Accident Analysis and Prevention 71 (2014) 93–105

Contents lists available at ScienceDirect

Accident Analysis and Prevention

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

Safety propensity index for signalized and unsignalized intersections: Exploration and assessment

Justin P. Schorr*, Samer H. Hamdar[1]

Department of Civil and Environmental Engineering, Center of Intelligent Systems Research, Traffic and Networks Research Laboratory, The George Washington University, 20101 Academic Way #201-I, Ashburn, VA 20147, USA

ARTICLE INFO

Article history:
Received 22 December 2013
Received in revised form 9 May 2014
Accepted 13 May 2014

Keywords:
Index
Intersections
Safety
Signals
Structural equations
Traffic

ABSTRACT

The objective of this study is to develop a safety propensity index (SPI) for both signalized and unsignalized intersections. Through the use of a structural equation modelling (SEM) approach safety is quantified in terms of multiple endogenous variables and related to various dimensions of exogenous variables. The singular valued SPI allows for identification of relationships between variables and lends itself well to a comparative analysis between models. The data provided by the Highway Safety Information System (HSIS) for the California transportation network was utilized for analysis. In total 22,422 collisions at unsignalized intersections and 20,215 collisions at signalized intersections (occurring between 2006 and 2010) were considered in the final models. The main benefits of the approach and the subsequent development of an SPI are (1) the identification of pertinent variables that effect safety at both intersection types, (2) the identification of similarities and differences at both types of intersections through model comparison, and (3) the quantification of safety in the form of an index such that a ranking system can be developed. If further developed, the adopted methodology may assist in safety related decision making and policy analysis.

1. Introduction

Creating a safer, more efficient transportation network is a main goal of transportation researchers across the world leading to countless studies that analyze different safety aspects as they pertain to transportation. In these studies the quantification of safety (a difficult and ambitious undertaking) is approached in a variety of manners and has a vast array of applications. One such application is the analysis of safety at roadway intersections. Intersection related collisions are of particular concern to transportation researchers as they accounted for 47% of all vehicles involved in collisions and 28% of those involved in fatal collisions on United States roadways in 2010 (NHTSA, 2012). Signalized intersections accounted for 25% of total vehicles in collisions and 8% of vehicles involved in fatal collisions while unsignalized intersections accounted for 22% of total vehicles involved and 20% of vehicles involved in fatal collisions (NHTSA, 2012).

Research in this study is aimed at quantifying and analyzing safety at both signalized and unsignalized intersections by exploring a comprehensive modelling framework that quantifies safety in terms of multiple endogenous variables and considers the combined effects of various exogenous variables on safety as well as on one another. One such framework that allows for the inclusion of these parameters is that of structural equation modelling (SEM). The main benefit of this approach is the quantification of safety in terms of a safety propensity index (SPI). The SPI is a singular value that relates all exogenous variables to all endogenous variables and allows for the development of a ranking system in order to understand and evaluate safety at intersections. Additionally, the estimation of structural models for both signalized and unsignalized conditions creates a powerful comparative framework from which additional analysis can be conducted. This methodology has been utilized by Hamdar et al. (2008) to examine aggressiveness at signalized intersections as well as by Hamdar and Schorr (2013) in order to compare and contrast safety in differing flow scenarios (interrupted and uninterrupted). Within the SEM framework, research in this study will explore further applications of the modelling technique – and a refined analysis approach is proposed to better understand and compare the factors affecting safety in both driving scenarios.

* Corresponding author. +1 2158501886.
E-mail addresses: justin11@gwu.edu (J.P. Schorr), hamdar@gwu.edu (S.H. Hamdar).
[1] Tel.: +1 202 994 6652; fax: +1 202 994 0127.

http://dx.doi.org/10.1016/j.aap.2014.05.008
0001-4575/

Specific research goals to be accomplished throughout this study are as follows: (1) to systematically identify the factors that affect safety propensity at both signalized and un-signalized intersections in a given area; (2) to utilize existing public data repositories (Highway Safety Information System (HSIS)) to study the safety implications of changes in network geometry as an evolving system (in time and space); (3) to validate the formulated structural equation model against alternative model structures estimated using the existing signalized and un-signalized intersection's incident data; and (4) to analyze the validated model to compare the results obtained given the intersection type. The findings may help in understanding how better transportation system performance can be achieved and strategies can be proposed to improve traffic safety and operations.

