Interrupted versus uninterrupted flow: A safety propensity index for driver behavior

Samer H. Hamdar*, Justin Schorr

Department of Civil and Environmental Engineering, Center for Intelligent Systems Research, Traffic and Networks Research Laboratory, The George Washington University, Exploration Hall, 20101 Academic Way, Ashburn, VA 20147, USA

ARTICLE INFO

ABSTRACT

The objective of this paper is to develop a quantitative safety propensity index (SPI) that captures the overall propensity of a given surrounding environment to cause unsafe driving. The study is conducted in two different flow conditions: interrupted and uninterrupted. Using structural modeling techniques, the index can be estimated from observed geometric, weather-related, vehicular, driver-related, and traffic-related characteristics. To illustrate the adopted approach, extensive effort was conducted to "sync" data from different sources including the Virginia Department of Transportation and the FARS/GES crash data libraries. The Virginia Department of Transportation provided traffic data for 10 freeway sections with interrupted flow and 9 highway sections with interrupted flow in the Northern Virginia area, USA. Two different structural equations models were found allowing insights to the safety impact of different surrounding elements/dimensions. The SPI provides (a) a basis for quantifying the effects of the aforementioned characteristics on safety, (b) a basis for comparing the differences between the factors affecting safety in different flow scenarios and (c) ranking the corresponding roadway sections/locations for improved safety performance. The framework and methodology used to develop this index have the potential to support safety policy analysis and decision making.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Roadway infrastructure impacts driving behavior which, in turn, has significant implications when analyzing vehicle to vehicle interactions and assessing macroscopic transportation network performance. Previous studies have separately assessed different characteristics (infrastructure, driver, vehicle, traffic, etc.) effects on safety. No comprehensive model exists that takes into account the combined effect of multiple characteristics' types on safety as well as their effect on one another.

Research conducted in this study will develop a safety propensity index in a framework linking road infrastructure and weather conditions to observed crash and traffic data. This framework will allow for a better understanding of the safety implication of road infrastructure and weather features such as: pavement characteristics, number of lanes, lane width, curb width, curvature, grade, precipitation, visibility and roadway surface friction. Crash data as well as highway infrastructure/traffic characteristics of the collection sites (through the Virginia Department of Transportation (VDOT)) are used for verifying the suggested approaches with

the structural equation method (SEM) technique. The different research findings will be used to develop surrogate safety measures on both interrupted and uninterrupted flow roadways and allow for comparison between the corresponding illustrated models: the commonalities and differences between the factors that influence safety under both scenarios will be presented.

The ultimate goals of the research are to (1) systematically identify the network characteristics that influence safety under different traffic situations; (2) study the response to changes in network geometry as an evolving system with temporal and spatial elements with particular attention to the corresponding safety implications; (3) validate the formulated behavioral traffic models against statistical models estimated using existing national incident data (NHTSA, 2010); (4) develop and compare these models for both interrupted and uninterrupted flow scenarios; and (5) observe the models to gain a deeper understanding of how better transportation system performance can be achieved and strategies can be proposed to improve traffic safety and operations.

2. Conceptual framework and background

Creating a safer driving environment is a main objective for transportation researchers in the United States, and worldwide. On roadways in the United States, specifically, there were over

* Corresponding author. Tel.: +1 202 994 6652; fax: +1 202 994 0127.
E-mail addresses: hamdar@gwu.edu, samerhamdar@gmail.com (S.H. Hamdar), justin11@gwmail.gwu.edu (J. Schorr).

