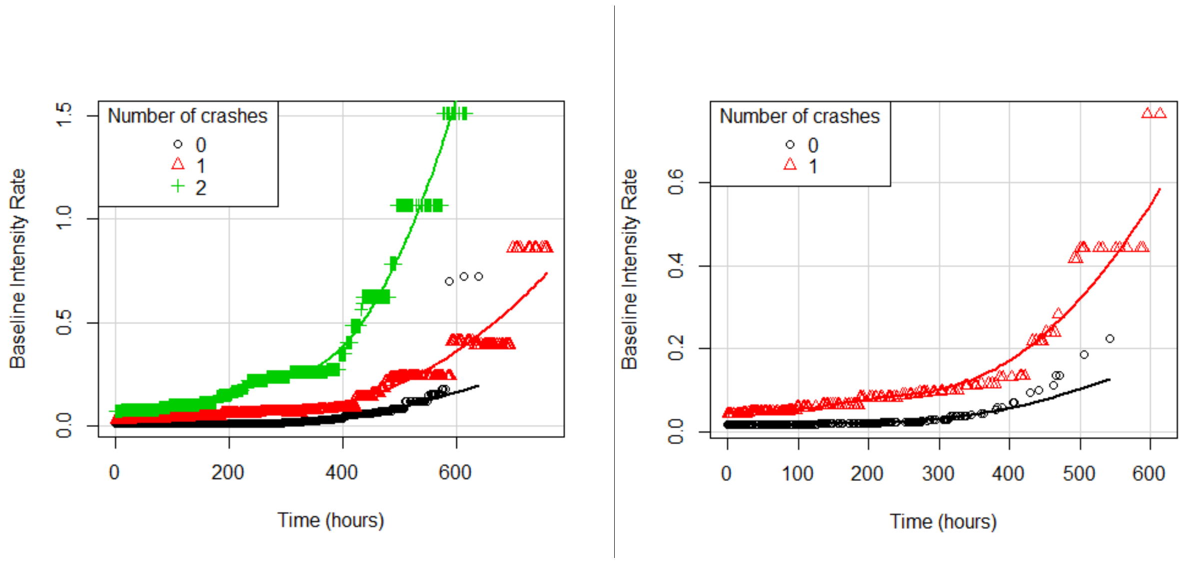


Figure 18. Graph. Baseline intensity rate estimation of SCI (left) and NC (right).

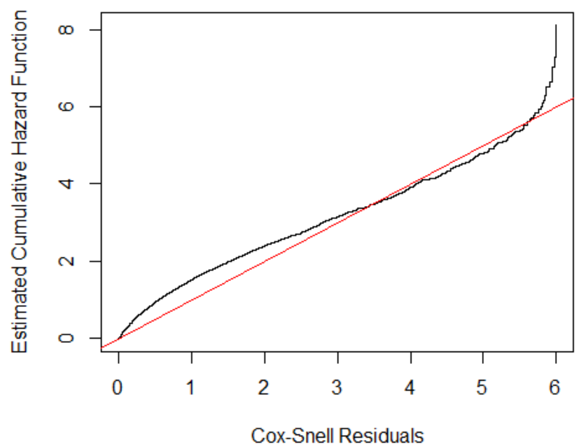
5.3 MODEL FITTING EVALUATION

The four models in the previous section do not reach an agreement on crash influence estimation, which makes model comparison important. As described in Section 3.5, Cox-Snell residuals have been widely used for overall model fitting evaluation.

