Accident Analysis and Prevention 132 (2019) 105277

Contents lists available at ScienceDirect

Accident Analysis and Prevention

journal homepage: www.elsevier.com/locate/aap

Exploring microscopic driving volatility in naturalistic driving environment prior to involvement in safety critical events—Concept of event-based driving volatility

Behram Wali[a], Asad J. Khattak[b,*], Thomas Karnowski[c]

[a] Postdoctoral Scholar, Massachusetts Institute of Technology, USA
[b] Beaman Distinguished Professor & Transportation Program Coordinator, University of Tennessee, Knoxville, TN 37996, USA
[c] Research Staff Member, Imaging, Signals, and Machine Learning Group, Oak Ridge National Laboratory, TN, USA

ARTICLE INFO

Keywords:
Naturalistic driving studies
Event-based volatility
Vehicular jerk
Crash
Near-crash
Crash propensity
Crash risk
Fixed and random parameters
Logit models

ABSTRACT

The sequence of instantaneous driving decisions and its variations, known as driving volatility, prior to involvement in safety critical events can be a leading indicator of safety. This study focuses on the component of "driving volatility matrix" related to specific normal and safety-critical events, named "event-based volatility." The research issue is characterizing volatility in instantaneous driving decisions in the longitudinal and lateral directions, and how it varies across drivers involved in normal driving, crash, and/or near-crash events. To explore the issue, a rigorous quasi-experimental study design is adopted to help compare driving behaviors in normal vs unsafe outcomes. Using a unique real-world naturalistic driving database from the 2[nd] Strategic Highway Research Program (SHRP), a test set of 9593 driving events featuring 2.2 million temporal samples of real-world driving are analyzed. This study features a plethora of kinematic sensors, video, and radar spatio-temporal data about vehicle movement and therefore offers the opportunity to initiate such exploration. By using information related to longitudinal and lateral accelerations and vehicular jerk, 24 different aggregate and segmented measures of driving volatility are proposed that captures variations in extreme instantaneous driving decisions. In doing so, careful attention is given to the issue of intentional vs. unintentional volatility. The volatility indices, as leading indicators of near-crash and crash events, are then linked with safety critical events, crash propensity, and other event specific explanatory variables. Owing to the presence of unobserved heterogeneity and omitted variable bias, fixed- and random-parameter discrete choice models are developed that relate crash propensity to unintentional driving volatility and other factors. Statistically significant evidence is found that driver volatilities in near-crash and crash events are significantly greater than volatility in normal driving events. After controlling for traffic, roadway, and unobserved factors, the results suggest that greater intentional volatility increases the likelihood of both crash and near-crash events. A one-unit increase in intentional volatility is associated with positive vehicular jerk in longitudinal direction increases the chance of crash and near-crash outcome by 15.79 and 12.52 percentage points, respectively. Importantly, intentional volatility in positive vehicular jerk in lateral direction has more negative consequences than intentional volatility in positive vehicular jerk in longitudinal direction. Compared to acceleration/deceleration, vehicular jerk can better characterize the volatility in microscopic instantaneous driving decisions prior to involvement in safety critical events. Finally, the magnitudes of correlations exhibit significant heterogeneity, and that accounting for the heterogeneous effects in the modeling framework can provide more reliable and accurate results. The study demonstrates the value of quasi-experimental study design and big data analytics for understanding extreme driving behaviors in safe vs. unsafe driving outcomes.

1. Introduction

The Global Status Report on Road Safety indicates that an estimated 1.25 million people annually die in road traffic crashes (RTCs) and approximately 50 million sustain injuries (WHO, 2015). This high toll of annual RTCs imposes substantial costs on our societies, with annual crash costs totaling to $240 billion within the United States (NHTSA, 2015). Among other factors, driving behavior and/or human factors in

* Corresponding author.
E-mail addresses: bwali@mit.edu (B. Wali), akhattak@utk.edu (A.J. Khattak), karnowskitp@ornl.gov (T. Karnowski).

https://doi.org/10.1016/j.aap.2019.105277
Received 4 August 2017; Received in revised form 17 July 2019; Accepted 19 August 2019
Available online 09 September 2019
0001-4575/

general are considered a leading cause of RTCs (Dingus et al., 2006; Liu and Khattak, 2016; FHWA, 2017). Recent statistics suggest that a majority of crashes are influenced in a major way by driver behavior (FHWA, 2017). Thus, for several decades researchers have focused on understanding the behavioral correlates of crash risk or crash propensity. For the most part, the analysis of behavioral factors correlated with crash propensity mainly builds upon questionnaire surveys and/or controlled experiments (Schneider et al., 2001, 2004; Machin and Sankey, 2008; Ivers et al., 2009; Antonopoulos et al., 2011; Qu et al., 2014; Scott-Parker and Oviedo-Trespalacios, 2017). While analysis of such a nature is important for identifying driver-related factors associated with higher crash risk, it does not shed light on the actual driving tasks and/or decisions that typically precede drivers' involvement in a crash (Kim et al., 2016). As such, it is crucial to gain insights regarding the sequence of microscopic instantaneous driving decisions (e.g., acceleration/deceleration) preceding drivers' involvement in a near-crash or crash situation. However, an analysis of such a nature was not possible until very recent mainly due to data unavailability.

The rapid technological advancements in recent years have enabled collection of huge amounts of spatiotemporal data about vehicle and human movement. With recent innovations ranging from realization of vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) technologies such as Dedicated Short Range Communication (DSRC) and WI-FI, to continuous video and radar surveillance, the collection of countless terabytes of real-world driving data is now a reality (Campbell, 2012; Henclewood, 2014). The generated large-scale empirical data by such technologies has significant potential in facilitating deeper understanding of instantaneous driving decisions prior to occurrence of unsafe outcomes, such as crashes (Kamrani et al., 2017). Relevant in this regard is the concept of "driving volatility" that captures the extent of variations in driving, especially hard accelerations/braking and jerky maneuvers and can serve as a leading indicator of near-crash and crash events (Khattak et al., 2015; Liu et al., 2015a; Wang et al., 2015; Liu and Khattak, 2016). Broadly, through monitoring and analysis of real-world driving data, proactive safety approaches can be formulated by giving warnings and alerts to drivers and which can reduce such volatility potentially improving safety.

With these forethoughts in mind, the main objective of this study is to investigate correlations between driving volatility and crash propensity. Crash propensity is usually defined as the tendency of a driver to get involve in an unsafe outcome, and which is mostly defined as a crash (Abdel-Aty and Pande, 2005; Christoforou et al., 2011). However, an important goal within this broader unsafe outcome perspective is to identify and analyze situations resulting in near-crashes (or near misses), as such "close calls" may foreshadow actual future crashes. Thus, in this study an unsafe outcome is defined as a crash or near-crash event. To explore the issue, a rigorous observational study design is adopted to help compare driving behaviors in normal vs unsafe outcomes. Such a study design is crucial to understanding the microscopic extreme driving behaviors in unsafe events and normal driving events. As such, the study builds upon a unique Naturalistic Driving Study (NDS) database of thousands of driving events in which a driver was involved in a safe driving event (baseline or normal event), crash event, or a near-crash event. For all such driving events, large-scale microscopic instantaneous driving decision data prior to involvement in both safe and unsafe outcomes are analyzed, and volatility indices created based on different driving performance measures. The volatility indices are then linked with crash propensity, event specific variables such as drivers' pre-event maneuvers and behaviors, secondary tasks, roadway and traffic flow related factors. Both simple and advanced statistical methods are employed to generate new knowledge critical to formulation of proactive warnings and alerts in case an unsafe outcome is anticipated. From a methodological perspective, discrete choice models are estimated for modeling crash propensity as a function of several variables including driving volatility, and which accounts for important issues of unobserved heterogeneity (Mannering et al., 2016) and omitted variable bias (discussed later in detail).

2. Literature Review

2.1. Crash frequency, crash rate and associated factors

At an aggregate level, a broad spectrum of studies have established relationships between crash frequency (or crash rates) and traffic related factors (Ivan et al., 2000; Martin, 2002; Qin et al., 2004; Anastasopoulos et al., 2008; Ma et al., 2017; Sarker et al., 2017), roadway factors (Qin et al., 2004; Anastasopoulos et al., 2008; Dong et al., 2014; Ma et al., 2017), built-environment factors (Ivan et al., 2000; Lee et al., 2014; Chen, 2015), weather related factors (Anastasopoulos et al., 2008; Dong et al., 2014; Hassan et al., 2017), and driver behavior (Lee et al., 2014; Hassan et al., 2017). Among other factors, driver behavior (or risky driving) is concluded to be the main contributing factor for crashes (Neyens and Boyle, 2007; Boyle et al., 2008; Lee and Abdel-Aty, 2008; Yan et al., 2008; Lee et al., 2014; Hassan et al., 2017). As a surrogate of driving behavior, aggregate measures such as residence characteristics of drivers, socio-economic and age-related factors, and/or ticket violations are usually used to relate driving behavior with crash frequency (Ivan et al., 2000; Weng and Meng, 2012; Lee et al., 2014; Mitchell et al., 2014; Liu et al., 2015b; Hassan et al., 2017). Traditional police crash report forms do not typically include detailed information about driving behavior related factors. As such, studies have also used self-reported questionnaire surveys to investigate (or infer) links between driving behavior and crash risk (Mannering, 2009; Tronsmoen, 2010; Smorti and Guarnieri, 2014; Hassan et al., 2017; Scott-Parker and Oviedo-Trespalacios, 2017). Different driver related factors (age, gender, nationality), vehicle types, mobile-phone use, drink driving, risk perception, and safety attitudes are found correlated with crash involvement (Tronsmoen, 2010; Hassan et al., 2017; Scott-Parker and Oviedo-Trespalacios, 2017).

As opposed to using crashes as a safety tool, near-crash traffic events are usually acknowledged but not used as safety tools. This is primarily due to the degree of subjectivity involved in identification of such events (Hayward, 1972). However, for drawing a complete picture, it is important to analyze situations that may result in near-crashes as such events are typically precursors to actual crashes. Collectively, while the previous studies provided information about important variables related to crash occurrence and/or crash rates, crucial information is missing regarding pre-crash critical vehicle maneuvers or operation. An understanding of the actual driving mechanism related with occurrence of crash or near-crash event is crucial for designing actionable proactive behavioral countermeasures.

2.2. Crash surrogates

There has been an extensive body of literature accumulated over the last 40 years or more concerning identification and development of measures related to "crash proximity", typically termed as crash surrogates. The idea is to identify situations that can potentially result in a traffic crash. As traffic crashes itself are rare events, the primary goal of developing crash surrogates has been to identify road sites with potential for safety improvement, or to identify sites where crashes may not have historically happened but perhaps are waiting to happen. Traditionally, the traffic conflict technique has been widely used to develop crash surrogates at intersections (Perkins and Harris, 1968; Salman and Al-Maita, 1995). Once, conflicts are identified, prediction models are then developed to relate traffic conflicts to traffic volumes and total accidents (Salman and Al-Maita, 1995). While useful, the traditional traffic conflict technique usually does not rely on real-world microscopic driving data, rather it relies on police-reported crashes and conflict data obtained at specific intersections. Usually, tight experimental or operational protocols are followed typically in a simulator or test track environments (Wu and Jovanis, 2013). As such, observability of unsafe events and its generalization to other driving contexts (not captured in the experiment design) is challenging. The emergence of naturalistic driving data in recent years offers the invaluable opportunity to observe both crash and near-crash events in a real-world setup (Shrestha et al., 2017). Among many other unique features, vehicles in a naturalistic driving or a connected vehicles context are instrumented with a plethora of sensing, radar, and video technologies that provide second-by-second data on vehicle kinematics and biometrics over a course of a trip (McDonald, 2017). Thus, the raw high resolution vehicle kinematics data can be processed and analyzed to detect "unusual" or "unsafe" driving events, and which can be termed as crash surrogates (Wu and Jovanis, 2013). Also, drivers in a naturalistic driving or connected vehicles setup are supposed to drive as they normally would, i.e., without specific operational and experimental protocols, and not in a simulator or test track environments. Typically, kinematic search criteria, based on longitudinal acceleration, lateral acceleration, yaw rate (to name a few), are developed and employed to detect such unsafe events. For an exhaustive overview of developing crash surrogates via naturalistic driving data, see (Jovanis et al., 2011; Wu and Jovanis, 2012, 2013), and the references therein. As a new type of crash surrogate, the present study presents the new concept of event-based volatility in a naturalistic or connected vehicles environment. In doing so, the overall concept of driving volatility is presented first (see below).

2.3. Real-world driving data and concept of driving volatility

Emerging technologies such as vehicle-to-vehicle and vehicle-to-infrastructure communication, and naturalistic driving studies facilitate the collection of high frequency real-world driving data. Towards this end, recent studies utilized real-world driving data integrated with sensor and radar technologies to propose the concept of "driving volatility", which is a measure of driving practice for characterizing instantaneous driving decisions, importantly extreme driving behaviors, and the dynamics of regimes in a typical driving profile (Liu et al., 2015a; Wang et al., 2015; Khattak and Wali, 2017). Specifically, such sensing captures the extent of variations in driving, especially hard accelerations/braking and jerky maneuvers (Khattak et al., 2015; Liu et al., 2015a; Wang et al., 2015; Liu and Khattak, 2016; Khattak and Wali, 2017; Murphey, 2008). The fundamental idea is to capture the magnitudes and amount of variations in driving decisions as larger variations (or heterogeneity) in microscopic decisions by the driver cannot only influence their own safety but also the operations of sorrounding traffic. For instance, a recent study developed a fundamental understanding of instantaneous driving decisions, and to distinguish normal from anomalous driving (Khattak and Wali, 2017). By conceptualizing microscopic driving decisions into distinct yet unobserved regimes, the focus was to quantify volatility in each regime and how driving regime allocation can be probabilistically mapped to the sorrounding traffic contexts (Khattak and Wali, 2017). A dynamic Markov regime switching methodology was presented to predict what a driver will do in short term in a connected vehicles environment, and which is fundamental to the development of driving feedback devices and control assist systems (Khattak and Wali, 2017). Compared to traditional behavioral measures (such as age, education, gender, socio-economics), the concept of individual level driving volatility provides personalized and actionable information for developing driving feedback devices, warning and control assists systems (Wang et al., 2015; Liu and Khattak, 2016; Khattak and Wali, 2017). The next section presents a brief synthesis of the new concept of "driving volatility matrix" that characterizes volatility in microscopic driving decisions at several hierarchies of the driving ecosystem.

2.4. Driving volatility and unsafe outcomes

While the afore-mentioned studies characterized driving practices by using rigorous data analytic methodologies (Liu et al., 2015a; Wang et al., 2015; Liu and Khattak, 2016; Khattak and Wali, 2017), the volatility was not linked with unsafe outcomes such as crashes. In this regard, recent studies by Kamrani et al. (2017) and Wali et al. (2018) extended the concept of driving volatility to specific locations (location-based volatility) and demonstrated how high resolution connected vehicles based driving data can be linked with historical crashes for designing proactive safety management tools (Kamrani et al., 2017; Wali et al., 2018b). Furthermore, in simulation-assisted frequentist as well as in Full-Bayesian setup, the studies demonstrated that the relationship between driving volatility and crashes vary across different locations (unobserved heterogeneity), and that it is necessary to control for omitted variables while establishing relationships between driving volatility and crash outcomes (Kamrani et al., 2017; Wali et al., 2018b). Furthermore, different statistical measures of location (intersection) specific volatilities were proposed to quantify location-based volatilities in connected (instrumented) vehicles environment (Wali et al., 2018b). In similar work, Kim et al. (2016) conducted an exploratory study to analyze the association between rear-end crash propensity and microscale driving behavior (Kim et al., 2016). Correlational statistics were computed and spatial distributions explored (Kim et al., 2016). All the three studies concluded that hard deceleration rates are associated with rear-end crashes on freeway ramps (Kim et al., 2016) and total crashes at signalized intersection (Kamrani et al., 2017; Wali et al., 2018b), and innovative proactive safety strategies were discussed (Kim et al., 2016; Kamrani et al., 2017; Wali et al., 2018b).

2.5. Research gap

The aforementioned studies contributed by providing data analytic and Bayesian statistical methodologies to link large-scale driving behavior data with historical crashes. However, important research gaps exist. First, these studies were aggregated level in the sense that location specific (intersections or freeway on/off ramps) driving behavior data were used to explain historical crashes at such locations (Kim et al., 2016; Kamrani et al., 2017; Wali et al., 2018b). Thus, insights regarding how individual driver's instantaneous driving decisions can be related to his/her crash involvement cannot be obtained. Second, due to data unavailability, short duration of high frequency driving data (two months' data in (Kamrani et al., 2017; Wali et al., 2018b) and three months' data in (Kim et al., 2016)) were used to explain multi-year crashes. Third, only crashes were used as tools for characterizing safety and near-crashes were not considered (Kim et al., 2016; Kamrani et al., 2017; Wali et al., 2018b). Given these research gaps, this study extends the concept of driving volatility to specific normal and safety-critical events, thus named "event-based volatility." Finally, we believe that methodological issues related to unobserved heterogeneity and omitted variable bias should be properly accounted for in analyses of such a nature. That is, it is important to control for unobserved factors that may influence unsafe outcomes but are not observed in data. If such unobserved factors could be included in a model, the correlations between driving volatility and unsafe outcomes can change, e.g., the magnitude or statistical significance of the relationship can change. The study by (Kim et al., 2016) was descriptive in nature and did not account for unobserved heterogeneity.

2.5.1. Research objective and contribution

Given the prevalent gaps in the literature, the objectives of this study are, (1) to characterize volatility in instantaneous driving decisions in normal driving events, crash events, and near-crash events, (2) to examine how volatility in instantaneous driving decision vary across drivers involved in normal, crash, and/or near-crash events, and (3) to understand correlations between driving volatility (intentional and unintentional) and crash propensity after controlling for other factors, unobserved heterogeneity and omitted variable bias.

These objectives seek to gain a fundamental understanding of instantaneous short-term driving decisions prior to involvement in unsafe outcomes, and therefore reveal how we can map driving volatility to drivers' involvement in unsafe outcomes, i.e., concept of "event-based volatility" is introduced. Crash propensity is defined as likelihood of drivers' involvement in crash- or near-crash events, compared to normal (baseline) driving events. Such an analysis is critical for designing proactive behavioral countermeasures as it can highlight moments of volatile (potentially unsafe) instantaneous driving decisions prior to involvement in an unsafe outcome.

For thousands of driving events in naturalistic driving studies, large-scale microscopic instantaneous driving decision data prior to involvement in both safe and unsafe outcomes are analyzed and volatility indices created based on different driving performance measures. Careful attention is given to the issue of intentional vs. unintentional volatility (discussed later in detail). Advanced statistical analysis is conducted to relate different intentional and unintentional driving volatility measures to crash propensity, generating new knowledge critical to formulation of proactive warnings and alerts. Given the important methodological concerns of unobserved heterogeneity and omitted variable bias, fixed- and random-parameter discrete choice models are developed to reach reliable conclusions.

3. Methodology

3.1. Conceptual framework

To understand driving volatility prior to involvement in safety critical events, detailed microscopic data on instantaneous driving decisions are needed (Liu and Khattak, 2016). The recently concluded SHRP2 Naturalistic Driving Study provides relevant data (TRB, 2013). Fig. 1 presents a conceptual framework that describes the overall study structure. The "Event Detail Table" in the framework consists of a table of critical safety events and baseline events, ranging from 20 s long to 30 s long. Specifically, 20 s of microscopic driving data are available for baseline events, whereas, 30 s data are available for safety-critical events (such as near-crashes and crashes). These events have been manually reviewed and categorized into a set of 74 descriptive variables (VTTI Insight Web Site). Each event is also accompanied by a set of measurements from the NDS sensors, sampled at 10 frames / second. The EDT used in this work was obtained in September 2014.