33,000 fatalities in 2009 and although this is a decrease in both total fatalities and fatalities per vehicle mile traveled (NHTSA, 2010), researchers are still seeking improvements to minimize this number. Before this marked decrease in fatalities (and in both fatal and total collisions) from 2008 to 2009, both the number of vehicle miles traveled and the number of total collisions had been following an ascending pattern for the previous 10 years (NHTSA, 2010). If vehicle miles traveled is considered a surrogate measure for congestion, and total collisions a measure for safety, there is a need to examine the different possible factors leading to the aforementioned patterns thus enabling a less congested transportation system and creating a safer driving experience. There are a variety of ways congestion can begin, even with freeway demand not at critical levels, including (but not limited to): shockwaves generated on uphill slopes due to trucks/trailers climbing at slower speeds; vision reduction for a few drivers by a rising sun in a straight highway segment; weaving near a freeway ramp; or exaggerated braking near an unconventional road design so that shockwaves propagate backwards and slow down or completely stop traffic down the road (Treiber et al., 2000). These examples illustrate complex interactions between roadway geometry, drivers' characteristics and environmental conditions that impact transportation efficiency and safety.

Previously conducted research in this area focuses on only one or two dimensions (such as geometry and traffic characteristics) (Karlaftis and Golias, 2002; Li et al., 1994) and mainly uses accident rates at the metric for evaluating safety (Joshua and Garber, 1990; Jones and Whitfield, 1988; Karlaftis and Golias, 2002; Li et al., 1994). In this paper, focusing on an empirical data-driven approach rather than on a simulation behavioral approach (Hamdar et al., 2008; Talebpour et al., 2012), through the FARS/GES crash libraries, different metrics (number of injuries, number of fatalities, etc.) are used to assess the safety of the roadway and multiple dimensions are considered in the analysis. Results from previous research can be used to develop an understanding of the effects that certain variables may have, but provide a weak basis for comparison of results as the type of analysis are fundamentally different. One such study (Lee et al., 2008) utilized a similar structural equation approach, but the dimensions selected (and the variables selected within those dimensions), the location of the study (Korean highways) and the flow conditions analyzed (only uninterrupted) all differed from the analysis conducted in this study.

Creating a safety propensity index (SPI) based on roadway geometry involves capturing the complex relationships outlined earlier. The interrelationships between latent and endogenous quantities on one hand and on the other, measured (observable) variables characterizing the environmental conditions, geometric features of the roadway, traffic situations, socio-demographics of the drivers, as well as instances of certain driving behaviors and collision scenarios. Fig. 1 presents an initial conceptual framework illustrating the main types of factors that enter into the formulations of the SPI, as well as its dependence on a set of complex relationships. This further illustrated in Fig. 2 through specific dimensions and example variables of measures that capture the dimension.

The complexities of the interrelationships followed by the dimensions and driving patterns mentioned above can be formulated us the structural equation modeling approach. Structural equation modeling (SEM) is a cause and effect approach to analyzing data where relationships between variables are postulated by the modeler based on theories and previous empirical results (Golob and Meurs, 1986a,b). The approach is "confirmatory rather than exploratory" (Golob, 2001), as the system of unidirectional effects of one variable on another is being constructed and then either accepted or rejected based on its validity (Golob, 2001). This approach is becoming increasingly popular in travel behavior research as user-friendly software becomes increasingly powerful and widely available (Golob, 2001).

While safety can be derived from a number of different metrics, the manner and degree in which it is affected by the aforementioned observable variables if difficult to quantify. This propensity for safety is captured though a latent scale and index, and related to observable variables through the SEM formulation. This index allows for the identification of the major contributing factors for a certain flow scenario, the manner in which those factors vary with flow scenario, and assessment of the relative importance of different determinants.

3. Statistical model

The data used for analysis was provided in multiple databases by the Virginia Department of Transportation (VDOT). The three databases that were combined and edited contained data on collisions, traffic and pavement characteristics respectively. Both interrupted and uninterrupted flow conditions were considered. In the uninterrupted flow situation, nine highway segments were chosen for analysis and are displayed in Table 1.

In the interrupted flow situation, ten roadway segments – all of which are components of a network of signalized intersections – were chosen for analysis and are displayed in Table 2.

3.1. Available data and additional limitations

The major limitation of this study was the availability of data in the state of Virginia. Initial conceptual framework had to be adjusted as a variety of necessary data was either not available or incomplete. In response to these major obstacles the following was implemented:

1. A friction variable was developed. In order to compute the friction values for the specific segments of roadway at the time of each individual collision, the weather conditions at time of collision, the pavement type, the vehicle speed and the time since last roadway rehabilitation were employed in unison. Cross referencing these values with a table of accepted friction values (Baker and Fricke, 1990; Table 3)

Table 1
Uninterrupted flow data collection locations.