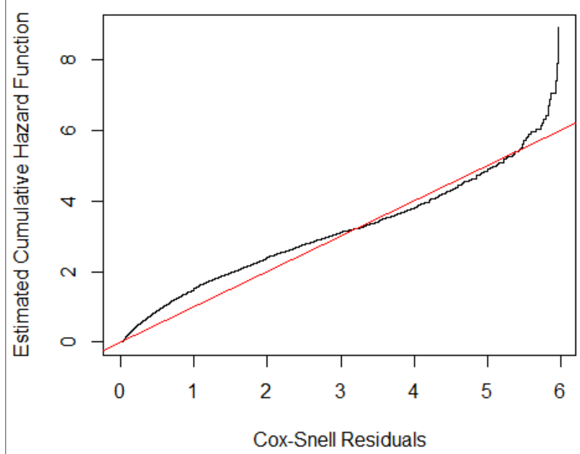
In the current study, the distribution of residuals for SCIs is heavy tailed compared to an exponential one distribution. For residuals larger than 6, the percentage varies from 1% to 2% from model to model. The probability associated with large (>6) Cox-Snell residuals is supposed to be 0.25% for an exponential one distribution. These extremely large residuals lead the fitting to depart from a straight line (Figure 20), which makes it hard to judge overall model fitting. Long intervals between two SCI events are the major source of the extreme residuals, such as a 20-hour gap compared to a 5-hour gap on average. We have examined those long gaps and it is possible that this was caused by missing event identification during the data reduction process. For this reason, we present the distribution of extremely large residuals in Table 14 and set the upper limit of the residual to 6. Figure 19 includes Cox-Snell residual plots of residuals less than 6 for four intensity-based models for SCI. It can be shown that the model fits reasonably well for the majority of data points.

Table 14. Distribution of large Cox-Snell residuals.

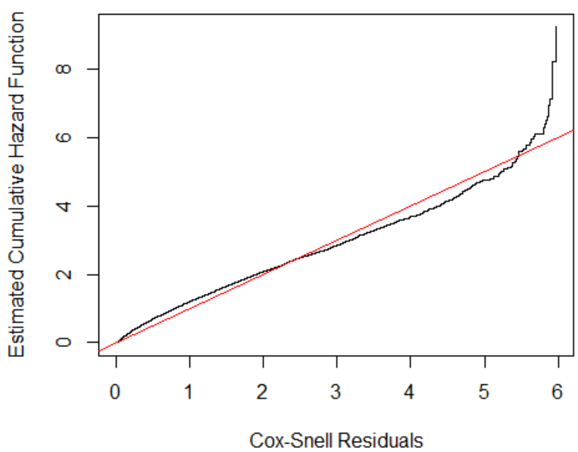
Quartile	95.0%	95.5%	96.0%	96.5%	97.0%	97.5%	98.0%	98.5%	99.0%	99.5%	100.0%
A-G	3.39	3.61	3.90	4.16	4.69	5.38	6.04	7.34	9.67	16.16	88.95
Stratified A-G	3.46	3.64	3.89	4.21	4.59	5.06	5.85	7.14	9.01	15.40	71.00
Frailty	3.40	3.55	3.77	4.02	4.32	4.64	5.12	5.68	6.76	8.73	36.73
Stratified frailty	3.35	3.54	3.71	3.97	4.21	4.52	4.88	5.43	6.37	8.11	34.72