By using large-scale data analytic techniques, a unique aspect of the current study is to combine traditional and emerging data sources in a meaningful way critical to development of proactive safety tools and behavioral countermeasures. In this study, both safety critical events (crash/near-crash) and baseline events (normal driving) are considered (defined later). This is important because understanding driving behavior in safety critical events vis-à-vis normal driving events can help determine meaningful behavioral differences. By using several performance measures, the magnitudes and extent of variations (termed as driving volatility) in driver's performance prior to involvement in safety critical and/or baseline events are quantified (Fig. 1). As a next step, the microscopic driving volatility information is then linked with data in the event detail table that provides event specific data such as pre-crash maneuvers, road inventory, weather factors, and traffic factors. By using advanced statistical methods, correlations between driving volatility and crash propensity are then explored. We hypothesis a positive correlation between driving volatility and crash propensity. Any correlation, if exists, can shed light on microscopic driving decisions, and how such decisions influence roadway safety (Fig. 1).

3.2. The concept of driving volatility matrix

The emerging data streams from connected vehicles, naturalistic driving sensor and telematics, and traditional transportation data can be combined to prospect opportunities for engineering smart and proactive transportation systems. Following upon the discussion presented in section 2.2, the key motivation behind analyzing driving volatility is to help predict what drivers will do in the short term. Consequently, Wali (2018) developed the new concept of "driving volatility matrix" which takes a systems approach to operationalizing driving volatility at different levels (Wali, 2018). In particular, the focus is to conceptualize and model the extent of variations in driving at several hierarchies of the real-world traffic ecosystem, i.e., 1) trip-based volatility, 2) event-based volatility, 3) location-based volatility, and 4) driver-based volatility, thus termed as driving volatility matrix. For completeness, we present a brief overview of the concept of driving volatility matrix in Fig. 2 and how it relates to the present study. For details, see (Wali, 2018).

The different elements of volatility matrix are illustrated at a very basic level in Fig. 2 in space-time dimension. The x-axis is space dimension (e.g., a road facility containing road segments and intersections) and y-axis is time dimension. Trip-based volatility relates to the extent of variations in microscopic driving decisions at an individual trip level. Referring to the first block in Fig. 2 (indicated by "A"), assume three persons initiate a trip from reference point (home) in Fig. 2 to a grocery store. The hypothetical speed profiles (in space-time dimension) are shown in Fig. 2. If we have microscopic driving behavior and telematics data (high-resolution speed, acceleration/deceleration, etc.) at our disposal for these three trips, then we can develop and apply

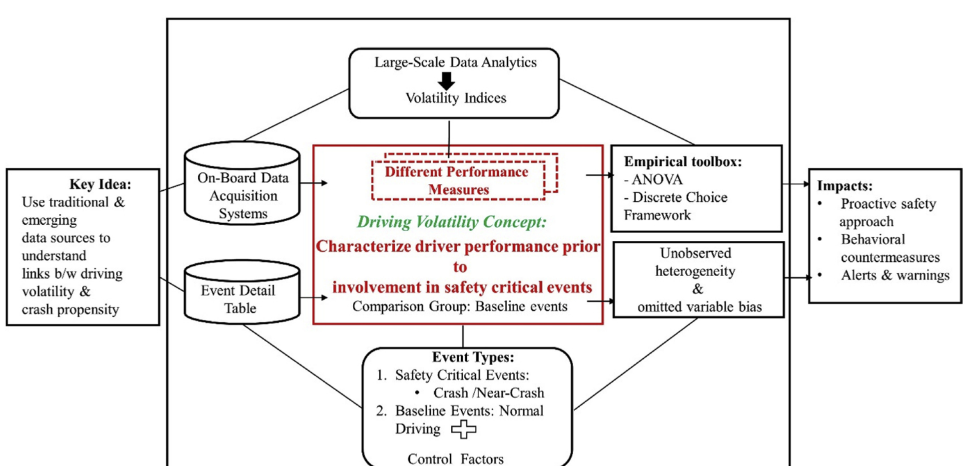

Fig. 1. Conceptual Framework.

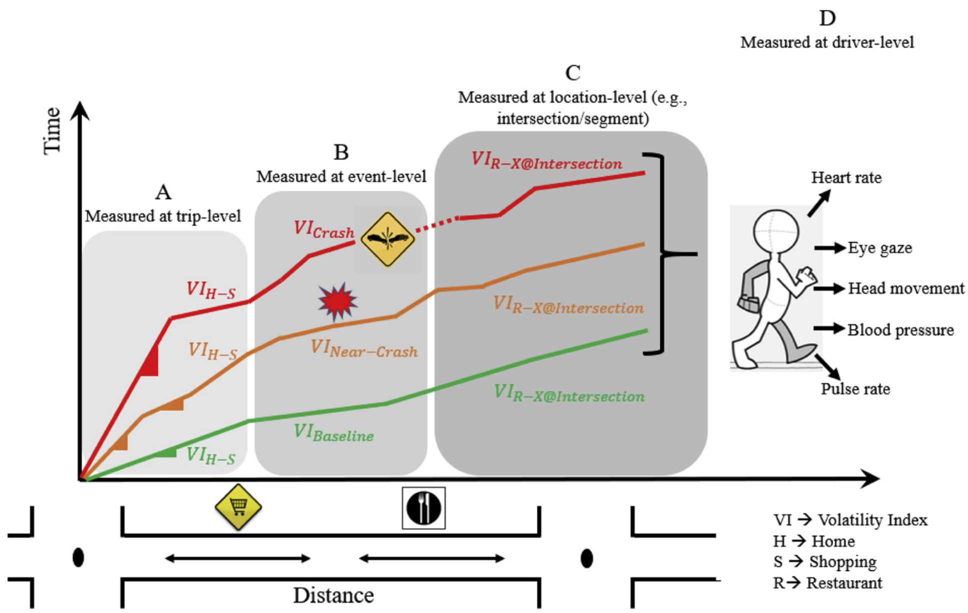

Fig. 2. Conceptualization of Driving Volatility Matrix in space-time dimension.
Note: (A → trip-based volatility, B→ event-based volatility, C→ location-based volatility, D→ driver-based volatility)

rigorous data analytic methodologies to quantify the extent of variations in microscopic driving decisions, and eventually develop volatility indices for each of the three trips (Fig. 2). At a very basic level, this is referred to as "trip-based volatility" where the volatility indices will quantify variations in driving decisions at individual trip level (Khattak and Wali, 2017).

At the next level, the idea of "event-based volatility" (as is presented in this study) introduces the notion that volatility in longitudinal and lateral directions prior to involvement in safety critical events (crashes/near-crashes) can be a leading and proactive indicator of safety (Wali et al., 2018c). For example, referring to the second block in Fig. 2 (indicated by "B"), the three persons now leave the grocery store(s), and start moving to restaurant(s). On their way, one of the persons (indicated by the red trajectory) gets into a crash, while the second person (indicated by the orange trajectory) gets into a near-crash event (Fig. 2). The third person (as is shown by green trajectory) does not get into any safety-critical event (crash or near-crash) on their way from grocery store to restaurant. Assuming that we have observed the two safety-critical events as well as the baseline event, we can now analyze the driving trajectories for the three trips to understand how volatility in microscopic driving decisions relate to the safety-critical (in this case a crash and near-crash) and baseline events, and whether such information can be used to predict occurrence of a crash and/or near-crash event, or subsequent injury outcomes given an unsafe event (Fig. 2) (Wali et al., 2018c).

Continuing analysis of high resolution connected and/or naturalistic driving data, another new idea relates to extending driving volatility to specific network locations, termed as *"location-based volatility"* (see the third block in Fig. 2) (Kamrani et al., 2017; Wali et al., 2018b). For illustration, continuing in the space-time dimension, the three persons now decide to individually leave the restaurant and go to some other place (see the third block in Fig. 2 indicated by "C"). As they move from the restaurant, they happen to pass through an intersection (see Fig. 2). In this case, assuming that we have microscopic driving trajectories, we can quantify volatility in instantaneous driving decisions for each of the three trips (or vehicle passings) and average the volatility indices for the three trips to generate "location-specific" volatility indices (Kamrani et al., 2017; Wali et al., 2018b). Thus, the idea of "location-based volatility" introduces the notion that high volatility and variability in microscopic driving decisions at a specific location can be related to the safety performance of that location, such as historical crashes (Kamrani et al., 2017; Wali et al., 2018b). Finally, the last element of volatility matrix is "driver-based volatility" (indicated by "D" in Fig. 2). As the name implies, driver-based volatility is person/driver specific and incorporates the volatility in driving decisions associated with each individual person. In this regard, the event-based volatility can also be deemed of as driver-based because we have person-specific individual vehicle passing trajectories before involvement into a baseline or safety-critical event. However, another equally important element of "driver-based volatility" can be the utilization of information on driver's biometrics and health data. For instance, how the heart rate, head movement, blood pressure, and pulse rate of a driver fluctuates as s(he) undertakes a specific trip (in a trip-based volatility domain), passes through a particular location (location-based volatility domain) or gets into a safety-critical event (event-based volatility domain).

As is evident, the concept of driving volatility matrix helps us understand the extent of variations in microscopic driving decisions at several hierarchies of the traffic ecosystem (Wali, 2018). The concept of driving volatility matrix provides a systems framework for characterizing the health of three fundamental elements of a transportation system: health of driver, environment, and the vehicle (Wali, 2018). By altering volatility in real-world microscopic driving decisions, vehicle kinematics, and roadway environment, the outcomes help improve transportation safety by proactively predicting crash occurrence and its severity given a crash. This paper and the rest of analysis presented relates to the new concept of "event-based volatility".

3.3. Data

Data from an on-going national Naturalistic Driving Study conducted as part of the 2nd Strategic Highway Research Program (SHRP) were used in this study (TRB, 2013). In this largest naturalistic driving study performed to date, the driving behaviors of approximately 3400 participant drivers were recorded with over 4300 years of naturalistic driving data collected between 2010 and 2013 (Hankey et al., 2016). The study data was collected from six naturalistic driving sites around the United States, with largest data collection sites in New York, Tampa, Seattle, Washington, Florida, and Buffalo (Hankey et al., 2016). The study used approximately 3300 participant vehicles (TRB, 2013;

Table 1
Different Volatility Measures Considered In this Study.

Direction	Performance Measure	Scheme 1 (K = 1): Entire 30-seconds data[1]	Scheme 2 (K = 2): First 20-seconds data[2]	Scheme 3 (K = 3): First 25-seconds data[3]
Longitudinal Direction	Positive vehicular jerk, $Longitudinal\ Vol_{pos-jerk,K}$	✓	✓	✓
	Negative vehicular jerk, $Longitudinal\ Vol_{negative-jerk,K}$	✓	✓	✓
	Acceleration, $Longitudinal\ Vol_{acceleration,K}$	✓	✓	✓
	Deceleration, $Longitudinal\ Vol_{deceleration,K}$	✓	✓	✓
Lateral Direction	Positive vehicular jerk, $Lateral\ Vol_{pos-jerk,K}$	✓	✓	✓
	Negative vehicular jerk, $Lateral\ Vol_{neg-jerk,K}$	✓	✓	✓
	Acceleration, $Lateral\ Vol_{acceleration,K}$	✓	✓	✓
	Deceleration, $Lateral\ Vol_{deceleration,K}$	✓	✓	✓

Notes: (1) Entire time series data, i.e., 20-seconds for baseline and 30-seconds for crash/near-crash events; (2) Of the 30-seconds data, the initial 20 s data are used while the 10 s data immediately prior to crash/near-crash are not used; (3) The initial 25 s data are used while the 5 s data immediately prior to crash/near-crash are not used; The three different data schemes, entire 30-seconds data, first 20-seconds data, first 25-seconds data, used in calculation of volatility indices are indexed by K = 1, K = 2, and K = 3 respectively (see text for explanation).

Hankey et al., 2016), using a data acquisition system (DAS) that collected four video views (driver's face, driver's hand, forward roadway, and rear roadway), vehicle network and status information (speed, brake, acceleration), and information from additional sensors networked with the DAS (e.g., accelerometers) (TRB, 2013).

Out of the many data categories collected in the SHRP 2 project, the data used in this study are "event data" and "continuous data". Event data provides detailed information regarding the different safety events in which a participant driver was involved. A notable feature of the SHRP 2 NDS data is the inclusion of three categories designated as events: crash, near-crash, and baseline events. Information about crash and near-crash events can provide richer estimates of prevalence and risk from different driver behaviors, roadway characteristics, and environmental conditions, whereas, baseline events are necessary for comparison purposes (TRB, 2013). For definitions of the three event severities, see (Hankey et al., 2016).

A case-cohort type sampling design is used for selection of baseline events where a random sampling scheme was conducted stratified by participant and proportion of time vehicle was driven. Proportion of time driven is defined as to include only vehicle speeds above 5 mph so as to eliminate the effect of long stopping times, and to focus on time periods where the vehicle was actually at risk of a crash or near-crash (Hankey et al., 2016). Regardless of whether involved in a crash or near-crash, all participants are included in the sample for baseline events to ensure that a minimum of one baseline event is included for each driver. Further details regarding the sampling design can be found in Hankey et al. (2016) (Hankey et al., 2016). The combination of crashes and near-crashes are referred to as safety-critical events (SCEs). Regarding identification of SCEs, SHRP 2 NDS used multiple methods such as, 1) Data collection site report, 2) Automatic Crash Notification (ACN), 3) Critical Incident (CI) button, 4) Analyst identified, and 5) Trigger execution. For example, the most systematic approaching to identifying an SCE was the method of trigger execution, which included post hoc processing of incoming data via custom algorithms called "triggers." These algorithms are characterized by different kinematic and behavioral signatures that have a highly probability of being present during specific SCEs. The SHRP 2 NDS used different thresholds based on project resources (as detailed in (Hankey et al., 2016)). Among many of other trigger types, longitudinal deceleration and lateral acceleration were also used. For instance, the initial trigger for lateral acceleration was specified if the lateral acceleration was greater than or equal to 0.75 g or less than or equal to -0.75 g, and that this threshold was exceeded for at least 0.2 s. For details about other trigger types, see (Jovanis et al., 2011; Hankey et al., 2016). Finally, once a SCE was identified through trigger execution method, video verification was then used to determine if an SCE occurred. Details about other methods used for identifying SCEs can be found in (Hankey et al., 2016). As can be seen, all the aforementioned techniques used in SHRP 2 NDS for identification of SCEs are rigorous. On top of that, the fact that majority of the SCEs identified through any of the aforementioned methods were video verified at the end further increases our confidence in the data.

A total of 9593 driving events are considered in this study, out of which 7589 are baseline, 673 are crashes, and 1295 are near-crash events. Note that the term "event" does not imply "trips". A participant along a single trip can have several events, e.g., baseline, crash, and/or near-crash. The 9593 events involve 1580 unique participants with some participants appearing more than once (i.e., involvement in more than one safety critical events). For each of the three event severities (baseline, crash, near crash), time-series data on vehicle motion (continuous data) is provided, i.e., 30 s instantaneous driving data (frequency of 10 Hz) for safety critical events (crashes and near-crashes) and 20 s instantaneous driving data with a frequency of 10 Hz for baseline events (Hankey et al., 2016). The time-series data contains information about longitudinal and lateral accelerations, speeds, gas pedal and steering wheel position, and wiper status.

As such, a total of 2.2 million records of real-world driving are analyzed in this study. By using information related to longitudinal and lateral accelerations and vehicular jerk, 24 different measures of driving volatility are calculated using the methods described next (Table 1). Finally, the event table provides detailed information on pre-incident maneuvers[1], legality of maneuvers, driver behavior, secondary tasks[2], start and end times, if applicable, of first, second, and third secondary events. Also included in the data is information about front-seat and rear-seat passengers, intersection and roadway type indicators, and traffic flow related factors. The detailed event data are finally linked with the event-specific volatility indices for subsequent analyses.

3.4. Components of volatility

Fig. 3 shows a 30-seconds longitudinal and lateral acceleration/deceleration (vehicular jerk) profiles prior to involvement in a crash event. By using large-scale data analytic techniques, driving volatility

[1] For Baseline events, this is the driving maneuver or action that the driver is engaged in for the last 2-6 seconds prior to the baseline anchor point (the point in video where the 20 seconds baseline driving data starts).

[2] Observable driver engagement in any of secondary tasks, and which begins at any point during the 5 seconds prior to the event start (crash, near-crash) through the end of the event (TRB, 2013). Secondary tasks primarily refer to distractions related to non-driving related glances away from the direction of vehicle movement (TRB, 2013). Some examples include radio adjustments, seat-belt adjustments etc. For Baselines, secondary tasks are coded for the last 5-6 seconds of the baseline epoch, which includes 5 seconds prior to baseline event start through one second after (to the end of the baseline). For further details, refer to (TRB, 2013).

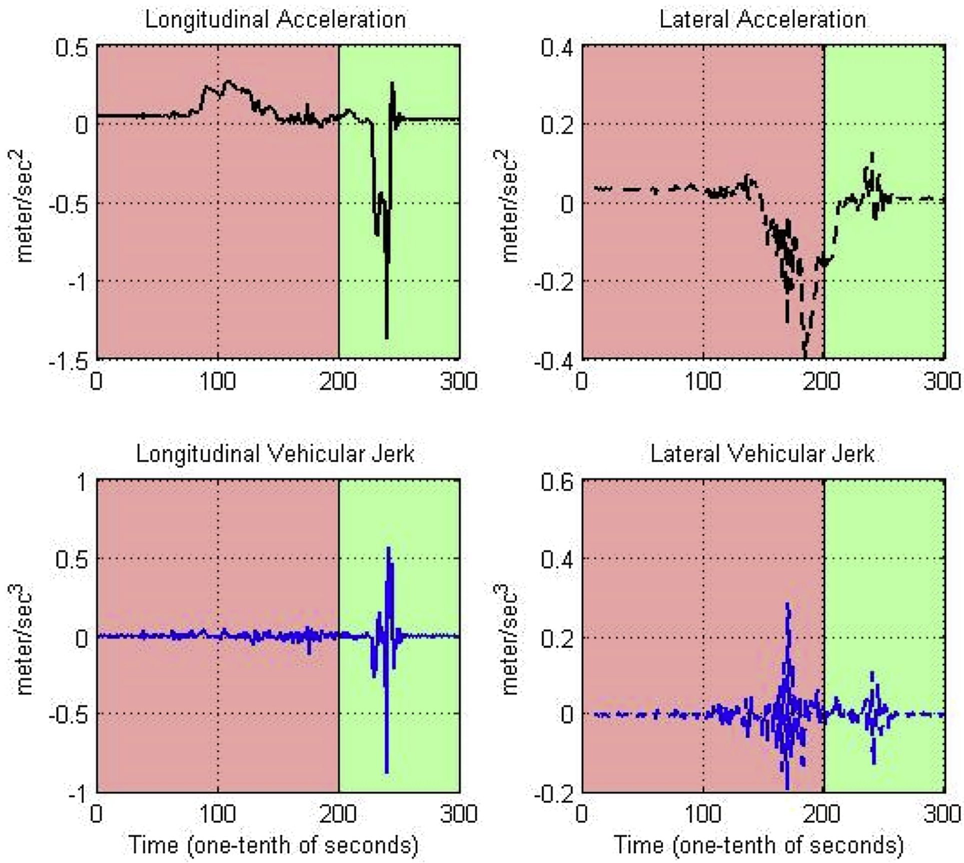

Fig. 3. Profiles of Instantaneous Driving Decisions Prior to Involvement in a Sample Crash Event (First portion of series in pink background indicate actual driving behavior and second portion of series in light green background may indicate driving decisions due to situational factors – see text for explanation). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

can be characterized for each of the events (i.e., baseline, crash, or near-crash events). Broadly speaking, the volatility indices for each event can be regarded as microscopic measures of driving performance (or erractic behavior) in normal or safety-critical events. The driving volatility indices developed using the entire 30-seconds data (for crash and near-crash events) can shed light on microscopic driving decisions that a driver undertook prior to involvement in safety-critical events. With real-world driving data based volatility indices (Fig. 3), proactive behavioral countermeasures can be planned for drivers that are consistently more volatile.