Segment number	Highway name	Mile markers
1	I-66 West	12–13
2	I-66 West	53–54
3	I-66 West	70–71
4	I-495 South	1.5–2.5
5	I-495 South	6–7
6	I-495 South	12–13
7	I-81 North	291–292
8	I-395 North	3–4
9	I-95 South	152–153

Table 2
Interrupted flow data collection locations.

Segment number	Main/through	Minor/cross
1	State Route 7	State Route 193
2	State Route 7	US 29
3	State Route 123	State Route243
4	US 50	Old Ox Road
5	State Route 28	Liberia Drive
6	US 29	US 15
7	US 17	State Route 28
8	State Route 123	Burke Lake Road
9	State Route 234	Prince William Parkway
10	State Route 123	State Route 309
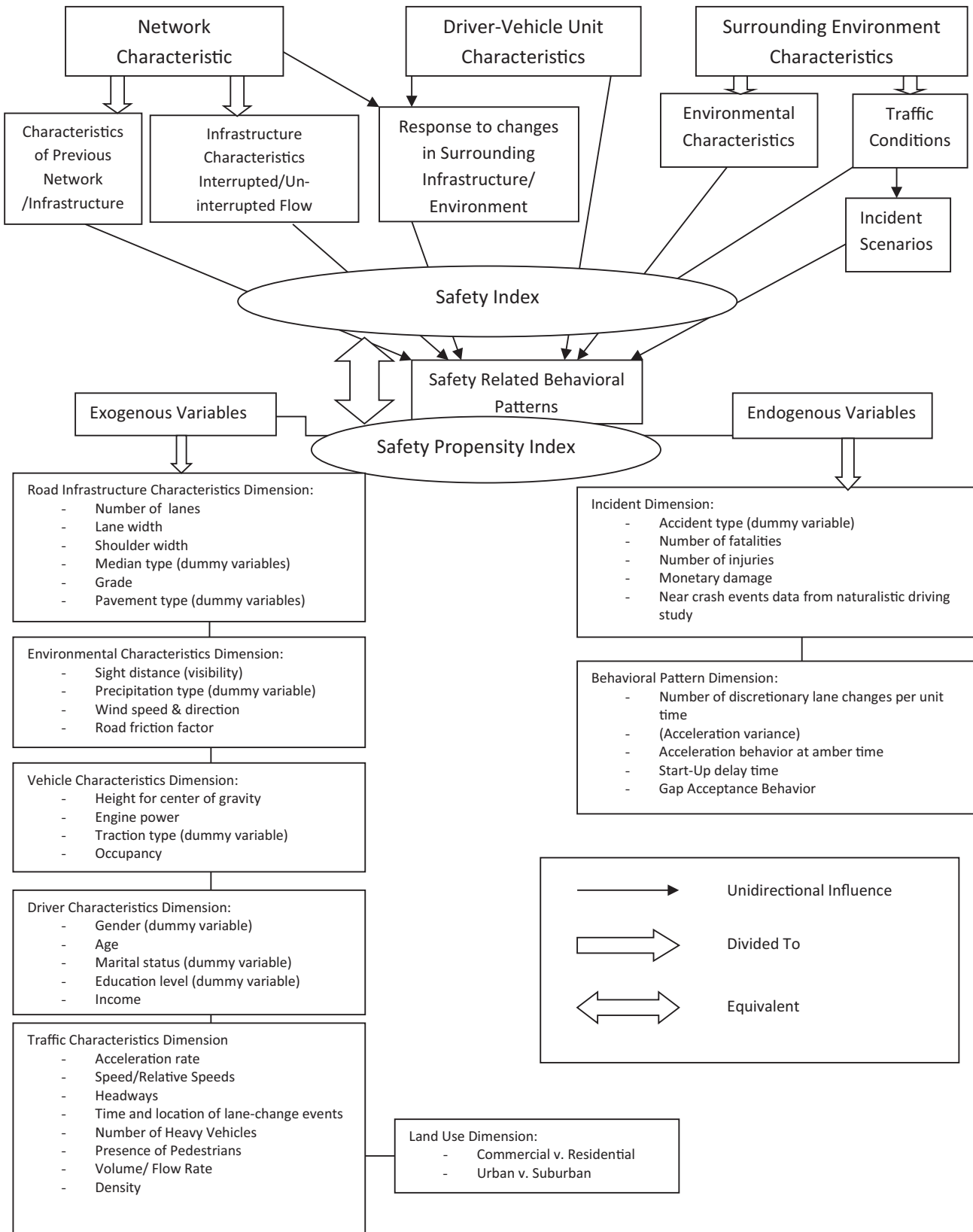
Figs. 1 and 2. (1) Safety propensity index framework – upper section. (2) Initial basic dimensions and patterns considered for safety index – lower section.

