Transportation Research Part C 92 (2018) 504–524

Contents lists available at ScienceDirect

Transportation Research Part C

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

How is driving volatility related to intersection safety? A Bayesian heterogeneity-based analysis of instrumented vehicles data

Behram Wali[a], Asad J. Khattak[a,*], Hamparsum Bozdogan[b], Mohsen Kamrani[a]

[a] Department of Civil & Environmental Engineering, The University of Tennessee, USA
[b] Department of Business Analytics & Statistics, The University of Tennessee, USA

ARTICLE INFO

Keywords:
Connected vehicles
Driving volatility
Intersection crashes
Proactive safety
Full Bayes estimation
Markov Chain Monte Carlo
Gibbs sampler
Heterogeneity
Random parameters
Poisson/Poisson log-normal regressions

ABSTRACT

Driving behavior in general is considered a leading cause of intersection related traffic crashes. However, due to unavailability of real-world driving data, intersection safety performance evaluations are largely reactive where state-of-the-art methods are applied to analyze historical crash data. In this regard, the emerging connected vehicles technology provides a promising opportunity for investigating intersection safety more from a proactive perspective. Driving volatility captures the extent of variations in instantaneous driving decisions when a vehicle is being driven. This study develops a fundamental understanding of microscopic driving volatility and how it relates to unsafe outcomes at intersections. Using high resolution driving data from a real-world connected vehicle testbed, Safety Pilot Model Deployment, in Ann Arbor, Michigan, a methodology is presented to quantify driving volatility at 116 intersections by analyzing more than 230 million real-world Basic Safety Messages. For proactive intersection safety evaluation, the large-scale connected vehicle data is then linked to detailed intersection data containing crashes, traffic exposure, and other geometric features. By using vehicular speed, acceleration/deceleration, and vehicular jerk based eight different volatility measures, descriptive analysis is performed to spot differences between driving volatility at signalized and un-signalized intersections. Then, in-depth statistical analysis is conducted separately for all intersections (signalized and un-signalized) and signalized intersections only. Importantly, not all factors that may influence crash frequency can be observed in the data. If unobserved factors could be included in a model, then correlations between driving volatility and crash frequency can change, e.g., the relationship can become statistically insignificant. Given the important methodological concerns of unobserved heterogeneity and potential omitted variable bias, hierarchical fixed- and random-parameter Poisson and Poisson log-normal models are estimated. Full Bayesian estimation via Markov Chain Monte Carlo (MCMC) based Gibbs sampling is performed, providing more efficient results. For all intersections, after controlling for traffic exposure, geometrics, and unobserved factors, a one-percent increase in intersection-level volatility calculated through two standard deviations threshold for acceleration/deceleration, passing level volatility captured through coefficient of variation of speed, and mean absolute deviance of vehicular jerk results in a 1.25%, 0.25%, and 0.35% increase in crash frequencies respectively. However, the relationships between intersection-specific volatility and crash frequencies are different for signalized intersections. Several of the exogenous factors are found to be normally distributed random parameters, suggesting that the effects of such variables vary across different intersections. The implications of the findings for proactive safety management are discussed.

* Corresponding author.
E-mail addresses: bwali@vols.utk.edu (B. Wali), akhattak@utk.edu (A.J. Khattak), bozdogan@utk.edu (H. Bozdogan), mkamrani@vols.utk.edu (M. Kamrani).

https://doi.org/10.1016/j.trc.2018.05.017
Received 26 May 2017; Received in revised form 11 May 2018; Accepted 19 May 2018

1. Introduction

Intersections are believed to be the most dangerous locations on roadways potentially due to the complex traffic movements that result in large number of vehicular conflicts, and the diverse set of operational and geometric features associated with them (Zheng and Liu, 2017; Persaud and Nguyen, 1998; Hashimoto et al., 2016; Rakha and Kamalanathsharma, 2011; Rakha et al., 2007). As such, improving roadway intersection safety is of high interest to the profession. Through application of diverse set of advanced empirical methods, researchers since decades have come up with intersection targeted safety performance models (Quddus et al., 2001; Muralidharan et al., 2016; El-Basyouny and Sayed, 2013). Typically, intersection safety performance evaluations are largely reactive, where state-of-the-art methods are applied to link historical crash data with crash specific, operational, and geometric related features, to name a few (Lord and Mannering, 2010). Given information about the afore-mentioned characteristics specific to each intersection, crashes can then be predicted based on which appropriate safety treatments are developed and recommended (Braitman et al., 2007; Tay and Rifaat, 2007).