However, as explained in Wali et al. (2018), using the entire 30-seconds driving data for crash and near-crash events aggregates the different components of volatility in instantaneous driving decisions prior to unsafe outcomes (Wali et al., 2018c). For instance, Fig. 3 illustrates the microscopic driving decisions in longitudinal and lateral direction prior to a crash event. However, in case of crash and near-crash events, the series can be divided into two series, where the first portion of the series will indicate the driver's speed choice (or acceleration, vehicular jerk) regardless of the event outcome, while the second portion of the series would indicate the "adjustments" or drivers' reaction to the event. That is, the volatility in the first component of the series is likely to reflect the actual driver behavior and can be regarded as "intentional volatility" by the driver, i.e., due to aggressive self-driving when the driver is in control. As shown in Fig. 3, the intentional volatility may be reflected in instantaneous driving data 20–30 seconds before the crash/near-crash (see the first part of the profile in Fig. 3 – indicated in pink background). Whereas, the second component may reflect "unintentional volatility" in the sense that the driver may have taken evasive maneuvers to avoid the crash or near-crash event or lost the control. The unintentional volatility can be reflected in driving data immediately before the unsafe outcome (say 5–10 seconds before the crash/near-crash event), as highlighted in light green background in Fig. 3. We call it "unintentional volatility" as the driver may have already anticipated the crash/near-crash in this case and is undertaking preventive measures to avoid the crash, i.e., evasive maneuvers. Also, the volatility immediately prior to the crash is likely to contain volatility due to loss of control before the crash (Wali et al., 2018c). We sincerely thank the four reviewers who highlighted this issue as well.

The fundamental objective of this study is to explore the links between driving volatility and crash propensity. As is evident, combining the two sources of volatility (intentional vs. unintentional) to explain the unsafe outcomes can lead to bias due to reverse causation, i.e., the volatility measures using 30-seconds data will not only reflect the actual driving behavior but also volatile driving behaviors due to risky situations/external events. Conversely, by aggregating the different components of volatility, evasive maneuvers that allowed a driver to avoid a crash would also be interpreted as increasing volatility. In this case, high volatility may then be associated with near-miss outcomes, again not due to driver behavior in general but due to a driver's reaction to unobserved situational variables. In this case, the first portion of the time series, i.e., intentional volatility can be used to explain the occurrence of unsafe outcomes (Wali et al., 2018c). As illustrated in Fig. 3, much of this bias may be eliminated by censoring data used to calculate different volatility measures in the time period immediately before a crash or near-crash outcome occurs; i.e., censoring to remove the influence of driver reactions immediately prior to a crash from the volatility measures while retaining volatility derived from driver behavior in the seconds leading up to, but not immediately before, a crash or near-rash event. As such, we also consider generating segmented volatility indices based on different time bins, and which can separate out how volatility in time to crash (or near-crash) relates to crash propensity (Fig. 3).

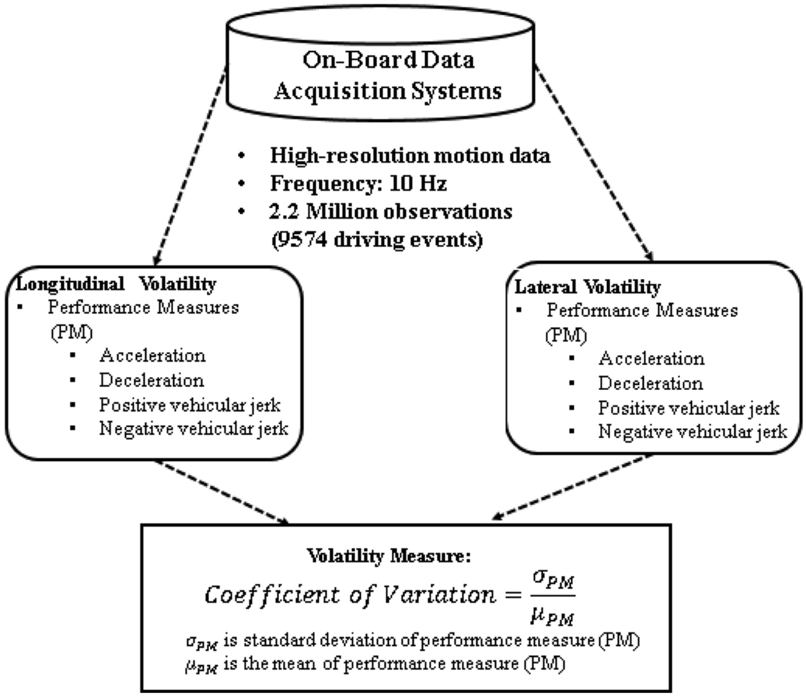

Fig. 4. Methodology for Characterizing Driving Volatility Prior to Involvement in Safety Critical Events.

3.5. Calculation of volatility

Driving volatility captures the extent of variations in driving, especially hard accelerations/braking and jerky maneuvers (Khattak et al., 2015; Liu et al., 2015a; Wang et al., 2015; Liu and Khattak, 2016; Kamrani et al., 2017; Wali, 2018; Wali et al., 2018b, c). Different instantaneous driving performance measures such as vehicle speeds, accelerations/decelerations, and/or steering angles can be used for estimation of volatility indices in longitudinal directions (Quddus, 2013; Kim et al., 2016; Liu and Khattak, 2016). For details, see (Wali et al., 2018b). Another important driving decision is in the lateral dimension, e.g., lane change decisions. Larger volatility in lateral dimension may also be associated with unsafe outcomes (Wali et al., 2018c). Another measure recently introduced in the literature for characterizing driving volatility is vehicular jerk (Wang et al., 2015; Wali et al., 2018c). Vehicle jerk is basically defined as the rate of change of vehicle acceleration with respect to time. Compared to accelerations/decelerations, vehicular jerk represents drivers' decisions to change marginal rate of acceleration or deceleration, and may better characterize driving volatility in instantaneous driving decisions (Wali et al., 2018b, c). In relevance to the current study, Feng et al. (2017) concluded the better potential of longitudinal vehicular jerk to identify "aggressive" drivers (Feng et al., 2017). However, vehicular jerk measures in lateral direction are not widely used (Wali et al., 2018c).

To fully characterize volatility in instantaneous driving decisions, we use both acceleration and vehicular jerk based performance measures (Fig. 3). As deceleration profiles usually have higher variations (Kamrani et al., 2017), separate volatility measures for acceleration and deceleration are used. Likewise, separate volatility indices are generated for positive and negative jerk values. While doing so, both longitudinal and lateral dimensions are considered in calculation of volatility prior to involvement in safety critical events, and which can better characterize the complex mechanism of instantaneous driving decisions in longitudinal and lateral directions (Fig. 3). Fig. 4 illustrates the methodology for characterizing driving volatility. For the sake of completeness, the formulae for velocity, acceleration, and vehicular jerk are shown in Eqs. (1)–(4):

$$r = Position \quad (1)$$

$$Velocity = v = \frac{\partial r}{\partial t} \quad (2)$$

$$Acceleration = a = \frac{\partial v}{\partial t} = \frac{\partial^2 r}{\partial^2 t} \quad (3)$$

$$Vehicular\ Jerk = j = \frac{\partial a}{\partial t} = \frac{\partial^2 v}{\partial^2 t} = \frac{\partial^3 r}{\partial^3 t} \quad (4)$$

Where: $\frac{\partial}{\partial t}$ indicates derivative of a performance measure (velocity, acceleration, etc.) with respect to time, and ∂t is a small change in time "t" (set to 0.1 s in our case).

Specifically, the on-board data acquisition systems installed in vehicles provide high-resolution motion data at a frequency of 10 Hz (TRB, 2013). Instantaneous longitudinal and lateral acceleration profiles are recorded for the entire trip. However, compared to drivers' performance throughout the entire trip, instantaneous driving decisions in the seconds leading up to, but not immediately before, a crash or near-rash event are more relevant and crucial. As such, the EDT provides instantaneous motion data, 30 s for every safety critical event (crash/near-crash) and 20 s for every baseline event (TRB, 2013). The 30-second driving behavior data can be interpreted as driving decisions undertaken immediately before the occurrence of crash or near-crash event.

A total of 2.2 million observations are used in this study for calculation of volatility indices for 9574 driving events (discussed later in detail). As shown in Fig. 4, for each event, acceleration and deceleration values are separated, and mean and standard deviations calculated for each. As a measure of volatility, coefficient of variation is used in the present study[3], i.e., the standard deviation(s) are then divided by mean

[3] Kamrani et al. (2017) (Kamrani et al., 2017) introduced coefficient of variation as a measure for characterizing driving volatility in connected vehicles environment. The coefficient of variation is a measure of variability used extensively in time-series finance literature (Weber et al., 2004) and recently in transportation literature (Kim et al., 2016; ; Wali et al., 2018b). Compared to standard deviation or variance, coefficient of variation is scale in-sensitive and this property allows meaningful comparisons between the volatility in instantaneous driving decisions in different safety critical and baseline events.

values to get an estimate of relative variability in instantaneous driving decisions across different events. Similar procedure is repeated for acceleration/decelerations in lateral direction, and for vehicular jerk (both positive and negative) in longitudinal and lateral directions (Fig. 4). For example, for each event, we first calculated vehicular jerks using the instantaneous vehicular acceleration data each one-tenth of a second (see Eq. (4) above). Then, we separated out positive and negative vehicular jerks for each event. Finally, for calculating volatilities associated with positive vehicular jerk, we calculated the standard deviation of positive vehicular jerk and divided it by its mean value. For negative vehicular jerk, we divided the corresponding standard deviation by the absolute mean value of negative vehicular jerk for each event. The same procedure was repeated for vehicular jerk in lateral direction. Finally, considering the discussion on different components of volatility (see Section 3.4), 24 different volatility measures are considered in this study based on the whether the entire 30 s, starting 20 s, or 25 s data are used[4] (Table 1). For ease of interpretation, all the 24 volatility measures are indexed (Table 1). For example, referring to Table 1, $Longitudinal\ Vol_{pos-jerk,K}$ indicates longitudinal volatility in positive vehicular jerk calculated using the entire 30-seconds data, the starting 20 s, or the starting 25 s driving data (indicated by subscript *"K"*). That is, the three different data schemes, entire 30-seconds data, first 20-seconds data, first 25-seconds data, used in calculation of volatility indices are indexed by K = 1, K = 2, and K = 3 respectively (see top row in Table 1). For instance, volatility in positive vehicular jerk in longitudinal direction calculated using the entire 30-seconds data is indexed as $Longitudinal\ Vol_{pos-jerk,K=1}$. Whereas, volatilities in positive vehicular jerk in longitudinal direction calculated using the first 20-seconds and first 25-seconds data are indexed as $Longitudinal\ Vol_{pos-jerk,K=2}$ and $Longitudinal\ Vol_{pos-jerk,K=3}$ respectively.

Note that the SHRP 2 NDS provides 30-seconds microscopic vehicle trajectory data for each of the safety-critical events (crashes/near-crashes), and which were used in this study. In addition, 29–30 seconds videos are also available for majority of the safety-critical events. However, we discovered that a safety-critical event does not need to occur exactly at the end of the 30-seconds trajectory data or at the end of the corresponding video files. In other words, the safety-critical event can take place before the end of the event data/video file, e.g., after 20 s while the event data is provided for the entire 30-seconds. Using the entire 30-seconds data in this case will distort the results, especially with respect to intentional vs. unintentional volatility. While it is not practical to manually check all of the video files for safety-critical events (as there are thousands of events), we manually checked the video files of a completely random sample of 100 crashes to exactly record the time at which the event occurred during the 30 s video files. For the decision of using all the 30 s data to be reasonable, we would expect the distribution of the event occurrence times to be left-skewed, i.e., majority of the event occurrence times would be expected to occur at the end of the trajectory/video files. Out of the 100 randomly sampled crashes, videos were available for 59 of them. For these 59 crashes, the video duration was 30 s for all except one crash (13 s). After extracting the event (crash) occurrence time from these 59 videos, we conducted a descriptive analysis to see the distribution of the data. The resulting distribution of event occurrence time (in seconds) was highly skewed to the left with a skewness parameter of -1.92 (skewness of less than -1 indicates heavily skewed distribution to the left) and kurtosis parameter of 10.887 (kurtosis value for a normal distribution is exactly 3). The mean event occurrence time was 24.70 s (median of 25 s) with a small standard deviation of 2.39 s. This analysis highlights that the decision to use the entire trajectory data for calculation of volatilities is reasonable, given that all the videos cannot be manually checked. However, out of the 59 crash events, we found one shorter duration video of 13 s for a crash event. That is, 1% of the crashes in the sample may be like these if we have the event videos for all of them.

From a perspective of calculating intentional volatility, the descriptive analysis of event occurrence times extracted from videos revealed a 25th percentile value of 24 s, i.e., for around 14 events (25% of 59 crash events) the crash-event occurred at 24 s or earlier in the event video files. Likewise, the 5th percentile value of event occurrence time was 21 s, which is also the smallest event occurrence time for the randomly sampled crashes (N = 2 events occurring at 21 s in event video files). Thus, given this data observability issue, and the impracticality of manually checking thousands of videos, one would conceptually prefer using the first 20-seconds of driving data in calculation of "intentional volatility" compared to using the first 25-seconds of the driving data, as using the latter scheme may not fully separate out intentional and unintentional volatility for the events with event occurrence time of less than 25 s in the video streams. For more than 90% of the sampled events, the event occurrence time was greater than 22 s. Thus, for the sake of conservative inferences, using the first 20-seconds driving data can better separate out intentional vs. unintentional volatility even for events that occurred earlier in the video files (such as at 22 s), keeping in view the average perception-reaction time of 2.5 s. However, this may also imply that we may not have some creep from unintentional volatility due to the crash, especially for crash events with event occurrence times around 21 s. The extent of this, nonetheless, would be low given that only 2% of the cases had an event occurrence time of 21 s in the video files if we have videos for all the events. Contrarily, using the first 20-seconds of driving data will also imply discarding a portion of "intentional" volatility for events that occurred at 25 s or after, in the event video files (the event occurrence time for 50% of the randomly sampled events was 25 s or greater). But, as mentioned earlier, for the sake of disentangling unintentional and intentional volatility to the extent possible, and to produce conservative inferences given that 50% of the events occurred before 25 s in the video files, it would be conceptually better to use the first 20-seconds driving data. We empirically examine, nonetheless, both of the data censoring schemes in the analysis presented next.

From a comparability perspective, the lengths of the durations for calculation of volatility indices for baselines and safety critical events are different. That is, the length of each baseline event is 20 s, whereas the length of safety critical events is 30 s. However, as discussed in detail above, a safety-critical event does not need to occur exactly at the end of the 30-seconds trajectory data or at the end of the corresponding video files. Ideally, equal lengths of the durations of driving data in the NDS for baselines and safety-critical events would be preferable. However, the volatility indices are still comparable because the duration of baseline events should not significantly alter the time-series distribution of microscopic driving decisions, assuming that the 20-seconds baseline driving profile on-average is representative of the entire (baseline) driving profile for a specific driver. Finally, given the data observability issue discussed above, the final models are based on 20-seconds driving data for safety critical events whereas 20 s driving data are used for baselines as well, thus making them pretty much comparable.

[4] Instead of a fixed censoring scheme used in this study (i.e., using the first 20 or 25 second of data in calculation of intentional volatility for all safety-critical events), a better choice could be to use the time a driver starts reacting as the time to differentiate "intentional volatility" and "unintentional volatility". In a separate and more recent research effort (results available from the authors) (Wali and Khattak, 2019), the authors used the driver's reaction time as the time to differentiate between intentional and unintentional volatility for all the events collected under NDS till date (N = ˜40,000 events). Using the time a driver starts reacting as the criterion, on average around 21.51 and 22.61 seconds of driving data were used for calculating intentional volatilities for near-crash and crash-events respectively (Wali and Khattak, 2019). Thus, our choice of using 20-seconds of data seems reasonable. In fact, keeping in view the above results, our estimates of "intentional volatility" are likely conservative.

3.6. Statistical models

Once the event specific volatilities (24 different volatility measures) are calculated for each event, the correlations between crash propensity and driving volatility are explored. Appropriate statistical models can shed light on microscopic driving decisions i.e., driving volatility, and how such decisions may be related to involvement in safety critical events. The potential outcomes related to crash propensity are baseline events, crash events, crash-relevant events, near-crash events, and non-subject conflict events. As can be seen, the response outcome is discrete and exhibits some ordering. However, un-ordinal probability models can be applied for the analysis of ordinal categorical outcomes (as is widely done in the literature), as such models are not afflicted with some of the key restrictions imposed by traditional ordinal probit and logit models (Savolainen and Mannering, 2007; Savolainen et al., 2011; Iranitalab and Khattak, 2017) and the references therein. Thus, un-ordinal multinomial logit discrete framework is used in this study. Following McFadden et al. (McFadden, 1973), a crash propensity function determining the outcome "*i*" of a specific event "*j*" can be defined as:

$$CP_{ij} = \boldsymbol{\beta_i X_{ij}} + \varepsilon_{ij} \tag{5}$$

Where: CP_{ij} is a crash propensity function determining the safety outcome "*i*" (if event was baseline, crash, crash-relevant, near-crash, or non-subject conflict) for event "*j*"; $\boldsymbol{X_{ij}}$ is the matrix of explanatory variables (driving volatility related variables, pre-crash maneuvers, weather, or traffic factors); $\boldsymbol{\beta_i}$ is the vector of estimable parameters related to each explanatory factor in $\boldsymbol{X_{ij}}$, and ε_{ij} are the error terms. Following McFadden (1973) (McFadden, 1973) and Train (Train, 2003), if ε_{ij} are assumed to be generalized extreme value distribution, the multinomial logit model can be formulated as (Wu and Jovanis, 2012):

$$P_j(i) = \frac{\exp[\boldsymbol{\beta_i X_{ij}}]}{\sum_I \exp[\boldsymbol{\beta_i X_{Ij}}]} \tag{6}$$

Where: $P_j(i)$ is the probability of specific outcome "*i*" (from the superset of all possible categories *"I"* defined earlier) for event "*j*". Following (Train, 2003), the following log-likelihood function can be solved to get estimates of $\boldsymbol{\beta_i}$:

$$LL = \sum_{j=1}^{J}\left(\sum_{i=1}^{I} \aleph_{ij}\left[\boldsymbol{\beta_i X_{ij}} - \boldsymbol{LN}\sum_{\forall \boldsymbol{I}} \boldsymbol{exp}(\boldsymbol{\beta_i X_{Ij}})\right]\right) \tag{7}$$

Where: $\aleph_{ij}$ is an indicator equal to 1 if observed response outcome for event "*j*" is "*i*", and 0 otherwise (Train, 2003). For further details about the fundamentals of the statistical methods used, see (Wu and Jovanis, 2012).