provided the friction value for each collision. For the interrupted flow scenarios the friction variable was not used for analysis because roadway rehabilitation data was not available.

2. An aggressive maneuver variable was developed. The data provided by VDOT was comprised of police reported data for collisions. In filing out their reports, officers were required to assess the action taken by the driver in the collision. Of the 43
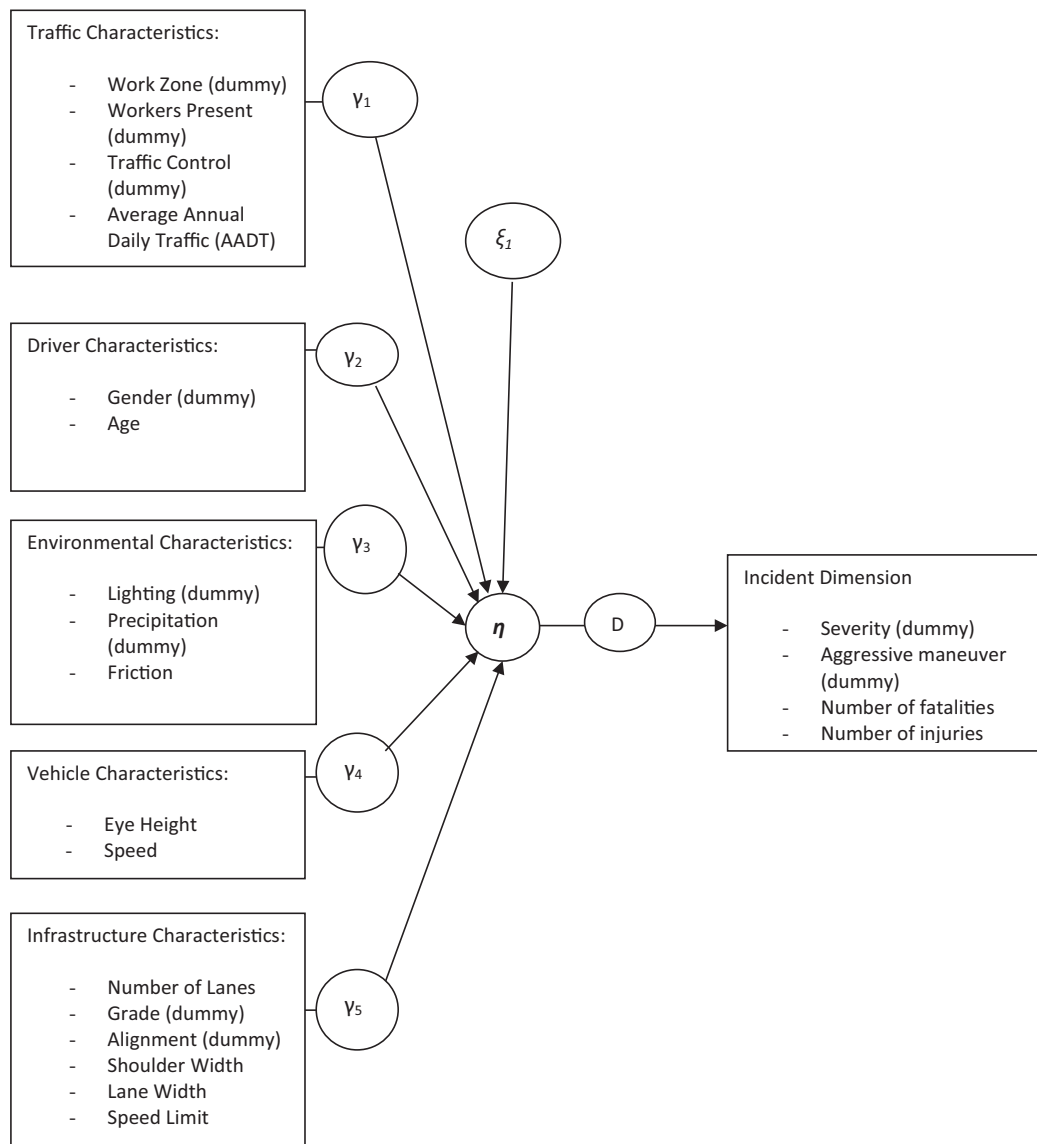
Fig. 3. Simplistic representation of the Initial theoretical model structure based on available data.