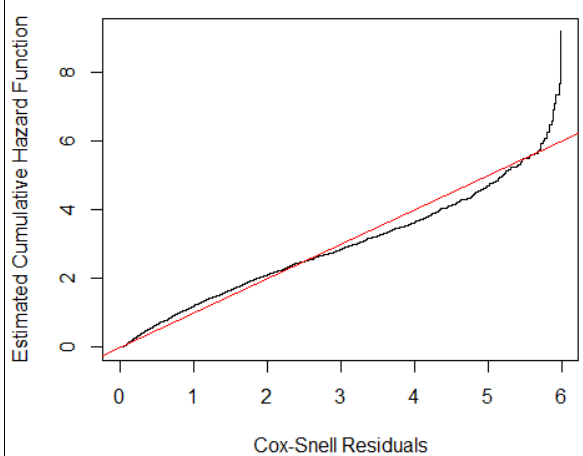
(a) A-G model



(b) Stratified A-G model

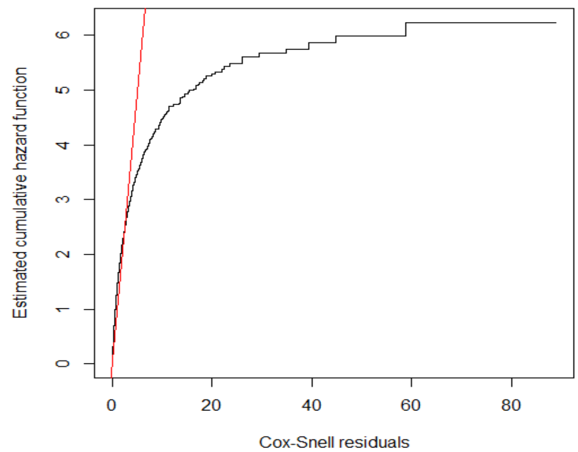


(c) Frailty model

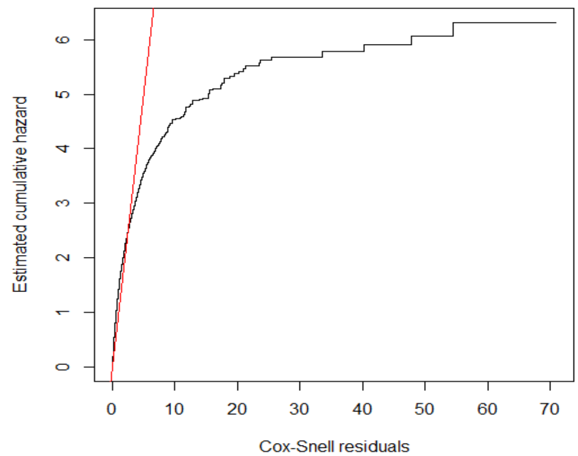


(d) Stratified frailty model

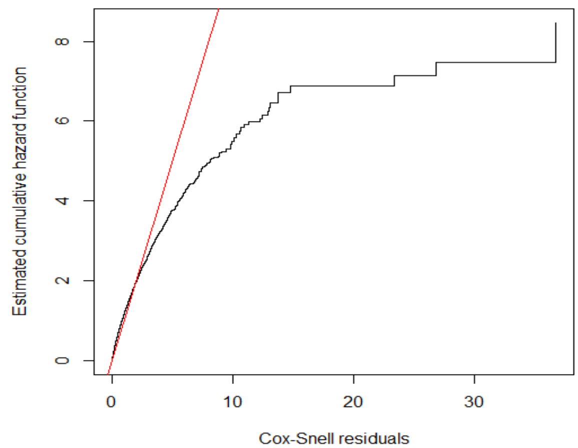
Figure 19. Graph. Plots of Cox-Snell residuals for SCI with large residuals removed.