3.6.1. Unobserved heterogeneity

The key focus of this study is to investigate correlations between driving volatility related measures and crash propensity. Crash propensity can be influenced by different factors, some of which are observed while other factors are unobserved in the data at hand. Given such unobserved factors, the correlation between explanatory factors (such as driving volatility indices) and crash propensity may vary across different events, and which is referred to unobserved heterogeneity in the literature (Train, 2003). In addition, the issue of possible omission of relevant and important explanatory factor(s) from the modeling framework has also serious implications (Washington et al., 2010; Jovanis et al., 2011). For example, if important explanatory factor (e.g., education, age, gender etc.) that may influence driver's performance is omitted from the model, it may happen that the observed correlation between driving volatility indices (observed explanatory factor) and crash propensity may be an outgrowth of those omitted factors, and not true correlation between volatility and crash propensity (Washington et al., 2010). Note that, the unobserved factors need not to be only driver-related but can also include other omitted and important variables, such as situational variables (Jovanis et al., 2011). For instance, while the SHRP 2 NDS database contains very detailed context-specific data, information on certain situational variables such as more complex situations, shorter sight distances, erratic driving be other motorists may not be available. These omitted situational factors may also be correlated with crash propensity. In our case, all such unobserved factors (driver-specific, vehicle-specific, environment-specific, and situation-specific, to name a few) that are likely to be correlated with crash propensity become a portion of the unobserved heterogeneity. As such, the statistical methods used in this study account for all different types of unobserved variables. To reach reliable conclusions about the correlation between driving volatility and crash propensity, it is crucial to account for the afore-mentioned methodological concerns. To account for these issues (Savolainen et al., 2011; Kamrani et al., 2017), a random parameter framework is adopted where the $\boldsymbol{\beta_i}$ are allowed to vary across different events according to some pre-specified distribution[5] . Following (Train, 2003), a mixing distribution is introduced in random parameters logit model, where the logit framework now becomes:

$$P_j(i) = \int \frac{\exp[\boldsymbol{\beta_i X_{ij}}]}{\sum_I \exp[\boldsymbol{\beta_i X_{Ij}}]} f(\beta\varphi)d\beta \tag{8}$$

Where: $f(\beta\varphi)$ is the density function of β conditional on the vector of parameters for the density function denoted φ. With the random parameter logit model in Eq. (8), β can now account for driver-specific variations of the effect of $\boldsymbol{X_i}$ on probabilities of different crash propensity outcomes, and with β determined by approximating the integral in Eq. (8) by drawing from a pre-specified density function $f(\beta\varphi)$ (Train, 2003). The estimation proceeds with Maximum Simulated Likelihood procedures where Halton draws (compared to random draws) are used in the simulation process. In this study, 1000 Halton draws are used for parameter estimation, nonetheless, 200 Halton draws are reported to produce accurate parameter estimates (Bhat,

[5] Over here, we clarify between the traditional multilevel (fixed and random effects models) and random parameter models. Note that the terminology of grouping variable used earlier for interpreting the results of traditional ANOVA is irrelevant to random parameter logit or mixed logit models (used in this study) and is only relevant to traditional multilevel (fixed- and random-effects) models (which are not used in this study). The fixed- and random-effect methods are typical techniques to incorporate unobserved heterogeneity in a "panel data" setting. In other words, using multilevel models (fixed- and random-effects) essentially implies that the analyst is working with panel data (Greene and Hensher, 2010; Wooldridge, 2010). That is, a "panel or group" data setting is usually employed for the application of fixed- and random-effects (multilevel) estimators (Wooldridge, 2010; ; Wali et al., 2018a). With traditional multilevel estimators, the individual-specific unobserved heterogeneity is captured in the constant term/intercept in a panel data setting (Greene and Hensher, 2010; Wooldridge, 2010). Subsequently, the rule of thumb is to have large enough sample size at level two grouping variable in a multilevel modeling framework for obtaining unbiased results (Wooldridge, 2010). However, this rule of thumb only applies to "panel data" multilevel models and does not apply to the cross-sectional application of simulation-based random parameter logit models estimated in this study. As opposed to multilevel (fixed- and random-effect) methods which necessitate panel data or grouping variable approach, simulation-based random parameter logit models can be applied both in cross-sectional (as is the case in present study) and panel setups. By using simulation-based maximum likelihood estimation (such as Monte Carlo integration or Markov Chain Monte Carlo methods), the random parameter modeling technique allows the possibility of coefficients on explanatory variables to vary across individual observations (which in our case is event severity – baselines, near-crashes, crashes) or groups (in case of panel data) (see Chapter 11 in (Greene, 2011) for details). As a result, the random parameter logit models represent an elegant extension in which the researcher can expand the amount of unobserved heterogeneity captured across individual observations (baseline, crash, or near-crash events in our case) (Train, 2003). For further details, see (Greene and Hensher, 2010;) (; ; Wali et al., 2018a).

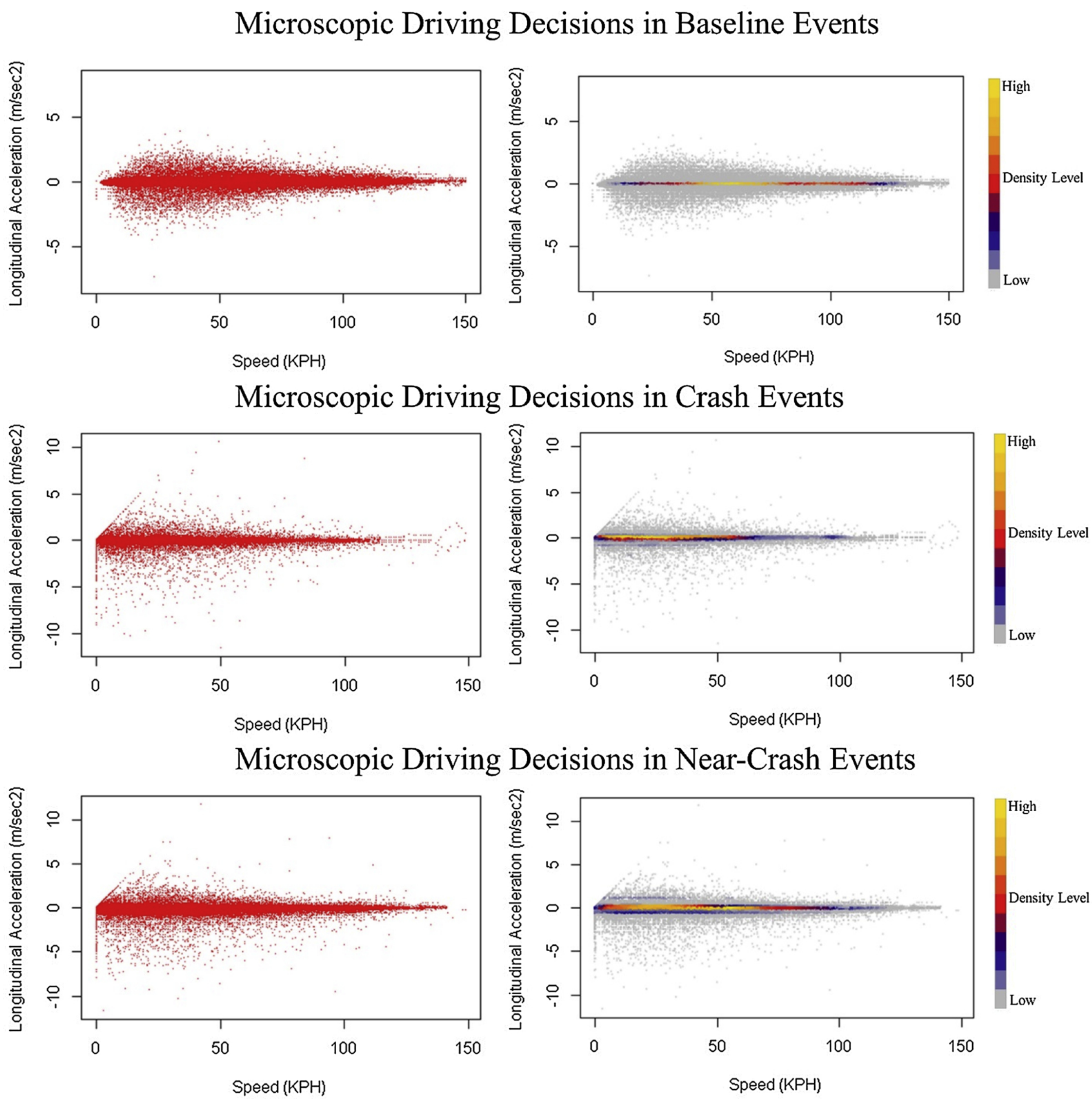

Fig. 5. Scatter and Density Plot Distributions of Longitudinal Acceleration and Speed in Baseline, Crash, and Near-Crash Events.

2003; Train, 2003). Regarding function form of the parameter density functions, we have tested normal, lognormal, triangular, uniform, and Weibull distributions. Further details can be found in (Train, 2003; Anastasopoulos and Mannering, 2011; Kamrani et al., 2017).

Although crash propensity function in all the estimated random parameter logit models are expressed in linear form, the logit transformation restricts direct interpretation of parameter estimates (Naik et al., 2016; Wali et al., 2018d). To intuitively interpret the modeling results, marginal effects are estimated for the fixed- and random-parameter logit model (Naik et al., 2016). For a certain change in value of explanatory factor, marginal effect suggests an instantaneous change in the probability of a crash propensity outcome while keeping all other factors at constant. Separate marginal effects are estimated for continuous and binary indicator variables. Following (Train, 2003), as marginal effects can be different at different levels of explanatory factors, therefore the average marginal effects over the sampled events are estimated.

4. Descriptive Analysis

4.1. Concept illustration and descriptive statistics

To understand how microscopic driving decisions vary across different events, Fig. 5 illustrates the distributions of longitudinal acceleration against speed for baselines/normal driving events, crash events, and near-crash events. As can be seen, high speeds (> 50-60 kph) are associated with smaller magnitudes of acceleration/deceleration as well as smaller dispersion in acceleration/deceleration values, i.e., lower volatility. Likewise, for each of baseline, crash, and near-crash events, Fig. 5 provides the density scatter plot of the relationship between longitudinal acceleration and speed. It can be seen that the bandwidth of acceleration values at high speeds is tighter than the bandwidth of acceleration values at low speeds. This finding corresponds with our understandings of basic Physics principles according to which the ability to accelerate a vehicle naturally decreases at higher traveling speeds (Fig. 5) (Liu and Khattak, 2016; Khattak and Wali, 2017).

Importantly, Fig. 5 also reveals a smaller dispersion in microscopic driver decisions in baselines events, compared to significantly greater dispersion in case of crash and/or near-crash events (Fig. 4). This finding suggests that drivers, on-average, are more volatile in crash and near-crash events.

Next, Fig. 6 presents the distributions of eight aggregate volatility measures considered in this study (i.e., eight volatility measures indexed with subscript $K = 1$, see Table 1). For brevity, the distributions of volatility measure calculated using the entire data (20 s for baseline

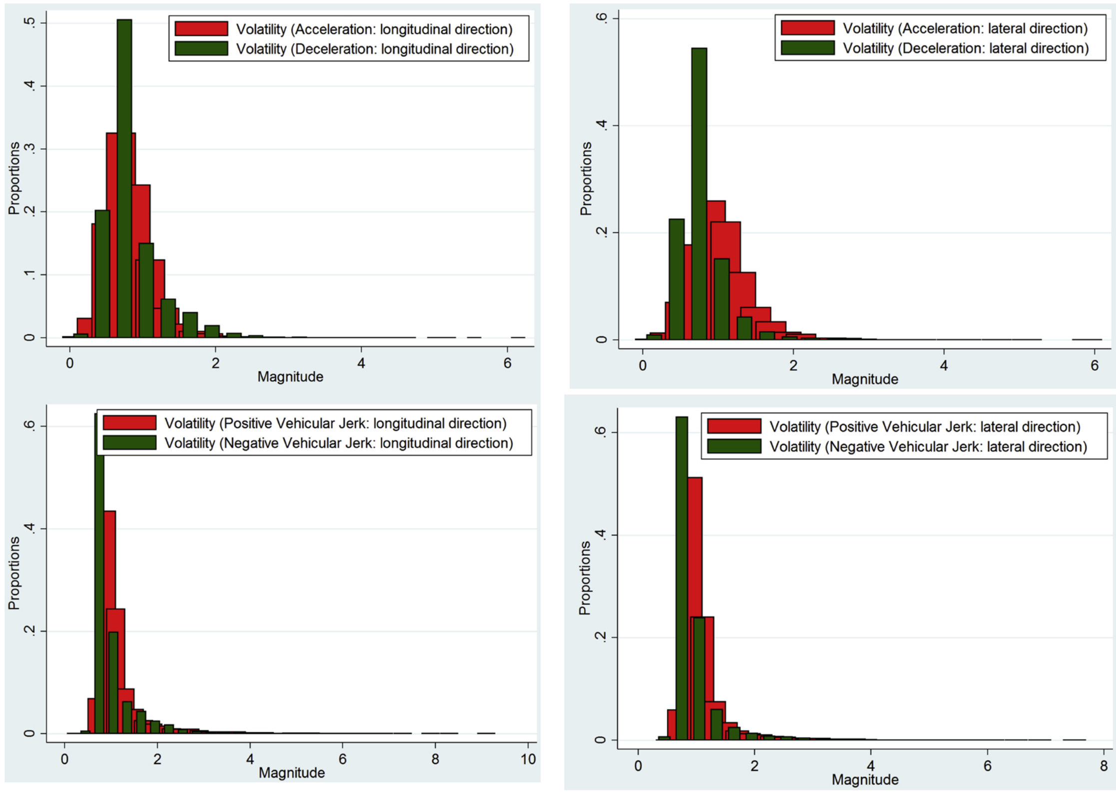

Fig. 6. Distributions of Volatility Measures in Naturalistic Driving Environment Calculated Using the Entire Data.

and 30 s for crashes/near-crashes) are shown (Fig. 5). All eight aggregate volatility measures are not normally distributed and in fact skewed to the right, with mean volatility statistic greater than the median. Recall that coefficient of variation (Fig. 4) is used as a measure to capture volatility in instantaneous driving decisions. Interestingly, the four volatility measures based on longitudinal and lateral accelerations (top two plots in Fig. 6) exhibit similar patterns, whereas, the volatility measures based on longitudinal and lateral vehicular jerks have greater magnitudes as well as range. Broadly, this suggests that vehicular jerk based measures may also better characterize the volatility in instantaneous driving decisions as it accounts for the sharp rate of change (within one-tenth of a second) in acceleration values (Liu and Khattak, 2016; Feng et al., 2017; Wali et al., 2018b, c).

Next, to see if there are statistically significant differences between aggregate driving volatility in safety critical events (crash, near-crash) and baseline events, ANOVA analysis is performed and results shown in Table 2. Our a-priori hypothesis is that compared to baseline (normal) driving events, drivers may exhibit greater volatility in crash and/or near-crash events.

Note that the below descriptive findings relate to driving volatility indices calculated using the entire time-series (20 s for baselines, and 30 s for crashes/near-crashes) and not the segmented time series data (see Section 3.4 for details). The rigorous statistical analyses presented in next section carefully addresses the issue of intentional vs. unintentional volatility and how it relates to crash propensity.

The top panel in Table 2 presents the summary statistics for all eight volatility measures (calculated using the entire data – refer to the third column in Table 1) and for each of the three event severities. Whereas, the bottom panel summarizes the within- and between-group variances across the three event severities. The within- and between-group variances are obtained from one-way ANOVA analysis (LP, 2005). In particular, the grouping variable is event severity, and which can be a baseline, near-crash, and crash. The motivation behind using event severity as a grouping unit is that we are interested in examining the magnitudes and variations in volatilities across different normal and/or safety-critical events. Of course, the one-way ANOVA analysis of driving volatilities can also be performed by considering other grouping units such as driver gender, age-groups, or education status, etc. However, we use the variable event severity given the primary focus of characterizing extent of variations in microscopic driving decisions within and across different events (baseline, crash, near-crash). The following observations can be made from Table 2:

- For all eight volatility measures (except one) tested, compared to baseline events, drivers on-average exhibit higher volatility in near-crash situations, whereas drivers' volatility in crash events is further greater than in near-crash situations.
- Interestingly, for near-crash events, the volatility in deceleration in longitudinal direction is greater than volatility in crash events (mean of 1.42 and 1.21 for near-crash and crash events respectively). This may be since drivers, in near-crash events, may react quickly (and with high volatility) to avoid an actual crash, and thus observed as near-crash.
- For volatility in longitudinal direction based on acceleration/deceleration measure, it is observed that drivers exhibit greater volatility in deceleration as compared to their volatility in acceleration (see mean values in top panel of Table 2 under longitudinal volatility). This is intuitive as drivers may react quickly to potential safety hazards in front of them and thus decelerate harder.
- For all volatility measures, compared to between-group variances, within-group variances are greater, suggesting larger variance in volatilities exhibited by different drivers involved in events of similar severity. This is intuitive given the driver-specific differences and that different drivers may respond differently even in same situations.
- Finally, for all the eight volatility measures, there is statistically significant evidence that driver volatilities in baseline, near-crash, and crash events are significantly different (see F-statistics and

Table 2
Descriptive and ANOVA Analysis of Driving Volatility Measures in Naturalistic Driving Environment.

	Longitudinal Volatility								Lateral Volatility							
*Performance Measures**	Positive vehicular jerk		Negative vehicular jerk		Acceleration		Deceleration		Positive vehicular jerk		Negative vehicular jerk		Acceleration		Deceleration	
Event Severity	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*	*Mean*	*SD*
Baseline	0.99	0.18	0.82	0.15	0.80	0.29	0.71	0.19	0.97	0.17	0.84	0.18	0.99	0.33	0.71	0.20
Crash	2.41	1.28	1.98	1.04	1.03	0.52	1.21	0.62	2.24	1.05	1.85	0.76	1.37	0.61	1.20	0.55
Near-Crash	1.77	0.77	1.50	0.47	0.87	0.31	1.42	0.45	1.30	0.46	1.06	0.33	1.14	0.46	0.92	0.36
	Analysis of Variance															
	B	**W**	**B**	**W**	**B**	**W**	**B**	**W**	**B**	**W**	**B**	**W**	**B**	**W**	**B**	**W**
SS	1754.4	2131.2	1238.2	1184.0	35.2	946.3	656.7	792.0	1048.2	1224.5	653.5	775.1	104.2	1328.7	175.9	667.3
DF	2	9556	2	9553	2	9426	2	9337	2	9,555	2	9,553	2	9409	2	9446
MS	877.2	0.2	619.1	0.1	17.6	0.1	328.4	0.1	524.1	0.1	326.8	0.1	52.1	0.1	88.0	0.1
F	3933.3	—	4995.3	—	175.5	—	3870.8	—	4089.6	—	4027.6	—	369.0	—	1245.2	—
Prob > F	0	—	0	—	0	—	0	—	0	—	0	—	0	—	0	—

Notes: (*) For definitions of volatility measures, refer to Fig. 2; SD is standard deviation; SS is Sum of Squares; DF is degrees of freedom; MS is mean of squares and is calculated as SS/DF; F is the corresponding F-statistic; B refers to between groups variance; W refers to within group variance; (—) means Not-Applicable; Unit for volatility in jerk is $\frac{m}{sec^3}$; Unit for volatility in acceleration and deceleration is: $\frac{m}{sec^2}$.

corresponding p-values in bottom panel of Table 2), with volatilities in near-crash and crash events significantly greater than volatility in baseline events.

Next, Table 3 provides the descriptive statistics of the significant variables in subsequent statistical models. The descriptive statistics of aggregate volatility related variables using the data (30-seconds for crash and near-crash and 20 s for baseline events) are presented in Table 3. Furthermore, owing to the presence of intentional and unintentional volatility (section 3.4), Table 3 to that effect also presents the descriptive statistics of volatility measures calculated using the first 20 s and 25 s driving data (as listed earlier in Table 1). Compared to volatility measures calculated using entire data, the volatility measures computed using the censored data are more likely to reflect the actual driving behavior and not the evasive maneuvers undertook by a driver due to situational factors.