possible selections, the following 10 were considered aggressive maneuvers: exceeding speed limit; overtaking on a hill; overtaking on a curve; improper passing of a school bus; cutting in; hit and run; eluding police; improper passing; and improper or unsafe lane changing. All others were considered non-aggressive driving maneuvers or did not appear in the data set.

3. Each collision was used as only one data point regardless of the number of vehicles involved. If collisions were used multiple times it would influence the exposure of the geometric, environmental and driver variables associated with them. For instance, it would not be prudent to include all vehicles in a 5 vehicle, low-speed, rear-end collision for fear that it would influence the significance of the results. Considering the size of the data set as well as the low-speed nature of many of the collisions, the vehicle with the highest police reported speed was used for analysis (driver and vehicle characteristics in addition to the speed being utilized for the friction calculation). Finally, injuries and fatalities were recorded as a count for the collision as a whole, rather than for each individual vehicle as they were originally reported.

4. For the interrupted flow scenarios, collisions occurring in *both* directions were considered for all segments. The nature of

the data did not allow for distinction between accidents that occurred in one direction or another. While certainly not ideal, the effects of this ambiguity were mitigated by the symmetrical nature of the segments used for analysis. Inspection of the provided data yielded no differences in geometric characteristics between opposite sides of the roadway for any of the segments analyzed. Traffic data was similarly ambiguous in nature as it too provided no directional distinction.

5. All data points (collisions) that were missing one or more of the variables considered for analysis were omitted.

6. Only collisions occurring after January 1, 2004 were considered so that results were current as roadway geometry is frequently changing.

Based on the data available, the final analysis was conducted using the following variables: surface width; shoulder width; number of lanes; alignment; weather; surface condition; lighting, severity; fatal count; injury count; work zone; workers present; speed limit; vehicle type; vehicle speed; driver age; driver sex; average annual daily traffic; friction; and divided/undivided (Fig. 3).

Table 3
Utilized friction values.

Years since rehab	Asphalt, dry		Asphalt, wet	
	Under 30 mph	Over 30 mph	Under 30 mph	Over 30 mph
2 and under	0.9	0.9	0.65	0.6
2–5	0.7	0.65	0.55	0.5
Over 5	0.65	0.55	0.55	0.5

Years since rehab	Cement, dry		Cement, wet	
	Under 30 mph	Over 30 mph	Under 30 mph	Over 30 mph
2 and under	0.9	0.85	0.65	0.6
2–5	0.7	0.65	0.6	0.55
Over 5	0.65	0.55	0.55	0.5

Years since rehab	Ice		Snow	
	Under 30 mph	Over 30 mph	Under 30 mph	Over 30 mph
2 and under	0.15	0.15	0.45	0.45
2 to 5	0.15	0.15	0.45	0.45
Over 5	0.15	0.15	0.45	0.45