2. Conceptual framework and background

A great amount of effort has gone into assessing the different factors that contribute to collisions at both signalized and unsignalized intersections. Previous studies have utilized a number of different modelling techniques for both analytical and predictive purposes. Throughout the literature safety is commonly characterized through a single endogenous metric such as collision rate (Vogt and Bared, 1998; Chin and Quddus, 2003; Wang and Abdel-Aty, 2006; Isebrands et al., 2010; Caliendo and Guida, 2012; Wu et al., 2013) or injury/collision severity (Abdel-Aty and Abdelwahab, 2004; Xie et al., 2009; Quddus et al., 2010; Jung et al., 2012) and expansions to this characterization often focus around the categorical grouping of collisions by collision type (Wang et al., 2003; Kim et al., 2006; Bham et al., 2012). The "safety framework" is further defined through the identification and selection of exogenous variables – which are typically related to one of the following categories: environmental conditions, geometric design, traffic characteristics, driver demographics or vehicle characteristics. Many studies examine a combination of two or three variable types such as geometric design/traffic characteristics (Li et al., 1994; Poch and Mannering, 1996; Karlaftis and Golias, 2001; Wang et al., 2003; Wu et al., 2013), driver demographics/environmental conditions (Bham et al., 2012; Jung et al., 2012), or geometric design/traffic characteristics/environmental conditions (Kim et al., 2006; Quddus et al., 2010). While these studies examine the combined effects of these variables on safety, the modelling framework utilized does not allow for examination of the effects individual variables have on one another or for analysis to be conducted on a dimensional level (geometric, environmental, etc.) as opposed to individual variables – all of which are captured through the SEM approach.

An additional area of interest not explored in the studies mentioned above is the comparison and evaluation of safety at both signalized and unsignalized intersections. Comparison of how certain factors (such as lane width, average annual daily traffic (AADT), median width, design speed and the level of actuation/control) have varying effects on safety at both intersection types can provide additional insights for future research and design.

This research separates itself from the aforementioned studies by imploring the analytical SEM approach which lends itself to a greater number and combination of exogenous measures. Furthermore, through the use of SEM formulation endogenous variables are grouped into different dimensions and their complex interactions are formulated. This confirmatory (as opposed to exploratory) approach requires the modeller to postulate the links between variables based on hypotheses and previous empirical results (Golob and Meurs, 1986; Golob, 2001). Once these links have been postulated, the SEM approach establishes causal directional relations between variables and then the model is either accepted or rejected based on its validity (Golob, 2001). When a model is accepted the propensity for safety is captured though a latent scale and index which are related to the observable variables through the SEM formulation. The structure of the model and the SPI itself allows for three major contributions: identification of variables that have a significant effect on safety at both signalized and unsignalized intersections; presenting the manner in which these variables and their effects vary with intersection type; and assessment of the relative importance of different determinants (Hamdar and Schorr, 2013).

In contrast to explicitly simulating drivers' behaviours (Hamdar et al., 2008; Paleti et al., 2010), this paper will feature an empirical data-driven approach (Hamdar and Schorr, 2013) through the HSIS data system. Although the SEM approach has been used to examine safety under different interrupted and uninterrupted flow conditions (Hamdar and Schorr, 2013) as well as at international unsignalized intersections alone (Lee et al., 2008), a model that compares different types of intersections will demonstrate the benefits of the approach itself and provide a new perspective on intersection safety. Comparing results with previous studies will lead to additional insights into potential effects of certain variables – though it is important to keep in mind that the research is fundamentally different and conducted in totally different geographic locations.

The formulation of the structural model given the data availability and the corresponding limitations can be found in Section 3. Section 4 provides the results of the factor analysis and the numerical simulation as well as the model statistics. These numerical results are discussed and the obtained models (signalized versus unsignalized) are compared in Section 5. The concluding remarks and the future research needs are provided in Section 6.

3. Statistical model

3.1. Available data and limitations

Data available through the HSIS for the California transportation network was utilized in this study. In order to build a comprehensive and inclusive data set, all collisions (regardless of collision type or severity) occurring at all intersections (regardless of the level of control) between 2006 and 2010 were considered for analysis. HSIS data was provided on a yearly basis in three separate data sets; the Accident File (containing the variables for lighting, precipitation, collision severity, total injuries, total fatalities and the number of vehicles), the Intersection File (containing the variables for AADT, channelization, intersection legs, control/actuation level and the number of lanes), and the Roadway File (containing the variables for lane width, shoulder width, surface width, median width, divided/undivided and design speed). Also contained in the Accident File was a variable stating where the collision had occurred in relation to an intersection (regardless of the type or level of control). Only collisions occurring in intersections or within 250 feet of an intersection were considered for analysis. Data sets were merged on a year by year basis first by merging the intersection and roadway files through variables for route name and milepost and then by merging this composite file with the accident file through the use of the same variables. For cases where the collision occurred within 250 feet of the intersection, the collision was assigned to the intersection where the absolute value of the difference between the collision milepost and the intersection milepost was at a minimum. Merged data sets for each year were then combined to yield the final set for analysis.

Once the master data set was compiled collisions were classified as having occurred at either an unsignalized or signalized intersection using the traffic control variable from the intersection