Several important insights can be obtained. First, the distributions of aggregate volatility measures (calculated using entire data) and segmented volatility measures (calculated using censored data) are on-average similar (see the descriptive statistics in Table 3). For instance, the mean coefficient of variation for positive vehicular jerk in longitudinal direction calculated using the entire data is 1.195 ($Longitudinal\ Vol_{pos-jerk,K=1}$) compared to the mean coefficient of variation of 1.168 computed using the first 25-seconds data ($Longitudinal\ Vol_{pos-jerk,K=3}$). Similar observations can be made for other volatility performance measures in longitudinal and lateral directions. This indicates that for the sampled events, drivers were not just volatile immediately before a crash (i.e., 5 s before the crash/near-crash) but also exhibited erratic or volatile behavior well before the crash (as reflected in volatility measures computed using first 20-seconds or 25-seconds data). Based on the discussion presented in earlier sections*, the critical question then becomes how intentional volatility may be associated with crash propensity?* Or in other words, is the relationship between aggregate driving volatility measures and crash propensity significantly different than the relationship between censored driving volatility measures and crash propensity. Note, however, that irrespective of the empirical results, the censoring of driving behavior data for eliminating the bias due to reverse causality is conceptually critical.

Regarding other factors, the average durations of first, second, and third secondary tasks are 2.09, 0.35, and 0.047 s respectively. Regarding secondary tasks, driver was observed to be texting while driving and looking at pedestrians in approximately 2.3% and 0.4% of the events (221 and 39 events). In 3.8% of the events, drivers changed lanes prior to getting involved in one of the three events. Interestingly, driver maneuver was observed to be safe but illegal in 201 events, whereas, maneuver was unsafe but legal in 259 events. The average number of front-seat (including driver) and rear-seat passengers are 1.279 and 0.106 respectively. Most of the events (around 70%) happened under free-flow traffic conditions. To check for multicollinearity, variance inflation factors are reported for all the explanatory variables. A VIF value of less than 10 indicates lack of problematic multicollinearity. In our case, VIF values for all explanatory factors are less than 5 (Table 3).

5. Modeling Results

This section presents the results of fixed- and random-parameter logit models, where crash propensity is modeled as a function of driving volatility related measures and other factors. In particular, owing to the issue of intentional and unintentional volatility (as discussed in Section 3.4), two sets of statistical models are estimated. The first set of statistical models contain the aggregate volatility measures (calculated using the entire time series data) as explanatory factors in addition to other factors. The use of aggregate volatility measures for explaining crash propensity can provide insights regarding how driving volatility relates to occurrence of unsafe outcomes. However, as noted earlier in

Table 3
Descriptive Statistics of Key Variables.

Variable	N	Mean	Std. Dev.	Min	Max	VIF
Key Volatility Indicators (entire 30 seconds driving data)*						
Longitudinal $Vol_{pos-jerk,K=1}$	9593	1.195	0.638	0.28	9.13	4.75
Longitudinal $Vol_{neg-jerk,K=1}$	9593	0.997	0.503	0	6.59	4.73
Lateral $Vol_{pos-jerk,K=1}$	9593	1.107	0.488	0.47	7.43	4.72
Lateral $Vol_{neg-jerk,K=1}$	9593	0.942	0.387	0.43	5.90	4.46
Longitudinal $Vol_{acceleration,K=1}$	9593	0.827	0.323	0	5.17	1.14
Longitudinal $Vol_{deceleration,K=1}$	9593	0.849	0.394	0	6.09	1.65
Lateral $Vol_{acceleration,K=1}$	9593	1.037	0.390	0	5.92	1.28
Lateral $Vol_{deceleration,K=1}$	9593	0.776	0.299	0	4.27	1.26
Key Volatility Indicators (first 20 seconds driving data)*						
Longitudinal $Vol_{pos-jerk,K=2}$	9593	1.024	0.270	0.28	5.70	2.15
Longitudinal $Vol_{neg-jerk,K=2}$	9593	0.836	0.184	0	3.30	1.86
Lateral $Vol_{pos-jerk,K=2}$	9593	1.012	0.309	0.47	14.07	2.36
Lateral $Vol_{neg-jerk,K=2}$	9593	0.854	0.202	0	4.03	2.20
Longitudinal $Vol_{acceleration,K=2}$	9593	0.800	0.303	0	4.16	1.03
Longitudinal $Vol_{deceleration,K=2}$	9593	0.721	0.211	0	2.41	1.12
Lateral $Vol_{acceleration,K=2}$	9593	0.986	0.365	0	6.99	1.07
Lateral $Vol_{deceleration,K=2}$	9593	0.717	0.211	0	2.47	1.16
Key Volatility Indicators (first 25 seconds driving data)*						
Longitudinal $Vol_{pos-jerk,K=3}$	9593	1.168	0.600	0.28	9.06	4.09
Longitudinal $Vol_{neg-jerk,K=3}$	9593	0.984	0.491	0	6.46	4.08
Lateral $Vol_{pos-jerk,K=3}$	9593	1.093	0.476	0.47	8.72	4.01
Lateral $Vol_{neg-jerk,K=3}$	9593	0.930	0.370	0.43	5.62	3.85
Longitudinal $Vol_{acceleration,K=3}$	9593	0.822	0.322	0	5.17	1.07
Longitudinal $Vol_{deceleration,K=3}$	9593	0.840	0.386	0	6.08	1.74
Lateral $Vol_{acceleration,K=3}$	9593	1.029	0.392	0	7.27	1.16
Lateral $Vol_{deceleration,K=3}$	9593	0.767	0.294	0	5.25	1.52
Drivers' Secondary Task Durations						
Secondary Task 1 (duration in seconds)	9593	2.092	2.720	0	24.12	1.16
Secondary Task 2 (duration in seconds)	9593	0.357	1.259	0	14.22	1.12
Secondary Task 3 (duration in seconds)	9593	0.047	0.465	0	15.00	1.07
Adjusting/monitoring other devices, 0 otherwise	9593	0.008	0.092	0	1	1.03
Engaged with cell-phone, dialing hand-held	9593	0.002	0.040	0	1	1.04
Cell-phone, Texting	9593	0.023	0.149	0	1	1.06
Looking at pedestrian	9593	0.004	0.061	0	1	1.09
Pre-Incident Maneuvers						
Changing lanes	9593	0.038	0.190	0	1	1.06
Making U-turn	9593	0.002	0.046	0	1	1.03
Merging	9593	0.003	0.059	0	1	1.03
Passing or overtaking	9593	0.005	0.071	0	1	1.09
Legality of Maneuvers						
Maneuver is safe and legal	9593	0.893	0.309	0	1	1.83
Maneuver is safe but illegal	9593	0.021	0.142	0	1	1.44
Maneuver is unsafe but legal	9593	0.027	0.162	0	1	1.31
Driver Behavior						
Unfamiliar/inexperience with roadway	9593	0.003	0.050	0	1	1.45
Aggressive Driving	9593	0.002	0.046	0	1	1.65
Drowsy, sleepy, fatigued	9593	0.013	0.112	0	1	2.12
Angry	9593	0.005	0.069	0	1	1.12
Number of occupants						
Front Seat Passengers	9593	1.279	0.449	1	3	1.12
Rear Seat Passengers	9593	0.106	0.416	0	5	1.08
Intersection-Roadway Influence						
Intersection influence: Traffic Signal	9593	0.124	0.329	0	1	1.13
Intersection influence: Uncontrolled	9593	0.034	0.181	0	1	1.05
Intersection influence: Stop sign	9593	0.034	0.181	0	1	1.08
Divided Roadway	9593	0.400	0.490	0	1	2.10
Not Divided - 2 way Traffic	9593	0.431	0.495	0	1	2.05
Traffic Flow Factors						
Level of Service: A1	9593	0.406	0.491	0	1	1.78
Level of Service: A2	9593	0.300	0.458	0	1	1.56
Unstable Flow	9593	0.021	0.142	0	1	1.07
Dead Flow	9593	0.010	0.100	0	1	1.04

Notes: N is sample size; SD is standard deviation; and VIF is Variance Inflation Factor; (*) For definitions of the volatility indices, see Table 1 and associated text for explanation; The three different data schemes, entire 30-seconds data, first 20-seconds data, first 25-seconds data, used in calculation of volatility indices are indexed by K = 1, K = 2, and K = 3 respectively (see text for explanation).

Table 4
Overview of Statistical Models Considered in this Study.

Category	Key Volatility Variables as Correlates	Statistical Models	
		Fixed Parameter Logit	Random Parameter Logit
	First set of models: Computed using Aggregate time series data		
Category 1	Acceleration/Deceleration based volatility measures	✓	✓
Category 2	Vehicular jerk based volatility measures	✓	✓
	Second set of models: Computed using censored time series data		
Category 3	Volatility measures computed using first 20 seconds time series data*	✓	✓
Category 4	Volatility measures computed using first 25 seconds time series data*	✓	✓

Notes: (*) the performance measures (vehicular jerk vs. acceleration/deceleration) that provided statistically better results in first set of models are used in the second set of models (category 3 and 4 models).

Section 3.4, the use of aggregate volatility measures will not allow us to relate “intentional volatility” with crash propensity. In this regard, the second set of statistical models contain segmented volatility measures (calculated using censored time series data) as explanatory factors. For convenience, we first present the goodness of fit results of statistical models with aggregate volatility measures followed by presentation of results of statistical models with segmented volatility measures.

5.1. Modeling scheme

Before presenting the results, we briefly present the overview of the modeling scheme. As mentioned earlier, the first set of models use aggregate driving volatility measures that are computed using the entire driving behavior time series data (30 s for crash/near-crash and 20 s for baseline events). Under this setting, for computing driving volatility, two sets of performance measures are considered, i.e., acceleration/deceleration and vehicular jerk. The statistical models with aggregate vehicular jerk or acceleration/deceleration-based volatility measures as explanatory variables are termed as Category 1 and Category 2 models, respectively (see Table 4). The rationale behind considering these two performance measures is to investigate if vehicular jerk-based driving volatility (compared to acceleration/deceleration based) measures better explain crash propensity or vice versa. For each of the two performance measures, fixed and random parameter logit models are developed.

Next, to separate out the two components of driving volatility, i.e., intentional vs. unintentional volatility, second set of statistical models are developed (Table 4). Under this setting, two schemes of censoring mechanisms for time series driving behavior data are considered. In particular, only the first 20 s time series data are used for computing driving volatility measures in the first censoring scheme (see Tables 1 and 4). Whereas, in the second censoring scheme, the first 25 s time series data are used for computing volatility indices and the last 5 s driving data are not considered (Table 4). As explained in detail earlier, the censoring of time series data can help in removing the influence of driver reactions immediately prior to crash or near-crash outcomes from the volatility measures while retaining volatility derived from driver behavior in the seconds leading up to but no immediately before the crash/near-crash event. Thus, using the best-fit performance measure under Category 1 and 2 models, censored versions of volatility indices are considered in Category 3 and 4 models. For each of the two categories, fixed and random parameter logit models are estimated (Table 4).

5.2. Estimation results

The results of statistical models are discussed next that quantify the correlations between crash propensity and driving volatility (aggregate and segmented), after controlling for other traffic, crash, and unobserved factors. First, a series of fixed-parameter logit models are estimated in which the parameter estimates were constrained to be fixed across all events. The fixed parameter multinomial logit models are derived from a systematic process to include most important variables (such as driving volatility related variables and others) on basis of statistical significance, specification parsimony, and intuition. For example, given the key focus, only volatility related variables were first inserted into the crash propensity functions to better understand the relationship between driving volatility and crash propensity. Driving volatility related variables that were statistically significant at 90% confidence level were retained in the corresponding crash propensity functions. Once this was done, other important variables as shown in Table 3 were inserted into the crash propensity functions in a step wise fashion. In doing so, variance inflation factors (VIF) of explanatory factors were examined to avoid multicollinearity issue. As discussed in detail in methodology section, unobserved heterogeneity and omitted variable bias are suspected and in presence of which accurate correlations between driving volatility related measures and crash propensity cannot be established. Therefore, random-parameter logit models are estimated where all the parameter estimates were allowed (and tested) to vary across different events. A parameter estimate that resulted in a statistically significant mean and/or standard deviation was retained as a random parameter in final model specification. If no standard deviation is presented for a particular variable, that variable should be interpreted as fixed parameter. Below, we briefly explain the key model comparison results from the first set of statistical models (using aggregate volatility measures) and the second set of statistical models (using segmented volatility measures). Note that the first set of models are briefly discussed for the sake of completeness and are treated as base models. As discussed in Section 3.4 in detail, a better way of quantifying the correlations between driving volatility and crash propensity is to eliminate the seconds of driving data immediately prior to crash/near-crash from the calculations of driving volatility, as done in the second set of statistical models below.

5.2.1. Statistical models using aggregate volatility measures as explanatory factors (volatility indices in Table 1 with $K = 1$ subscript)

Under the first set of statistical models, eight different volatility measures based on acceleration/deceleration and vehicular jerk in longitudinal and lateral direction are considered (see Table 4). In particular, fixed and random parameter models with acceleration/deceleration-based volatility indices are termed as Category 1 models, whereas models with vehicular jerk-based measures are termed as Category 2 models (Table 4). In the models above, both longitudinal and lateral volatility are considered. Conceptually, we hypothesize that vehicular jerk-based volatility measures may perform better than acceleration-based measures as the earlier accounts for the sharp rate of change (within one-tenth of a second) in acceleration values. Table 5 summarizes the goodness-of-fit results of fixed- and random-parameter Category 1 and 2 logit models. It is seen that random-parameter models in the two categories clearly outperformed their fixed-parameter counterparts. This is evident from the significantly lower AIC and BIC values for random-parameter models (Table 5). Also, the results of

Table 5
Model Comparison Using Aggregate Driving Volatility Measures.

	Category 1 Models		Category 2 Models	
	Random Parameters	Fixed Parameters	Random Parameters	Fixed Parameters
Number of Parameters	45	37	39	33
Log-likelihood at zero	−6202.12	−6202.12	−6202.12	−6202.12
Log-likelihood at convergence	−2568.27	−2609.51	−2001.71	−2044.20
AIC	5226.54	5293.02	4081.42	4154.41
Bozdogan's CAIC	5593.93	5595.10	4399.83	4423.82
Likelihood Ratio Test	Random vs. Fixed Parameters		Random vs. Fixed Parameters	
Likelihood Ratio χ^2 Statistics	82.48		84.98	
DF	8		6	
Critical χ^2 (DF, 0.01)	20.09		16.812	

Notes: Category 1 models include acceleration/deceleration-based volatility measures as explanatory factors; Category 2 models include vehicular jerk-based volatility measures as explanatory factors; AIC is Akaike Information Criterion; CAIC is Bozdogan's Consistent AIC; DF is Degree of Freedom.

Table 6
Model Comparison Using Segmented Driving Volatility Measures.

	Category 3 Models		Category 4 Models	
	Random Parameters	Fixed Parameters	Random Parameters	Fixed Parameters
Number of Parameters	35	33	38	33
Log-likelihood at zero	−6202.12	−6202.12	−6202.12	−6202.12
Log-likelihood at convergence	−4521.09	−4533.75	−2411.88	−2475.94
AIC	9112.18	9133.50	4899.76	5017.88
Bozdogan's CAIC	9397.93	9402.92	5210.00	5287.30
Likelihood Ratio Test	Random vs. Fixed Parameters		Random vs. Fixed Parameters	
Likelihood Ratio χ^2 Statistics	25.32		128.11	
DF	4		5	
Critical χ^2 (DF, 0.01)	13.277		15.086	

Notes: Category 3 models include vehicular jerk-based volatility measures calculated using first 20-seconds time series data as explanatory factors; Category 4 models include vehicular jerk-based volatility measures calculated using first 25-seconds time series data as explanatory factors; AIC is Akaike Information Criterion; CAIC is Bozdogan's Consistent AIC; DF is Degree of Freedom.

likelihood-ratio test are reported which suggest that random-parameter models are statistically superior to their fixed parameter counterparts at 99.5% confidence level (Table 5) (Washington et al., 2010). Specifically, the parameter estimates for eight variables each in Category 1 model, and six variables in Category 2 models are found to be normally distributed random parameters, suggesting significant heterogeneity in the effects of explanatory factors (including aggregate volatility measures) on crash propensity (Table 5). Importantly, both for fixed- and random-parameter approaches, vehicular jerk based longitudinal and lateral volatility measures (category 2 models) performed statistically superior to acceleration-based volatility measures (category 1 models). After adjusting for the degrees of freedom differences, this is evident from the significantly lower AIC and Bozdogan's CAIC values for category 2 models against category 1 models (Table 5). This finding is intuitive and expected as hypothesized earlier.

5.2.2. Statistical models using segmented volatility measures as explanatory factors (volatility indices in Table 1 with $K = 2$ or $K = 3$ subscripts)

To separate out the different components of driving volatility in time to crash/near-crash, this section briefly presents the results of statistical models with segmented driving volatility measures calculated either using first 20-seconds time series data (Category 3 models) or first 25-seconds time series data (Category 4 models). To account for unobserved heterogeneity and omitted variable bias, both fixed and random parameter logit models are estimated. As vehicular jerk-based volatility indices significantly outperformed acceleration/deceleration-based volatility indices (see earlier section), only vehicular jerk-based volatility indices are considered in this set of statistical models. Table 6 summarizes the goodness-of-fit results of fixed- and random-parameter Category 3 and 4 logit models. Several insights can be obtained from the results presented in Table 6. First, models with vehicular jerk-based volatility measures that are calculated using first 25 s time series data (Category 4 models) significantly outperformed the models with volatility measures calculated using the first 20-seconds time series data (Category 3 models). This can be seen from the significantly lower AIC and CAIC values for Category 4 models. However, given the important conceptual and data observability issues presented in section 3.4 regarding event occurrence times in the corresponding video files, we prefer the statistical models based on first 20 s time series data on vehicular jerk.[6]

Second, owing to the presence of unobserved heterogeneity and potential omitted variable bias, the random-parameter models in the two categories clearly outperformed their fixed-parameter counterparts, as indicated by lower AIC/CAIC values for random parameter models and likelihood ratio test statistics favoring random parameter models (Table 6).

Table 7 shows the results of random parameter Category 3 model, i.e., models estimated with segmented vehicular jerk-based volatility measures based on first 20-seconds driving data. To contrast the

[6] The finding that models with segmented volatility measures based on first 25-seconds data resulted in best-fit is intuitive because these models are using more driving behavior data points compared to models with volatility measures based on first 20-seconds driving data. In fact, the evasive maneuvers a driver does just prior to a crash/near-crash are indicators of an unsafe event. However, as we are interested in trying to quantify the effects of "intentional volatility" on crash propensity, we prefer and use the models with segmented volatility measures based on 20-seconds data (as discussed in Section).

Table 7
Estimation Results of Random Parameter Logit Models for Crash Propensity with Segmented Vehicular Jerk Based Driving Volatility Measures*.