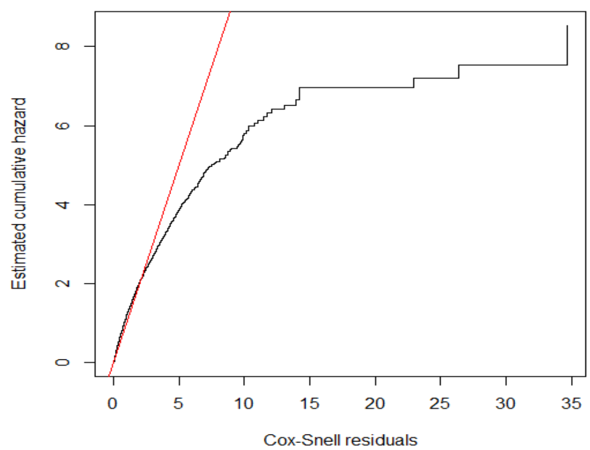
(a) A-G model



(b) Stratified A-G model



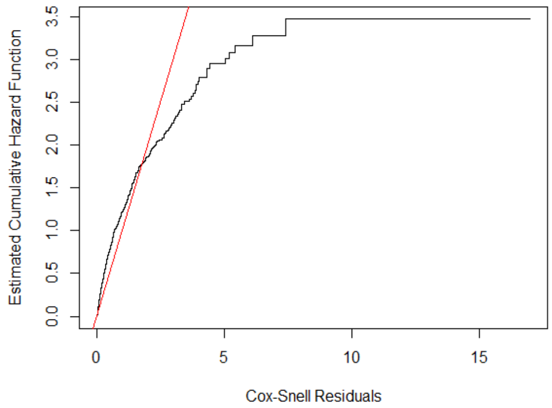
(c) Frailty model



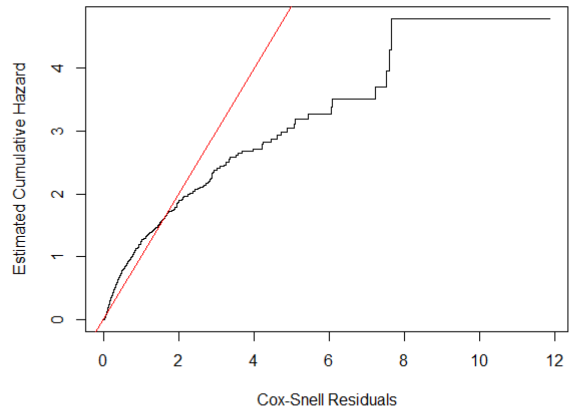
(d) Stratified frailty model

Figure 20. Graph. Plots of Cox-Snell residuals for SCI.

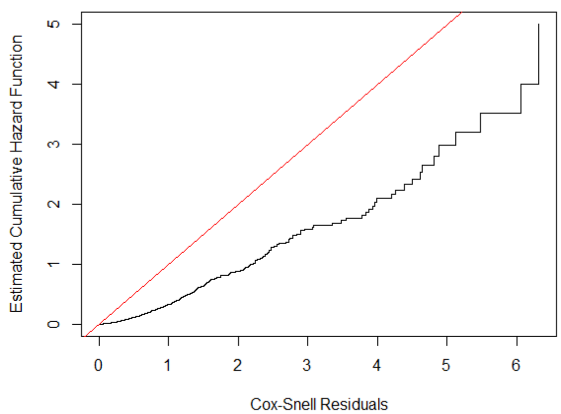
Plots of Cox-Snell residuals for NC are evaluated in Figure 21. The residual plots of the stratified frailty model are much closer to a straight line compared to the other three models, indicating the best model fitting.



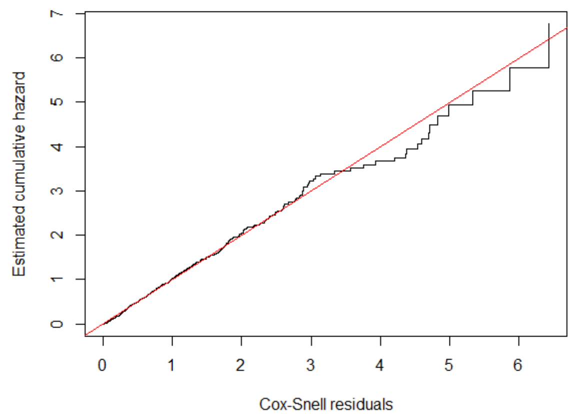
(a) A-G model



(b) Stratified A-G model



(c) Frailty model



(d) Stratified frailty model

Figure 21. Graph. Plots of Cox-Snell residuals for NC.

CHAPTER 6. SUMMARY AND DISCUSSION

This study evaluated the influences of crashes on driver distraction behavior and the driving risk using 100-Car NDS data. Driving behavior was measured by secondary driving tasks, and driving risk was measured through SCIs and NCs. Crash influence on driving behavior was evaluated with a count-based approach using a mixed binomial regression model. For testing influence on driving risk, SCIs and NCs were treated as counting processes and modeled through their intensity function. Four intensity-based models were evaluated and compared. The stratified frailty model fits the situation better than the other models and provides the best fit.

The results indicate that drivers' engagement in moderate and complex secondary tasks tends to be lower after crashes, especially within a 15-hour driving time window. This decreasing effect tends to diminish over time. Crash impact on NCs is similar for both males and females. This study confirmed that crashes have a positive effect on driver behavior. Drivers either learn from the crash experience or are more cautious while driving, which is reflected in the reduced SCI intensity within a short window after crashes. In addition, the results also indicate that female and male drivers showed different responses to crashes. Male drivers responded to both the first crash and the second crash with a lower SCI intensity after each crash. Females showed no significant response to the first crash but did show a decreased SCI intensity after the second crash. These findings provide crucial information for understanding drivers' response to dramatic driving events and can be critical for developing safety education programs and safety countermeasures.

The simulation study demonstrated that the stratified frailty model is capable of accommodating possible variation among groups without losing power to test for effects of interest. If subjects behave differently among various levels, aggregating them together will mask the effect at the individual level. We also observed robust performance of the stratified frailty model when subjects are not from different levels.