Driving behavior and/or human factors in general are considered a leading cause of intersection traffic crashes (Akamatsu et al., 2003; Zimmerman and Bonneson, 2004). Importantly, volatility in instantaneous driving decisions can be a leading indicator for understanding the occurrence of unsafe outcomes such as incidents or crashes (Wali et al., 2018d; Khattak and Wali, 2017). The concept of driving volatility captures the extent of variations in instantaneous driving decisions (such as variations in speed) when a vehicle is being driven at a specific roadway location (Wang et al., 2015; Khattak et al., 2015; Liu and Khattak, 2016). Such information can help in identification of intersection locations where crashes may not have happened yet but are perhaps waiting to happen (Schneider et al., 2004). In this regard, connected vehicles technology provides a promising avenue for investigating intersection safety, more from a proactive perspective. With monitoring, processing, and adequate integration of connected vehicle data with historical crash data, the generated large-scale empirical data from connected vehicle systems have significant potential in facilitating deeper understanding of instantaneous driving decisions, and to link microscopic driving decisions to unsafe safety outcomes. The Safety Pilot Model Deployment (SPMD) offers detailed instantaneous driving data generated by connected vehicles in real-world environment. This pilot, sponsored by US-DOT, is currently on-going in Ann Arbor, Michigan, and intends to display vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communication systems in real-life environment. Of specific interest are the Basic Safety Messages (BSMs) that provide high-frequency (usually ten times per second) information packets containing detailed data on vehicle's motion, location, driving context, and instantaneous driving decisions (e.g., speed).

This study focuses on extending the concept of driving volatility to specific intersections, thus termed as Intersection-Based Volatility, by using real-world large-scale connected vehicle data. Given that navigating through an intersection is a complex task, we posit that the concept of intersection-based volatility can provide critical insights regarding the correlations between driving behaviors (its extent and variation) at a specific intersection and key safety outcomes. Using large-scale real world microscopic driving data, a methodology for conceptualizing and quantifying driving volatility at individual intersections is presented. Then, for proactive intersection safety management, driving volatilities at specific intersections are linked to detailed intersection data containing crashes, traffic exposure, and other geometric features. From a methodological perspective, appropriate count data models are developed within Full Bayesian framework, and which accounts for the important issues of omitted variable bias and unobserved heterogeneity (discussed later in detail).

2. Literature review

A careful review of literature reflects the prompt response by government agencies, automotive industry and academia to such disruptive yet beneficial connected and automated vehicles innovation. This innovation has an unquestionable potential to significantly improve the current transportation systems (Ghiasi et al., 2017a; Wali et al., 2018d; Liu and Khattak, 2016; Khattak and Wali, 2017; Zeng et al., 2017), with some recent studies showing benefits in form of comprehensive crash savings, fuel efficiency, parking benefits, and travel time reduction to approach $4000 per year for each CAV operated on road (Fagnant and Kockelman, 2014). From a research perspective, a wide range of reliable transportation connectivity solutions are explored to address real world safety challenges, mobility issues, and environmental challenges (Zulkefli et al., 2014; Zulkefli et al., 2017; Khattak et al., 2015; Liu and Khattak, 2016; Khattak and Wali, 2017; Letter and Elefteriadou, 2017; Zheng and Liu, 2017; Arvin et al., 2018; Wali et al., 2018d; Ghiasi et al., 2017b).

Connected vehicle solutions can potentially help in addressing transportation challenges by primarily targeting the human factor involved in surface transportation. From safety perspective, driving behavior and/or human factors in general are considered a leading cause of traffic crashes (Akamatsu et al., 2003; Zimmerman and Bonneson, 2004; Kamrani et al., 2014; Arvin et al., 2017). Since decades, by analyzing traditional crash, roadway, and geometric data, researchers have developed safety performance models for designing effective safety countermeasures. However, the intersection safety countermeasures in particular and road safety countermeasures in general are typically reactive in nature, i.e., roadway or intersection geometric improvements are developed once crashes happen, and are not specifically designed based on the driver's behavior which in turn is a major cause of unsafe outcomes at intersections. This largely can be attributed to the intrinsic limitations of traditional crash data which do not provide detailed information about drivers' performance, and typically surrogate measures such as speed limits or controlled simulation experiments are used to evaluate drivers' performance (Shah et al., 2018; Haglund and Åberg, 2000; Aarts and Van Schagen, 2006; Bao and Boyle, 2008; Wali et al., 2017b; Wali et al., 2017c). On the other hand, the rapid technological sensor and driving surveillance advancements in recent years have enabled collection of huge amounts of spatiotemporal data about vehicle and human movement (Ghasemzadeh et al., 2018). For example, by using Global Positioning System (GPS) based taxi data, Pei et al. (2012) investigated