Variable	Category 3 Model: Fixed Parameter Logit[1]				Category 3 Model: Random Parameter Logit[1]			
	Crash		Near-Crash		Crash		Near-Crash	
	β	z-score	B	z-score	β	z-score	β	z-score
Constant	−4.192	−15.84	−2.718	−11.59	−4.378	−13.99	−2.785	−9.67
Key Segmented Volatility Indicators (Based on first 20 seconds data bin)*								
Longitudinal $Vol_{pos-jerk,K=2}$	1.679	9.82	1.223	6.9	1.854	8.77	1.281	5.9
Longitudinal $Vol_{neg-jerk,K=2}$	—	—	0.963	4.79	—	—	1.242	3.98
Lateral $Vol_{pos-jerk,K=2}$	2.222	12.64	2.081	11.19	2.180	9.92	2.297	8.07
standard deviation	—	—	—	—	*0.702*	*2.44*	—	—
Lateral $Vol_{neg-jerk,K=2}$	—	—	−0.899	−4.49	—	—	−1.305	−2.85
standard deviation	—	—	—	—	—	—	*1.998*	*3.34*
Drivers' Secondary Task Durations								
Secondary Task 1 (duration in seconds)	0.222	15.05	0.190	15.42	0.235	13.13	0.213	9.61
Secondary Task 2 (duration in seconds)	0.186	6.63	0.193	7.91	0.193	6.14	0.216	6.4
Legality of Maneuvers								
Maneuver is safe and legal	−2.297	−19.44	−2.140	−21.08	−2.417	−15.68	−2.445	−9.91
Maneuver is safe but illegal	−2.306	−7.54	−3.422	−8.32	−2.397	−7.14	−3.884	−6.42
Driver Behavior								
Unfamiliar/inexperience with roadway	2.049	4.26	—	—	2.393	4.38	—	—
Aggressive Driving	—	—	3.220	4.02	—	—	3.625	3.72
Front Seat Passengers	−0.420	−3.97	−0.598	−6.94	−0.444	−3.85	−0.662	−6.11
Intersection-Roadway Influence								
Intersection influence: Traffic Signal	0.766	6.21	1.080	12.08	0.777	5.71	1.306	7.52
Intersection influence: Uncontrolled	1.690	9.19	2.200	14.59	1.709	8.03	2.565	8.11
Divided Roadway	−0.903	−7.29	—	—	−0.980	−6.92	—	—
Not Divided - 2-way Traffic	−0.540	−4.99	0.005	0.06	−0.623	−5.1	−0.208	−1.21
standard deviation	—	—	—	—	—	—	1.038	2.91
Traffic Flow Factors								
Level of Service: A1	—	—	−1.402	−15.67	—	—	−1.697	−7.61
Level of Service: A2	−0.513	−4.56	−0.937	−10.68	−0.539	−4.37	−1.105	−6.99
Unstable Flow	—	—	1.094	6.28	—	—	1.108	3.49
standard deviation	—	—	—	—	—	—	2.238	2.59

Notes: (*) Baseline event is considered base category- all parameter estimates to be interpreted relative to baseline event; (1) refers to models with vehicular jerk-based volatility measures calculated using first 20-seconds time series data. (**) The K = 2 subscript in the indices of volatility measures indicate that first 20-seconds data are used in calculation of volatility measures (see Table 1 for details.).
Notes: (*) Baseline event is considered base category- all parameter estimates to be interpreted relative to baseline event; (1) refers to models with vehicular jerk-based volatility measures calculated using first 20-seconds time series data.

differences, results of fixed parameter Category 3 model are also presented in Table 7. Finally, to better interpret the results, marginal effects are provided in Table 8 for the random parameter Category 3 model shown in Table 7. Marginal effects quantify the potential increase (decrease) in the "exact probability" of observing a specific outcome (event severity) with one-unit increase in a continuous explanatory variable or a switch from 0 to 1 for a dummy variable, with all other variables held constant (Train, 2003; Zhao and Khattak, 2015; Wali et al., 2017a, b). However, if multiplied by 100, it quantifies the "percentage increase (decrease)" in the probability of observing a crash outcome (as an example) with one second increase in the value of a specific continuous variable. The statistics shown in Table 8 are marginal effects multiplied by 100 for ease of interpretation. In discussing the results, we use either of the two interpretations depending on the context. To demonstrate the differences in the effects of explanatory factors on crash propensity, Table 8 also provides the marginal effects for fixed parameter Category 3 model. Given the conceptual motivation presented in Section 3.4, we base our discussion on the results of Category 3 models only. Nonetheless, for completeness, the estimation results of random parameter Category 1 and 2 models are presented in Tables A1 Appendix A. In addition, the estimation results of fixed and random parameter Category 4 models are presented in Table A2 in Appendix A.

While the focus of this study is to explore the heterogeneous correlations between driving volatility and crash propensity, a validation analysis is also performed using hold-out sample to test the accuracy of the predictive power of the best-fit random parameter Category 3 logit model (Zhao and Khattak, 2015). The data are divided into training and testing samples based on a 70/30 split (N = 6685 and 2854 observations for training and testing samples respectively). The best-fit random parameter logit model (see Table 6) was calibrated using the training sample and predictions are then computed using testing, training as well as overall data. To obtain predicted probabilities for the best-fit random parameter logit model, we use 200 number of Halton draws for simulating the probabilities. To minimize the influence of starting chain values of Halton sequence on estimable predicted probabilities, we discard the initial 20 Halton draws as burn-in sample. The results of validation analysis are presented in Table 9.

Referring to the results of random parameter category 3 logit model presented in Table 7, the final model structure is provided below. With baseline event as a reference category, the probability of a crash event can be calculated using Eq. (9) below:

Table 8
Marginal Effects of Fixed- and Random-Parameter Category 3 Models.

Variable	Effects on probabilities of the event outcomes (multiplied by 100)					
	Fixed Parameter Logit Model			Random Parameter Logit Model		
	Baseline	Crash	Near-Crash	Baseline	Crash	Near-Crash
Crash Utility Function						
Longitudinal $Vol_{pos-jerk,K=2}$	−6.66	7.39	−0.73	−12.30	15.79	−3.49
Longitudinal $Vol_{neg-jerk,K=2}$	—	—	—	—	—	—
Lateral $Vol_{pos-jerk,K=2}$	−8.82	9.78	−0.96	−18.29	22.77	−4.48
Lateral $Vol_{neg-jerk,K=2}$	—	—	—	—	—	—
Secondary Task 1 (duration in seconds)	−0.88	0.97	−0.10	−0.86	1.17	−0.31
Secondary Task 2 (duration in seconds)	−1.50	−0.08	1.60	−0.71	0.96	−0.25
Front seat passengers	1.67	−1.85	0.18	1.40	−1.92	0.52
Maneuver is safe and legal	9.11	−10.11	0.99	17.22	−21.23	4.01
Maneuver is safe but illegal	9.15	−10.15	1.00	4.55	−6.33	1.78
Unfamiliar/inexperience with roadway	−8.13	9.01	−0.89	−18.41	23.27	−4.86
Intersection influence: Traffic Signal	−3.04	3.37	−0.33	−3.49	4.58	−1.09
Intersection influence: Uncontrolled	−6.71	7.44	−0.73	−10.62	13.55	−2.93
Divided Roadway	3.58	−3.97	0.39	3.27	−4.50	1.22
Not Divided - 2-way Traffic	2.14	−2.38	0.23	2.31	−3.11	0.80
Level of Service: A2	2.04	−2.26	0.22	1.82	−2.48	0.66
Near-Crash Utility Function						
Longitudinal $Vol_{pos-jerk,K=2}$	−9.86	−0.53	10.39	−10.72	−1.80	12.52
Longitudinal $Vol_{neg-jerk,K=2}$	−7.76	−0.42	8.18	−10.31	−1.74	12.05
Lateral $Vol_{pos-jerk,K=2}$	−16.77	−0.90	17.67	−23.07	−3.26	26.33
Lateral $Vol_{neg-jerk,K=2}$	7.25	0.39	−7.64	3.40	1.10	−4.51
Secondary Task 1 (duration in seconds)	−1.50	−0.08	1.60	−1.34	−0.28	1.61
Secondary Task 2 (duration in seconds)	−1.60	−0.08	1.60	−1.36	−0.28	1.64
Front Seat Passengers	4.82	0.26	−5.07	3.52	0.81	−4.33
Maneuver is safe and legal	17.25	0.92	−18.17	24.13	3.27	−27.40
Maneuver is safe but illegal	27.58	1.48	−29.06	10.27	3.05	−13.32
Aggressive Driving	−25.95	−1.39	27.34	−42.24	−4.93	47.18
Intersection influence: Traffic Signal	−8.70	−0.47	9.17	−10.41	−1.83	12.25
Intersection influence: Uncontrolled	−17.73	−0.95	18.68	−26.52	−3.63	30.15
Not Divided - 2-way Traffic	−0.04	0.00	0.04	−0.30	0.15	0.15
Level of Service: A1	11.30	0.61	−11.90	9.83	2.28	−12.11
Level of Service: A2	7.55	0.40	−7.95	6.46	1.34	−7.80
Unstable Flow	−8.82	−0.47	9.29	−14.95	−1.64	16.59

Note: Marginal effects (multiplied by 100) rounded to nearest two decimals; (—) indicates Not Applicable.

$$
\begin{aligned}
ln\left(\frac{\Pr(event = crash)}{\Pr(event = baseline)}\right) = {} & \beta_{10} + \beta_{11}(Longitudinal\ Vol_{pos-jerk,K=2}) \\
& + \beta_{12i}(Lateral\ Vol_{pos-jerk,K=2}) \\
& + \beta_{13}(\text{Secondary Task 1 (duration in seconds)}) \\
& + \beta_{14}(\text{Secondary Task 2} \\
& \text{(duration in seconds)}) \\
& + \beta_{15}(\text{Maneuver is safe and legal, } 1/0) \\
& + \beta_{16}(\text{Maneuver is safe but illegal, } 1/0) \\
& + \beta_{17}(\text{Unfamiliar/inexperience with} \\
& \text{roadway, } 1/0) \\
& + \beta_{18}(\text{Front Seat Passengers}) \\
& + \beta_{19}\ (\text{Intersection influence: Traffic Signal,} \\
& 1/0) \\
& + \beta_{1\_10}\ (\text{Intersection influence:} \\
& \text{Uncontrolled, } 1/0) \\
& + \beta_{1\_11}\ (\text{Divided Roadway, } 1/0) \\
& + \beta_{1\_12}\ (\text{ Not Divided} - 2 \\
& - \text{way Traffic, } 1/0) \\
& + \beta_{1\_13}\ (\text{Level of Service: A2, } 1/0)
\end{aligned}
\tag{9}
$$

Likewise, with baseline event as a reference category, the probability of a near-crash event can be calculated using Eq. (10):

$$
\begin{aligned}
ln\left(\frac{\Pr(event = near - crash)}{\Pr(event = baseline)}\right) = {} & \beta_{20} + \beta_{21}(Longitudinal\ Vol_{pos-jerk,K=2}) \\
& + \beta_{22}(Longitudinal\ Vol_{neg-jerk,K=2}) \\
& + \beta_{23}(Lateral\ Vol_{pos-jerk,K=2}) \\
& + \beta_{24i}(Lateral\ Vol_{neg-jerk,K=2}) \\
& + \beta_{25}(\text{Secondary Task 1 (duration in seconds)} \\
& + \beta_{26}(\text{Secondary Task 2 (duration in seconds)}) \\
& + \beta_{27}(\text{Maneuver is safe and legal, } 1/0) \\
& + \beta_{28}(\text{Maneuver is safe but illegal, } 1/0) \\
& + \beta_{29}(\text{Aggressive driving, } 1/0) \\
& + \beta_{2\_10}(\text{Front Seat Passengers}) \\
& + \beta_{2\_11}\ (\text{Intersection influence: Traffic} \\
& \text{Signal, } 1/0) \\
& + \beta_{2\_12}\ (\text{Intersection influence:} \\
& \text{Uncontrolled, } 1/0) + \beta_{2\_13i}\ (\text{ Not Divided} \\
& - 2 - \text{way Traffic, } 1/0) \\
& + \beta_{2\_14}\ (\text{Level of Service: A1, } 1/0) \\
& + \beta_{2\_15}\ (\text{Level of Service: A2, } 1/0) \\
& + \beta_{2\_15i}\ (\text{Level of Service: Unstable Flow, } 1/0)
\end{aligned}
\tag{10}
$$

Finally, the two probabilities for crash and near-crash outcomes (calculated through Eqs. (9) and (10)) can be subtracted from 1 to get

Table 9
Predictive Accuracy of Best-Fit Random Parameter Model with Segmented Vehicular-Jerk Based Volatility Measures As Explanatory Factors (Category 3 Model).

Model type and category	Predicted Outcome			
	Baseline	Crash	Near-crash	Total
Overall data				
Baseline	7,374	45	137	7,556
Crash	424	116	135	675
Near-crash	793	71	455	1,319
Total	8,591	232	727	9,550
Percent correctly classified = 7958/9550 = 83.3%				
Percent correctly classified - baselines = 7374/7556 = 97.5%				
Percent correctly classified - crashes = 135/675 = 17.19%				
Percent correctly classified - near-crashes = 455/1319 = 34.5%				
Kendall's τ b = 0.4814				
Training data (in-sample)				
Baseline	5,163	34	96	5,293
Crash	283	92	95	470
Near-crash	549	51	322	922
Total	5,995	177	513	6,685
Percent correctly classified = 5577/6685 = 83.4%				
Percent correctly classified - baselines = 5163/5293 = 97.5%				
Percent correctly classified - crashes = 92/470 = 19.5%				
Percent correctly classified - near-crashes = 322/922 = 34.9%				
Kendall's τ b = 0.4895				
Testing data (out-of-sample)				
Baseline	2,211	11	41	2,263
Crash	141	24	40	205
Near-crash	244	20	133	397
Total	2,596	55	214	2,865
Percent correctly classified = 2368/2865 = 82.6%				
Percent correctly classified - baselines = 2211/2263 = 97.7%				
Percent correctly classified - crashes = 24/205 = 11.7%				
Percent correctly classified - near-crashes = 133/397 = 33.5%				
Kendall's τ b = 0.4624				

Notes: Kendall's τ b is the Kendall's rank correlation coefficients between the nominal observed and predicted crash propensity, ranging between -1 and 1 showing perfectly positive or perfectly negative correlation.
11. APPENDIX A.

the change in probability for baseline outcome. Note that the *"i"* subscript with some of the β parameter estimates in Eqs. (9) and (10) indicate normally distributed random parameters (see Table 7).

6. Discussion

6.1. Safety effects of driving volatility

As mentioned in earlier section, we base our discussion regarding the safety effects of driving volatility on the results obtained from random parameter Category 3 model, i.e., model with vehicular jerk-based volatility measures in longitudinal and lateral directions calculated using first 20-seconds driving data and contrast the results with fixed parameter Category 3 model to highlight the implications of ignoring unobserved heterogeneity and possible omitted variable bias. Furthermore, as the segmented volatility measures are likely to be reflecting the "intentional" driving behavior, we will refer to it as "intentional volatility" in interpreting the findings below.

For crash outcome, the parameter estimates of volatility in positive vehicular jerk both in longitudinal and lateral direction are positive and statistically significant at 95% confidence level (Table 7). This suggests that, compared to baseline events, greater "intentional" volatility is associated with higher likelihood of involvement in a crash event. For example, a one-unit increase in segmented volatility associated with positive vehicular jerk in longitudinal direction ($Longitudinal\ Vol_{pos-jerk,K=2}$) is associated with a 15.79% increase in probability of observing a crash outcome (Table 8), compared to a 7.39% increase indicated by the fixed parameter counterpart (Table 8). That is, if we ignore heterogeneity in the effects of volatility related factors due to systematic variations in unobserved factors, the resulting magnitude of association can be underestimated by almost half. Likewise, a one-unit increase in volatility associated with positive vehicular jerk in lateral direction ($Lateral\ Vol_{pos-jerk,K=2}$) increases the probability of crash outcome by 22.77% (see marginal effects in Table 8). However, the parameter estimate for $Lateral\ Vol_{pos-jerk,K=2}$ (segmented volatility in positive vehicular jerk in lateral direction) is a normally distributed random parameter with a mean of 2.18 and standard deviation of 0.70 (Table 7). This suggests that the associations are not fixed and vary across different events. To visualize the heterogeneity in the effects of random parameters, Fig. 7 shows the distributions of all random parameters in Category 4 model, e.g., see box-plot 1 in Fig. 7 for positive vehicular jerk in lateral direction ($Lateral\ Vol_{pos-jerk,K=2}$).

These findings are important because it imply that greater "intentional volatility" in positive vehicular jerk in time to crash/near-crash makes unsafe outcomes a more probable outcome. In addition, it shows that intentional volatility in positive vehicular jerk in lateral direction has more negative consequences than volatility in positive vehicular jerk in longitudinal direction. We did not find statistically significant association for volatility in negative vehicular jerk in longitudinal and lateral directions[7],[8] .

Coming to the effects of volatility on near-crash outcome, the results intuitively suggest that an increase in segmented volatilities in positive

[7] The important findings that increase in volatility increases the probability of a crash outcome should be interpreted with some caution. Note that some level of volatility will happen naturally due to surrounding driving environments and will not be contributing to crashes (Khattak et al., 2015; Khattak and Wali, 2017). That is, zero volatility is not achievable practically on an average vehicular trip, and does not necessarily imply safer driving. For example, drivers have to respond to other road users and driving tasks required from driving environments by decelerating and accelerating all the time. Thus, the question then becomes what could be a reasonable volatility or a threshold indicating high crash risk? In the literature, several measures are used to identify aggressive or calm driving styles including the standard deviation of acceleration/vehicular jerk, ratio of the standard deviation and the average acceleration in a specific time window, and percentage of time acceleration exceeds a specific threshold (Langari and Won, 2005; ; ; Liu and Khattak, 2016). Along these lines, several key cutoff points for aggressive driving behavior based on acceleration have been reported in the literature. For instance, in the context of urban driving, 1.47 m/s^2 and 2.28 m/s^2 were found as estimates of aggressive and extremely aggressive accelerations thresholds (Kim et al., 2006; Kim and Choi, 2013), compared to the threshold of 0.85 – 1.10 m/s^2 for aggressive driving on urban routes (De Vlieger et al., 2000). Likewise, based on analysis of police crash reports or crash causation studies, strong positive connections between aggressive driving and safety are also reported in the literature (Renski et al., 1999; Paleti et al., 2010). A relatively recent naturalistic driving study by Kim et al. (2016) found a strong correlation between rear-end crash rates and the propensity of hard decelerations (below 4 m/s^2) (Kim et al., 2016). As is evident, the thresholds reported in the literature are (intuitively) inconsistent given that microscopic driving decisions are highly context-specific.

[8] Thus, as part of future work, context-specific thresholds for driving volatility indicating high crash risk in naturalistic settings should be developed. Fundamentally, this translates to the categorization of continuous variables (driving volatility based on acceleration/vehicular jerk, etc.) in crash risk prediction. The number of categories (thresholds) into which a continuous volatility-related variable can be categorized depends mainly on the nature of relationship between volatility and crash risk. As a result, the need for adequate characterization of the relationship must be borne in mind. To this end, one way could be to employ innovative machine-learning based Generalized Additive Modeling framework with tensor- or cubic-based splines to empirically identify thresholds for driving volatility from the data at hand. Importantly, potential safety-critical 'trigger' and 'ceiling' effects can also be identified for key volatility measures (Li et al., 2010;, 2019). Also, the thresholds for low (or high) volatility can vary by environmental conditions and vehicle and roadway types (among other factors) used in the trip. These sources of variations in instantaneous driving decisions suggest that such thresholds should be customized to the context and circumstances.