There are a couple of limitations of this study. First, the individual driver risk variation might be confounded with the observed effect. Second, the study is based on a relative small number of crashes with mild crash severity. With larger NDS data sets becoming available, such as the Second Strategic Highway Research Program (SHRP 2) Naturalistic Driving Study, more concrete evidence will be available on the impacts of crashes on driver behavior and potentially the impact of crashes by severity.

REFERENCES

af Wahlberg, A. (2012). Changes in driver celeration behavior over time: Do drivers learn from collisions? *Transportation Research Part F, 15*(5), 471–479.

Andersen, P. K., & Gill, R. D. (1982). Cox's regression model for counting processes: A large sample study. *The Annals of Statistics*, *10*(4), 1100–1120.

Chipman, M. L. (1982). The role of exposure, experience and demerit point levels in the risk of collision. *Accident Analysis & Prevention*, *14*(6), 475–483.

Cox, D. R. (1972). Regression models and life tables. *Journal of the Royal Statistical Society: Series B, 34*, 187–202.

Curry, A. E., Hafetz, J., Kallan, M. J., Winston, F. K., and Durbin, D. R. (2011). Prevalence of teen driver errors leading to serious motor vehicle crashes. *Accident Analysis & Prevention*, *43*(4), 1285–1290

Dingus, T. A., Klauer, S. G., Neale, V. L., Petersen, A., Lee, S. E., Sudweeks, J., Perez, M. A., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z. R., Jermeland, J., and Knipling, R. R. (2006). *100-Car Naturalistic Driving Study – Phase II Results of the 100-Car Field Experiment* (DOT HS 810 593). Washington, DC: National Highway Traffic Safety Administration.

Guo, F., & Hankey, J. (2009). *Modeling 100-car safety events: A case-based approach for analyzing naturalistic driving data* (Report No. 09-UT-006, 2009). Blacksburg, VA: National Surface Transportation Safety Center for Excellence.

Guo, F., Klauer, S. G., Hankey, J. M., & Dingus, T. A. (2010). Near crashes as crash surrogate for naturalistic driving studies. *Transportation Research Record: Journal of the Transportation Research Board*, *2147*, 66–74.

Guo, F., Fang, Y. (2013). Individual driver risk assessment using naturalistic driving data. *Accident Analysis & Prevention*, *61*, 3-9

Kaneko, T., & Jovanis, P. P. (1992). Multiday driving patterns and motor carrier accident risk: A disaggregate analysis. *Accident Analysis & Prevention*, *24*, 437–456.

Levy, D. T. (1990). Youth and traffic safety: The effects of driving age, experience, and education. *Accident Analysis & Prevention*, *22*(4), 327–334.

Lin, M. R., Huang, W., Hwang, H. F., Wu, H. D., & Yen, L. L. (2004) The effect of crash experience on changes in risk taking among urban and rural young people. *Accident Analysis & Prevention*, *36*, 213–222.

Lucas, J. L. (2003). Drivers' psychological and physical reactions after motor vehicle accidents. *Transportation Research Part F: Traffic Psychology and Behavior*, *6*(2), 135–145.

McGilchrist, C. A., & Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics, 47*, 461–466.

Nielsen, G. G., Gill, R. D., Andersen, P. K., & Sorensen, T. I. A. (1992). A counting process approach to maximum likelihood estimation in frailty models. *Scandinavian Journal of Statistics*, *19*, 25–43.

Rajalin, S., & Summala, H. (1997). What surviving drivers learn from a fatal road accident? *Accident Analysis & Prevention*, *29*(3), 277–283.

Treat, J. R. (1980). A study of precrash factors involved in traffic accidents. *HSRI Research Review*, *10*(6), 35.

Therneau, T. M., Grambsch, P. M., & Pankratz, V. S. (2003). Penalized survival models and frailty. *Journal of Computational and Graphical Statistics*, *12*(1), 156–175.

Waller, P. F., Elliott, M. R., Shope, J. T., Raghunathan, T. E., & Little, R. J. (2001). Changes in young adult offense and crash patterns over time. *Accident Analysis & Prevention*, *33*(1), 117–128.

Wienke, A. (2010). *Frailty models in survival analysis*. Chapman & Hall/CRC Biostatistics Series. Hoboken: CRC Press.