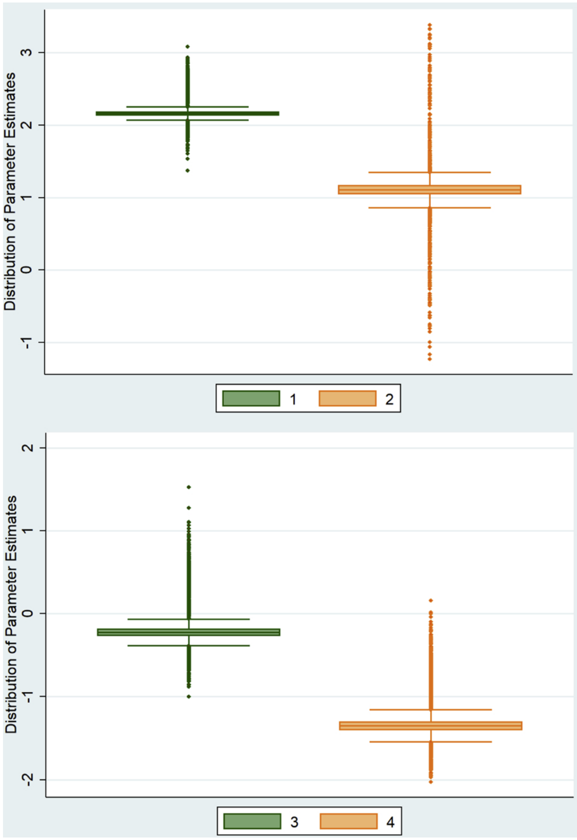

Fig. 7. Distributions of Random-Parameters in Category 3 Model.
Notes:
(1) Volatility (Positive vehicular Jerk: lateral direction) $Lateral\ Vol_{pos-jerk,K=2}$ – utility function of crash outcome;
(2) Unstable traffic flow – utility function of near-crash outcome;
(3) Not Divided – 2-way Traffic – utility function of near-crash outcome;
(4) Volatility (Negative vehicular jerk: lateral direction) $Lateral\ Vol_{neg-jerk,K=2}$ – utility function of crash outcome.

and negative vehicular jerk in longitudinal direction increases the probabilities of observing near-crash events (Table 7). For example, a one-unit increase in segmented volatility associated with positive ($Longitudinal\ Vol_{pos-jerk,K=2}$) and negative vehicular jerk in longitudinal direction ($Longitudinal\ Vol_{neg-jerk,K=2}$) increases the probability of near-crash outcome by 12.52 and 12.05 percentage points, respectively (see marginal effects in Table 8). Likewise, for near-miss outcomes, increase in volatility in positive vehicular jerk in lateral direction ($Lateral\ Vol_{pos-jerk,K=2}$) is significantly and positively correlated with the probability of near-crash outcome (Table 7). In particular, the effect of volatility in positive vehicular jerk in lateral direction ($Lateral\ Vol_{pos-jerk,K=2}$) is significantly greater than the effect of volatility in positive vehicular jerk in longitudinal direction ($Longitudinal\ Vol_{pos-jerk,K=2}$). That is, a one-unit increase in volatility associated with positive vehicular jerk (lateral direction) increases the chance of near-crash outcome by 26.33%, almost double the effect of positive vehicular jerk (longitudinal direction) on near-crash outcome (see Table 8). Again, the parameter estimates for this variable are found to be normally distributed random parameters, suggesting that the effects vary significantly across the observations (see Fig. 7 for the distribution). Finally, an increase in volatility associated with negative vehicular jerk in lateral direction ($Lateral\ Vol_{neg-jerk,K=2}$) is negatively correlated with near-crash occurrence, despite significant heterogeneity not just in magnitude of effects but direction as well. That is, with a mean of -1.305 and standard deviation of 1.998, the direction of association is negative for 74.3% of the observations and positive for the rest (Table 7).

6.2. Safety effects of secondary task durations, passengers, & legality of maneuvers

The estimation results in Table 7 also shed light on the associations between secondary task durations, legality of maneuvers, and crash propensity. Secondary tasks in this study primarily refer to distractions related to non-driving related glances away from the direction of vehicle movement[9] (TRB, 2013). The results suggest that a one-second increase in the duration of the first secondary task increases the probability of crash outcome and near-crash outcome by 0.0117 and 0.0161 units respectively (note that the marginal effects reported in Table 8 are multiplied by 100 for ease of interpretation). Likewise, a one-second increase in the duration of the second secondary task increases the probability of crash and near-crash outcome by 0.96 and 1.64 percentage points respectively (Table 8). These findings are in agreement with previous studies (Klauer et al., 2006), and intuitive as any driver distraction for larger amount of time are likely to result in unsafe outcomes.

Regarding legality of maneuvers, some interesting findings surfaced from the analysis. It is found that if a maneuver is safe (irrespective of being legal or illegal), the likelihood of crashes and near-crashes decreases[10] (Table 7). Likewise, if the driver is unfamiliar with the roadway, the probability of crash outcome increases by 23.27 percentage points, compared to only 9.01 percentage points increase in the fixed-parameter model (Table 8). This finding is in agreement with previous studies, e.g., (Klauer et al., 2006). Finally, Category 3 model suggests that a one-unit increase in number of front seat passengers decreases the probability of crash and near-crash outcome by 0.0192 and 0.0433 units respectively (Table 8). This is intuitive as accompanying front seat passenger may alert or warn the driver in case the driver is anticipated to undertake an unsafe maneuver or action, or perhaps the driver may just have less risky behaviors given the presence of a passenger.

6.3. Safety effects of roadway and traffic flow factors

Compared to baseline events, the random parameter modeling results suggest that the probability of crashes or near-crashes increases at intersections (Table 7). Within intersections, the likelihood of near-crashes is higher on un-controlled intersections compared to intersections with traffic signals (Table 8). Contrary to intersections which have higher likelihood of safety critical events, the likelihood of crashes or near-crashes is on-average lower on roadways (Table 7). These findings are intuitive as intersections generally involve more complex movements and larger number of conflicts (Poch and Mannering, 1996; Chin and Quddus, 2003; Ye et al., 2009; Kamrani et al., 2017). For instance, compared to other roadway types, the chance of crash outcome decreases by 3.11 percentage points on undivided two-way roadways (Table 8). However, the association between undivided two-way roadways and near-crash outcome exhibits significant heterogeneity with a mean of -0.208 and relatively larger standard deviation of 1.038, translating to a negative effect for 57.9% of observations and positive effect for the rest (See Table 7). Note that this variable was statistically insignificant in the fixed parameter counterpart (Table 7). This finding

[9] Some examples of secondary tasks are radio adjustments, seatbelt adjustments, or looking outside at pedestrians. Note that secondary tasks do not include tasks that are critical to the driving task such as speedometer checks, mirror/blink spot checks, activating wipers/headlights, or shifting gears (TRB, 2013).

[10] Note that the maneuver is a vehicle-kinematic based measure and is not related to driver's engagement in secondary tasks and/or distractions. For example, the variable related to legality of maneuvers refer to what the vehicle does (movement and position of the vehicle) such as going straight or changing lane and is not related to what the driver is doing inside the vehicle (such as texting and driving).

is important in the sense that it indicates that ignoring unobserved heterogeneity can lead to inaccurate or misleading results.

Coming to traffic related factors, the results suggest that the likelihood of crashes and/or near-crashes decreases in free-flow traffic either with no lead traffic present (Level of Service: A1) or with lead traffic present (Level of Service: A2) (Table 7). Contrarily, the likelihood of near-crashes increases in unstable traffic flow conditions[11], despite significant heterogeneity in the effects of magnitudes (see parameters in Table 7). These findings are also intuitive as potential of conflict in unstable flow traffic conditions is higher, and thus probability of near-miss may increase (Table 8).

Regarding the predictive accuracy of the best-fit random parameter model, the results in Table 9 reveal reasonably well predictive power of the model with correct classification percentages of 83.3%, 83.4%, and 82.6% for overall, training, and testing data, respectively (Table 9). These overall correct classification percentages are greater than those reported elsewhere, e.g., see (Zhao and Khattak, 2015). Around 35% of the near-crashes and 19.5% of the crashes are correctly classified in the training sample (Table 9). However, the correct classification percentage for crashes dropped to 11.7% in the testing sample (see Table 9). This drop is expected since crashes are rare-events making their prediction relatively challenging. Also, note that while the simulation-based methods employed in this study can be used for predictions, they are largely inferential in the sense that they capture the complex dependencies embedded in the data and may not perform as well as machine learning methods in predicting outcomes (Iranitalab and Khattak, 2017).

7. limitations/future work

The present study focused on exploring the links between event-based volatility and crash propensity irrespective of different types of crash events, such as rear-end, sideswipe, angled, roadway departure, etc. In future, a natural extension of the present study would be to examine how event-based driving volatility varies across different crash event types. As it was not practical to manually check all the thousands of video files for examining the data observability issue (section 3.5), we manually checked the video files of a completely random sample of 100 crashes to exactly record the time at which the event occurred during the 30 s video files. If we could have checked all the video files manually, the resulting distribution of event occurrence times may be different. Also, information on several important factors such as driving age, driver experience, risk perception, is not available in the dataset used in this study. Such important factors should be explicitly controlled for in the analysis. While such factors are not available to the authors, note that the statistical methods used in this study account for all different types of unobserved factors including driving age and/or experience. Finally, while the validation analysis presented provided valuable insights into the predictive power of the best-fit model, it does not explain the model's predictive power in predicting 'low-volatility' crashes and 'high-volatility' baseline events. Thus, building upon the discussion presented in footnotes 5 and 6, future efforts can focus on understanding the critical volatility cut-off points indicating high crash risk with a focus on "low-volatility" crashes and "high-volatility" baselines.

8. Conclusions

Driving behavior in general is considered a leading cause of road traffic crashes. Relevant in this regard is the concept of "driving volatility" that captures the extent of variations in driving, especially hard accelerations/braking and jerky maneuvers. To understand driving volatility prior to involvement in safety critical events, detailed microscopic data on instantaneous driving decisions and safety outcomes are needed. The present study extended the concept of driving volatility to specific events, thus termed as event-based volatility. The key contribution is to find an important predictive relationship between crash propensity and driving volatility (a leading indicator of crash). The SHRP2 Naturalistic Driving Study provides relevant sensor, video, and radar based real-world microscopic driving data in this regard. The present study adopts a rigorous quasi-experimental study design to help compare driving behaviors in normal vs unsafe outcomes. Specifically, crash propensity is defined as likelihood of drivers' involvement in crash- or near-crash events, compared to normal (baseline) driving events. With these forethoughts in mind, the key objective of this study was to examine how driving volatility in time-to-crash or near-crash correlates with crash propensity? To achieve this, an innovative methodology is proposed for characterization of volatility in instantaneous driving decisions in normal and safety-critical events. A total of 2.2 million records of real-world driving for 9593 driving events are analyzed in this study. By using information related to longitudinal and lateral accelerations and vehicular jerk, 24 different aggregate and segmented measures of driving volatility are proposed. In doing so, intentional as well as unintentional event-based driving volatility is characterized, i.e., the issue of actual driving behavior (and the volatility therein) and volatility due to evasive maneuvers of driver to avoid an unsafe outcome (reverse causality) is carefully addressed. Given the important methodological concerns of unobserved heterogeneity and omitted variable bias, fixed- and random-parameter discrete choice models were estimated to reach reliable conclusions.

For all eight aggregate volatility measures, there is statistically significant evidence that driver volatilities in baseline, near-crash, and crash events are significantly different, with volatilities in near-crash and crash events significantly greater than volatility in baseline events. Owing to the issue of intentional vs. unintentional driving volatility, and data observability issue, the final fixed and random parameter models relate crash propensity with driving volatility-based measures calculated using first 20-seconds microscopic driving data, and other observed and unobserved factors. After controlling for traffic, roadway, situational, and other unobserved factors, the results suggest that greater "intentional" volatility is positively correlated with both crash and near-crash events. Importantly, the findings show that greater intentional volatility in positive vehicular jerk (lateral direction) has more negative consequences than volatility in positive vehicular jerk in longitudinal direction. For example, a one-unit increase in (intentional) volatility of positive vehicular jerk in lateral direction increases the probability of crash and near crash outcome by 22.77 and 26.33 percentage points respectively, compared to statistically significant 15.79 and 12.52 percentage point increase in probability of crash and near crash outcome in case of a unit increase in positive vehicular jerk in longitudinal direction respectively. Compared to acceleration/deceleration based volatility measures, empirical evidence suggests that vehicular jerk based volatility models best explain crash propensity. Finally, the correlations established in this study exhibit significant heterogeneity, i.e., the effects of explanatory factors (such as driving volatility) varies across different events and that accounting for the heterogeneous effects in the modeling framework can provide more accurate and reliable results.

The above volatility related findings have important implications for proactive safety. For instance, instantaneous driving decisions can be monitored in real-time and warnings and alerts can be issued to drivers in case driver's decisions in longitudinal and lateral directions exhibit greater volatility (especially in braking). Given that instantaneous driving decisions during deceleration are more volatile and that the effect of volatility in lateral direction on safety outcome is more severe, such alerts and warnings can potentially help in improving safety. From a behavioral perspective, the findings originating from our analysis using segmented volatility indices indicates that it is not just

[11] The variable related to unstable traffic flow conditions was found statistically insignificant in utility function of crash outcome.

the driving volatility immediately prior to crash/near-crash outcome that is critical, but more importantly the volatility in driving decisions when the driver is presumably in control of the vehicle. Given that the volatility in driving decisions well before the driver anticipated an unsafe outcome can be regarded as “intentional” volatility, proactive real-time warning and control assist applications can significantly enhance safety. To this effect, driving monitoring systems and connected and automated vehicles (CAVs) can potentially help in providing alerts and also warn surrounding drivers. The effectiveness of such countermeasures needs to be explored in detail in future research.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The data for this study were provided through a collaborative effort between Virginia Tech Transportation Institute, the U.S. Federal Highway Administration (FHWA), and Oak Ridge National Laboratory (ORNL). The timely assistance and guidance of the ORNL team about data elements is highly appreciated. The contribution of Ms. Ali Boggs in proof-reading the manuscript is highly appreciated. This paper is based upon work supported by the US National Science Foundation under grant No. 1538139. Additional support was provided by the US Department of Transportationthrough the Collaborative Sciences Center for Road Safety, a consortium led by The University of North Carolina at Chapel Hill in partnership with The University of Tennessee. Any opinions, findings, and conclusions or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of the sponsors.

Appendix A

Table A1
Estimation Results of Random Parameter Logit Models for Crash Propensity with Aggregate Driving Volatility Measures*.

Variable	Category 1 Model (Random Parameters Logit)[1]				Category 2 Model (Random Parameters Logit)[2]			
	Crash β	z-score	Near-Crash β	z-score	Crash β	z-score	Near-Crash β	z-score
Constant	−15.15	−13.90	−10.40	−17.02	−29.07	−11.09	−13.10	−16.34
standard deviation	*1.73*	*4.38*	—	—	2.86	5.44	—	—
Key Volatility Indicators**								
Longitudinal $Vol_{pos-jerk,K=1}$	—	—	—	—	3.32	7.11	2.82	8.46
standard deviation	—	—	—	—	*1.23*	*4.19*	*0.83*	*5.11*
Longitudinal $Vol_{neg-jerk,K=1}$	—	—	—	—	8.70	11.35	7.76	14.75
Lateral $Vol_{pos-jerk,K=1}$	—	—	—	—	4.95	6.66	2.55	5.76
Lateral $Vol_{neg-jerk,K=1}$	—	—	—	—	5.63	5.88	−0.20	−0.46
Longitudinal $Vol_{acceleration,K=1}$	1.55	5.37	−0.50	−2.15	—	—	—	—
standard deviation	*0.81*	*3.59*	—	—	—	—	—	—
Longitudinal $Vol_{deceleration,K=1}$	6.61	17.23	8.49	19.70	—	—	—	—
standard deviation	—	—	*0.8*	*3.58*	—	—	—	—
Lateral $Vol_{acceleration,K=1}$	1.85	6.45	0.65	3.66	—	—	—	—
standard deviation	*1.06*	*3.66*	—	—	—	—	—	—
Lateral $Vol_{deceleration,K=1}$	5.92	12.98	3.35	13.10	—	—	—	—
Drivers' Secondary Task Durations								
Secondary Task 1 (duration in seconds)	0.27	8.89	0.21	8.97	0.27	5.32	0.24	8.9
Secondary Task 2 (duration in seconds)	0.17	2.95	0.15	3.24	0.28	2.8	0.15	2.97
Legality of Maneuvers								
Maneuver is safe and legal	−2.67	−9.84	−2.45	−11.74	−4.16	−7.69	−2.85	−12.55
standard deviation	—	—	—	—	*2.93*	*6.37*	—	—
Maneuver is safe but illegal	−2.39	−4.40	−4.17	−6.18	−2.63	−2.61	−5.53	−6.45
Driver Behavior								
Unfamiliar/inexperience with roadway	2.18	2.23	—	—	3.46	2.06	—	—
Aggressive Driving	—	—	2.89	2.80	—	—	3.68	3.4
Front Seat Passengers	−0.31	−1.60	−0.44	−2.91	—	—	−0.49	−3.1
Intersection-Roadway Influence								
Intersection influence: Traffic Signal	0.84	3.64	1.50	8.72	—	—	1.09	6.05
Intersection influence: Uncontrolled	1.38	4.03	1.96	7.40	1.99	3.3	2.80	9.2
Intersection influence: Stop sign	0.82	2.21	1.07	3.47	—	—	—	—
Divided Roadway	−0.41	−2.28	−0.41	−2.28	−1.72	−3.34	−0.43	−2.14
Not Divided – 2-way Traffic	−0.83	−5.01	−0.37	−1.84	−1.13	−2.85	−0.28	−1.29
standard deviation	—	—	*0.85*	*2.36*	—	—	*1.01*	*3.72*
Traffic Flow Factors								
Level of Service: A1	—	—	−1.69	−7.38	—	—	−2.40	−8.81
standard deviation	—	—	*1.04*	*3.19*	—	—	*1.13*	*3.17*
Level of Service: A2	−1.13	−2.83	−0.83	−5.01	−0.71	−1.99	−0.93	−5.36
standard deviation	*1.39*	*2.20*	—	—	—	—	—	—
Unstable Flow	—	—	1.42	4.49	—	—	0.97	2.77
Dead Flow	—	—	1.19	2.32	—	—	0.78	1.61
standard deviation	—	—	*1.45*	*1.99*	—	—	—	—

Notes: (*) Baseline event is considered base category- all parameter estimates to be interpreted relative to baseline event; (1) refers to model with acceleration/deceleration based volatility measures as explanatory factors; (2) refers to model with vehicular jerk based volatility measures as explanatory factors; (*****)*** See Table 1 for definitions of the volatility indices.
Notes: (1) refers to model with acceleration/deceleration-based volatility measures as explanatory factors; (2) refers to model with vehicular jerk based volatility measures as explanatory factors.

Table A2
Estimation Results of Random Parameter Logit Models for Crash Propensity with Segmented Vehicular Jerk Based Driving Volatility Measures*.

Variable	Category 4 Model: Fixed Parameter Logit[1]		Category 4 Model: Random Parameter Logit[1]					
	Crash β	Near-Crash z-score	Crash β	z-score	Near-Crash β	z-score	β	z-score
Constant	−12.95	−30.93	−8.80	−26.39	−27.12	−8.8	−9.73	−21.57
standard deviation	—	—	—	—	*4.52*	*6.85*	—	—
Key Segmented Volatility Indicators (Based on first 25 seconds data bin)*								
Longitudinal $Vol_{pos-jerk,K=3}$	1.92	8.82	1.88	9.17	2.21	4.26	2.15	8.64
standard deviation	—	—	—	—	*2.56*	*5.15*	—	—
Longitudinal $Vol_{neg-jerk,K=3}$	5.51	20.73	5.37	21.56	7.18	10.72	6.71	19.08
Lateral $Vol_{pos-jerk,K=3}$	2.41	8.14	1.95	6.84	5.76	5.79	1.65	4.62
Lateral $Vol_{neg-jerk,K=3}$	1.51	4.9	−0.18	−0.63	5.62	5.08	−0.27	−0.75
Drivers' Secondary Task Durations								
Secondary Task 1 (duration in seconds)	0.27	12.64	0.22	12.97	0.51	6.67	0.24	11.43
Secondary Task 2 (duration in seconds)	0.17	4.26	0.15	4.55	0.29	2.52	0.18	4.43
Legality of Maneuvers								
Maneuver is safe and legal	−2.30	−13.39	−2.22	−16.27	−3.95	−6.44	−2.59	−14.84
standard deviation	—	—	—	—	*−0.79*	*−2.2*	—	—
Maneuver is safe but illegal	−2.23	−4.74	−3.59	−7.27	−2.67	−2.36	−4.71	−6.99
Driver Behavior								
Unfamiliar/inexperience with roadway	1.55	2.79	—	—	5.84	2.97	—	—
Aggressive Driving	—	—	2.85	3.24	—	—	3.09	3.2
Front Seat Passengers	−0.43	−2.74	−0.52	−4.48	−0.49	−2.27	−0.55	−3.94
Intersection-Roadway Influence								
Intersection influence: Traffic Signal	0.50	2.75	0.93	7.20	−2.46	−1.99	1.04	6.72
standard deviation	—	—	—	—	*−4.87*	*−3.45*	—	—
Intersection influence: Uncontrolled	1.89	7.57	2.19	11.06	2.97	4.21	2.61	10.84
Divided Roadway	−0.79	−4.88	—	—	−3.12	−4.54	—	—
Not Divided - 2-way Traffic	−0.76	−4.97	−0.34	−3.23	−2.02	−4.03	−0.67	−3.97
standard deviation	—	—	—	—	—	—	*1.26*	*5.45*
Traffic Flow Factors								
Level of Service: A1	—	—	−1.35	−12.65	—	—	−1.99	−11.88
Level of Service: A2	−0.35	−2.19	−0.73	−6.15	−1.14	−2.42	−0.96	−6.55
Unstable Flow	—	—	0.86	3.58	—	—	0.97	3.19

Notes: (*) Baseline event is considered base category- all parameter estimates to be interpreted relative to baseline event; (1) refers to models with vehicular jerk based volatility measures calculated using first 25-seconds time series data; (**) See Table 1 for definitions of the volatility indices.
Notes: (*) Baseline event is considered base category- all parameter estimates to be interpreted relative to baseline event; (1) refers to models with vehicular jerk-based volatility measures calculated using first 25-seconds time series data.

References

Abdel-Aty, M., Pande, A., 2005. Identifying crash propensity using specific traffic speed conditions. J. Saf. Res. 36 (1), 97–108.
Anastasopoulos, P.C., Mannering, F.L., 2011. An empirical assessment of fixed and random parameter logit models using crash-and non-crash-specific injury data. Accid. Anal. Prev. 43 (3), 1140–1147.
Anastasopoulos, P.C., Tarko, A.P., Mannering, F.L., 2008. Tobit analysis of vehicle accident rates on interstate highways. Accid. Anal. Prev. 40 (2), 768–775.
Antonopoulos, C.N., Germeni, E., Bacopoulou, F., Kalampoki, V., Maltezos, S., Skalkidis, I., Daskalopoulou, S., Negri, E., Petridou, E., 2011. Assessing the impact of risk-taking behavior on road crash involvement among University students residing in two Mediterranean countries. Saf. Sci. 49 (6), 933–938.
Bhat, C.R., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp. Res. Part B Methodol. 37 (9), 837–855.
Boyle, L.N., Tippin, J., Paul, A., Rizzo, M., 2008. Driver performance in the moments surrounding a microsleep. Transp. Res. Part F Traffic Psychol. Behav. 11 (2), 126–136.
Campbell, K.L., 2012. The SHRP 2 naturalistic driving study: addressing driver performance and behavior in traffic safety. TR News 282.
Chen, P., 2015. Built environment factors in explaining the automobile-involved bicycle crash frequencies: a spatial statistic approach. Saf. Sci. 79, 336–343.
Chin, H.C., Quddus, M.A., 2003. Applying the random effect negative binomial model to examine traffic accident occurrence at signalized intersections. Accid. Anal. Prev. 35 (2), 253–259.
Christoforou, Z., Cohen, S., Karlaftis, M.G., 2011. Identifying crash type propensity using real-time traffic data on freeways. J. Saf. Res. 42 (1), 43–50.
De Vlieger, I., De Keukeleere, D., Kretzschmar, J., 2000. Environmental effects of driving behaviour and congestion related to passenger cars. Atmos. Environ. 34 (27), 4649–4655.
Dingus, T.A., Klauer, S.G., Neale, V.L., Petersen, A., Lee, S.E., Sudweeks, J., Perez, M., Hankey, J., Ramsey, D., Gupta, S., 2006. The 100-car Naturalistic Driving Study, Phase II-results of the 100-car Field Experiment.
Dong, C., Richards, S.H., Clarke, D.B., Zhou, X., Ma, Z., 2014. Examining signalized intersection crash frequency using multivariate zero-inflated Poisson regression. Saf. Sci. 70, 63–69.
Feng, F., Bao, S., Sayer, J.R., Flannagan, C., Manser, M., Wunderlich, R., 2017. Can vehicle longitudinal jerk be used to identify aggressive drivers? An examination using naturalistic driving data. Accid. Anal. Prev. 104, 125–136.
FHWA, 2017. Human Factors. Retrieved 05/26/2017, from. Federal Highway Administration Research and Technology. https://www.fhwa.dot.gov/research/topics/safety/humanfactors/.
Greene, W.H., 2011. Econometric Analysis (Seventh Edition). Pearson Education.
Greene, W.H., Hensher, D.A., 2010. Modeling Ordered Choices: a Primer. Cambridge University Press.
Hankey, J.M., Perez, M.A., McClafferty, J.A., 2016. Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets. Virginia Tech Transportation Institute.
Hassan, H.M., Shawky, M., Kishta, M., Garib, A.M., Al-Harthei, H.A., 2017. Investigation of drivers' behavior towards speeds using crash data and self-reported questionnaire. Accid. Anal. Prev. 98, 348–358.
Hayward, J.C., 1972. Near Miss Determination Through Use of a Scale of Danger.
Henclewood, D., 2014. Safety Pilot Model Deployment – One Day Sample Data Environment Data Handbook. U. D. O. T. Research and Technology Innovation Administration. Research and Technology Innovation Administration, McLean, VA.
Iranitalab, A., Khattak, A., 2017. Comparison of four statistical and machine learning methods for crash severity prediction. Accid. Anal. Prev. 108, 27–36.
Ivan, J.N., Wang, C., Bernardo, N.R., 2000. Explaining two-lane highway crash rates using

land use and hourly exposure. Accid. Anal. Prev. 32 (6), 787–795.
Ivers, R., Senserrick, T., Boufous, S., Stevenson, M., Chen, H.-Y., Woodward, M., Norton, R., 2009. Novice drivers' risky driving behavior, risk perception, and crash risk: findings from the DRIVE study. Am. J. Public Health 99 (9), 1638–1644.
Jovanis, P., Aguero-Valverde, J., Wu, K.-F., Shankar, V., 2011. Analysis of naturalistic driving event data: omitted-variable bias and multilevel modeling approaches. Transp. Res. Rec. J. Transp. Res. Board 2236, 49–57.
Kamrani, M., Wali, B., Khattak, A.J., 2017. Can Data Generated by Connected Vehicles Enhance Safety? Proactive Approach to Intersection Safety Management (Accepted). Transp. Res. Rec. J. Transp. Res. Board.
Khattak, A., Nambisan, S., Chakraborty, S., 2015. Study of Driving Volatility in Connected and Cooperative Vehicle Systems. National Science Foundation https://nsf.gov/awardsearch/showAward?AWD_ID=1538139." from http://tesp.engr.utk.edu/research21/projects.php.
Khattak, A.J., Wali, B., 2017. Analysis of volatility in driving regimes extracted from basic safety messages transmitted between connected vehicles. Transp. Res. Part C Emerg. Technol. 84, 48–73.
Kim, E., Choi, E., 2013. Estimates of Critical Values of Aggressive Acceleration From a Viewpoint of Fuel Consumption and Emissions.
Kim, H., Oh, J.-S., Jayakrishnan, R., 2006. Application of activity chaining model incorporating a time use problem to network demand analysis. Transp. Res. Rec. 1977 (1), 214–224.
Kim, S., Song, T.-J., Rouphail, N.M., Aghdashi, S., Amaro, A., Gonçalves, G., 2016. Exploring the association of rear-end crash propensity and micro-scale driver behavior. Saf. Sci. 89, 45–54.
Klauer, S.G., Dingus, T.A., Neale, V.L., Sudweeks, J.D., Ramsey, D.J., 2006. The Impact of Driver Inattention on near-crash/crash Risk: an Analysis Using the 100-car Naturalistic Driving Study Data.
Langari, R., Won, J.-S., 2005. Intelligent energy management agent for a parallel hybrid vehicle-part I: system architecture and design of the driving situation identification process. IEEE Trans. Veh. Technol. 54 (3), 925–934.
Lee, C., Abdel-Aty, M., 2008. Testing effects of warning messages and variable speed limits on driver behavior using driving simulator. Transp. Res. Rec. J. Transp. Res. Board 2069, 55–64.
Lee, J., Abdel-Aty, M., Choi, K., 2014. Analysis of residence characteristics of at-fault drivers in traffic crashes. Saf. Sci. 68, 6–13.
Li, X., Lord, D., Zhang, Y., 2010. Development of accident modification factors for rural frontage road segments in Texas using generalized additive models. J. Transp. Eng. 137 (1), 74–83.
Liu, J., Khattak, A., Wang, X., 2015a. The role of alternative fuel vehicles: using behavioral and sensor data to model hierarchies in travel. Transp. Res. Part C Emerg. Technol. 55, 379–392.
Liu, J., Khattak, A.J., 2016. Delivering improved alerts, warnings, and control assistance using basic safety messages transmitted between connected vehicles. Transp. Res. Part C Emerg. Technol. 68, 83–100.
Liu, J., Khattak, A.J., Richards, S.H., Nambisan, S., 2015b. What are the differences in driver injury outcomes at highway-rail grade crossings? Untangling the role of precrash behaviors. Accid. Anal. Prev. 85, 157–169.
LP, S.C., 2005. Stata Statistical Software Release 9. Stata Press Publication.
Ma, Z., Zhang, H., Steven, I., Chien, J., Wang, J., Dong, C., 2017. Predicting expressway crash frequency using a random effect negative binomial model: a case study in China. Accid. Anal. Prev. 98, 214–222.
Machin, M.A., Sankey, K.S., 2008. Relationships between young drivers' personality characteristics, risk perceptions, and driving behaviour. Accid. Anal. Prev. 40 (2), 541–547.
Mannering, F., 2009. An empirical analysis of driver perceptions of the relationship between speed limits and safety. Transp. Res. Part F Traffic Psychol. Behav. 12 (2), 99–106.
Mannering, F.L., Shankar, V., Bhat, C.R., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.
Martin, J.-L., 2002. Relationship between crash rate and hourly traffic flow on interurban motorways. Accid. Anal. Prev. 34 (5), 619–629.
McDonald, A., et al., 2017. Using kinematic driving data to detect sleep apnea treatment adherence. J. Intel. Transport. Syst. 21 (5), 422–434. https://doi.org/10.1080/15472450.2017.1369060.
McFadden, D., 1973. Conditional Logit Analysis of Qualitative Choice Behavior.
Mitchell, R., Bambach, M., Friswell, R., 2014. Work and non-work-related vehicle crashes: the contribution of risky driving practices. Saf. Sci. 68, 65–72.
Murphey, Y.L., 2008. Intelligent vehicle power management: an overview. Computational Intelligence in Automotive Applications. Springer, pp. 169–190.
Naik, B., Tung, L.-W., Zhao, S., Khattak, A.J., 2016. Weather impacts on single-vehicle truck crash injury severity. J. Saf. Res. 58, 57–65.
Neyens, D.M., Boyle, L.N., 2007. The effect of distractions on the crash types of teenage drivers. Accid. Anal. Prev. 39 (1), 206–212.
NHTSA, 2015. National highway traffic safety administration: the economic and societal impact of motor vehicle crashes, 2010 (Revised). Ann. Emerg. Med. 66 (2), 194–196.
Paleti, R., Eluru, N., Bhat, C.R., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accid. Anal. Prev. 42 (6), 1839–1854.
Perkins, S.R., Harris, J.L., 1968. Traffic conflict characteristics-accident potential at intersections. Highway Res. Rec. 225.
Poch, M., Mannering, F., 1996. Negative binomial analysis of intersection-accident frequencies. J. Transp. Eng. 122 (2), 105–113.
Qin, X., Ivan, J.N., Ravishanker, N., 2004. Selecting exposure measures in crash rate prediction for two-lane highway segments. Accid. Anal. Prev. 36 (2), 183–191.
Qu, X., Yang, Y., Liu, Z., Jin, S., Weng, J., 2014. Potential crash risks of expressway on-ramps and off-ramps: a case study in Beijing, China. Saf. Sci. 70, 58–62.
Quddus, M., 2013. Exploring the relationship between average speed, speed variation, and accident rates using spatial statistical models and GIS. J. Transp. Saf. Secur. 5 (1), 27–45.
Renski, H., Khattak, A.J., Council, F.M., 1999. Effect of speed limit increases on crash injury severity: analysis of single-vehicle crashes on North Carolina interstate highways. Transp. Res. Rec. 1665 (1), 100–108.
Salman, N.K., Al-Maita, K.J., 1995. Safety evaluation at three-leg, unsignalized intersections by traffic conflict technique. Transp. Res. Rec. 1485 (28), 177–185.
Sarker, A.A., Paleti, R., Mishra, S., Golias, M.M., Freeze, P.B., 2017. Prediction of secondary crash frequency on highway networks. Accid. Anal. Prev. 98, 108–117.
Savolainen, P., Mannering, F., 2007. Probabilistic models of motorcyclists' injury severities in single-and multi-vehicle crashes. Accid. Anal. Prev. 39 (5), 955–963.
Savolainen, P.T., Mannering, F.L., Lord, D., Quddus, M.A., 2011. The statistical analysis of highway crash-injury severities: a review and assessment of methodological alternatives. Accid. Anal. Prev. 43 (5), 1666–1676.
Schneider, R., Khattak, A., Zegeer, C., 2001. Method of improving pedestrian safety proactively with geographic information systems: example from a college campus. Transp. Res. Rec. J. Transp. Res. Board 1773, 97–107.
Schneider, R.J., Ryznar, R.M., Khattak, A.J., 2004. An accident waiting to happen: a spatial approach to proactive pedestrian planning. Accid. Anal. Prev. 36 (2), 193–211.
Scott-Parker, B., Oviedo-Trespalacios, O., 2017. Young driver risky behaviour and predictors of crash risk in Australia, New Zealand and Colombia: same but different? Accid. Anal. Prev. 99, 30–38.
Shrestha, D., Lovell, D., Tripodis, Y., 2017. Hardware and software for collecting microscopic trajectory data on naturalistic driving behavior. J. Intel. Transport. Syst. 21 (3), 202–213. https://doi.org/10.1080/15472450.2017.1283224.
Smorti, M., Guarnieri, S., 2014. Sensation seeking, parental bond, and risky driving in adolescence: some relationships, matter more to girls than boys. Saf. Sci. 70, 172–179.
Train, K., 2003. Discrete Choice Methods With Simulation. Cambridge university press.
TRB, 2013. The 2nd strategic highway research program naturalistic driving study dataset. T. R. B. o. t. N. A. o. Science Available from the SHRP2 NDS InSight Data Dissemination.
Tronsmoen, T., 2010. Associations between driver training, determinants of risky driving behaviour and crash involvement. Saf. Sci. 48 (1), 35–45.
Wali, B., 2018. Harnessing Big Data for Characterizing Driving Volatility in Instantaneous Driving Decisions – Implications for Intelligent Transportation Systems. Ph.D. Thesis Dissertation. The University of Tennessee, Knoxville.
Wali, B., Ahmed, A., Ahmad, N., 2017a. An ordered-probit analysis of enforcement of road speed limits. Proc. Inst. Civil Eng. Transp. 1–10.
Wali, B., Ahmed, A., Iqbal, S., Hussain, A., 2017b. Effectiveness of enforcement levels of speed limit and drink driving laws and associated factors–Exploratory empirical analysis using a bivariate ordered probit model. J. Traffic Transp. Eng. 4 (3), 272–279.
Wali, B., Chapman, J., Frank, L., 2019. Task 4 – Benchmarking Methodology and Pilot Application. National Environmental Database – Walk Index, Benchmarks, & Investments. Urban Design 4 Health Inc. Project No.: 74518. Supported by the Robert Wood Johnson Foundation, Rochester (NY).
Wali, B., Khattak, A., 2019. Is driving volatility in school zones related to crash propensity in a naturalistic driving environment? Harnessing cyber physical systems for proactive safety management (under-review in journal). Presented at the Transportation Research Board 98th Annual Meeting.
Wali, B., Khattak, A., David, G., Liu, J., 2018a. Fuel economy gaps within & across garages: a bivariate random parameters seemingly unrelated regression approach (forthcoming). Int. J. Sustain. Transp.
Wali, B., Khattak, A.J., Bozdogan, H., Kamrani, M., 2018b. How is driving volatility related to intersection safety? A Bayesian heterogeneity-based analysis of instrumented vehicles data. Transp. Res. Part C Emerg. Technol. 92, 504–524.
Wali, B., Khattak, A.J., Karnowski, T., 2018c. How driving volatility in time to collision relates to crash severity in a naturalistic driving environment? Presented at the Transportation Research Board 97th Annual Meeting 2018.
Wali, B., Khattak, A.J., Li, X., 2018d. Exploring non-linear dependencies in correlates of roadway crashes: application of generalized additive & piecewise linear count data models (paper No.: No. 18-00064). Transportation Research Board 97th Annual Meeting.
Wang, X., Khattak, A.J., Liu, J., Masghati-Amoli, G., Son, S., 2015. What is the level of volatility in instantaneous driving decisions? Transp. Res. Part C Emerg. Technol. 58, 413–427.
Washington, S.P., Karlaftis, M.G., Mannering, F., 2010. Statistical and Econometric Methods for Transportation Data Analysis. CRC press.
Weber, E.U., Shafir, S., Blais, A.-R., 2004. Predicting risk sensitivity in humans and lower animals: risk as variance or coefficient of variation. Psychol. Rev. 111 (2), 430.
Weng, J., Meng, Q., 2012. Effects of environment, vehicle and driver characteristics on risky driving behavior at work zones. Saf. Sci. 50 (4), 1034–1042.
WHO, 2015. Global Status Report on Road Safety 2015. World Health Organization.
Wooldridge, J.M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT

press.
Wu, K.-F., Jovanis, P.P., 2012. Crashes and crash-surrogate events: exploratory modeling with naturalistic driving data. Accid. Anal. Prev. 45, 507–516.
Wu, K.-F., Jovanis, P.P., 2013. Defining and screening crash surrogate events using naturalistic driving data. Accid. Anal. Prev. 61, 10–22.
Yan, X., Abdel-Aty, M., Radwan, E., Wang, X., Chilakapati, P., 2008. Validating a driving simulator using surrogate safety measures. Accid. Anal. Prev. 40 (1), 274–288.
Ye, X., Pendyala, R.M., Washington, S.P., Konduri, K., Oh, J., 2009. A simultaneous equations model of crash frequency by collision type for rural intersections. Saf. Sci. 47 (3), 443–452.
Zhao, S., Khattak, A., 2015. Motor vehicle drivers' injuries in train–motor vehicle crashes. Accid. Anal. Prev. 74, 162–168.