4. Measurement models

Measurements models are specified in two sets of equations. The first set (the exogenous measurement model) is represented as follows:

$$\mathbf{X} = \boldsymbol{\Lambda}\mathbf{X}(\gamma) + \boldsymbol{\omega} \tag{1}$$

$\mathbf{X}$, vector of exogenous variables; $\boldsymbol{\Lambda}\mathbf{X}$, matrix of structural coefficients for latent exogenous variables to their observed indicator variables; γ, vector of latent exogenous constructs; $\gamma 1$, safety propensity for aggressive driving associated with "Traffic Characteristics dimension"; $\gamma 2$, safety propensity for aggressive driving associated with "Driver Characteristics dimensions"; $\gamma 3$, safety propensity for aggressive driving associated with "Environmental Characteristics dimensions"; $\gamma 4$, safety propensity for aggressive driving associated with "Vehicle Characteristics dimensions"; $\gamma 5$, safety propensity for aggressive driving associated with "Infrastructure Characteristics"; $\boldsymbol{\omega}$, vector of measurement error terms for observed variables.

The latent exogenous variables are a direct reflection of the dimensions initially considered in the framework of the study. The observed exogenous variables are described in the following Table 4A, including how they are measured, and the associated variable name by which they will be designated in the next section. These variables were selected based on availability within the data set as well as through the examination of previously conducted studies (Paleti et al., 2010; Hamdar et al., 2008; Lee et al., 2008).

The second set (endogenous measurement model) of equations are summarized in Eq. (2):

$$\mathbf{Y} = \boldsymbol{\Lambda}\mathrm{Y}(\boldsymbol{\eta}) + \boldsymbol{\tau} \tag{2}$$

where $\mathbf{Y}$, vector of observed endogenous variables; $\boldsymbol{\Lambda}\mathbf{Y}$, matrix of structural coefficients for latent endogenous variables to their observed indicator variables; $\boldsymbol{\eta}$, vector of latent endogenous variable; $\eta 1$, safety propensity index; $\boldsymbol{\tau}$, vector of measurement error terms for observed endogenous variables. The observed endogenous variables are described in Table 4B.

5. Structural model

A structural model relating the endogenous latent variable $\eta 1$ to the exogenous latent variables $\gamma 1$, $\gamma 2$, and $\gamma 3$ can be expressed as (Golob, 1988):

$$\boldsymbol{\eta} = \Delta\gamma + \xi \tag{3}$$

Table 4A
Exogenous variables description.

Exogenous variable	Description
X_1 (Work Zone)	Dummy variable corresponding to the presence of a work zone: Work zone not present, $X_1 = 0$ Work zone present, $X_1 = 1$
X_2 (workers present)	Dummy variable corresponding to the presence of workers: Workers not present, $X_2 = 0$ Workers present, $X_2 = 1$
X_3 (control)	Dummy variable corresponding to the type of highway control: Lane markings only, $X_3 = 0$ Non-standard highway control, $X_3 = 1$
X_4 (AADT)	Average annual daily traffic
X_5 (gender)	Dummy variable corresponding to the gender of the driver: Male, $X_4 = 0$ Female, $X_4 = 1$
X_6 (age)	Age of the driver
X_7 (lighting)	Dummy variable corresponding to the lighting at time of collision: Natural light, $X_6 = 0$ Street lamps, $X_6 = 1$
X_8 (precipitation)	Dummy variable corresponding to the precipitation at time of collision: No precipitation, $X_7 = 0$ Precipitation, $X_7 = 1$
X_9 (Friction)	Friction values for the specific segments of roadway at the time of each individual collision. The weather conditions at time of collision, the pavement type, the vehicle speed and the time since last roadway rehabilitation were employed in unison. Cross-referencing these values with the table of accepted friction values (Traffic Accident Investigation Manual, Baker, 1990) provided the friction value for each collision. Detailed description found in Table 3
X_{10} (eye height)	Driver eye height. Using average values based on the type of vehicle, a height of 45 inches was assigned to any passenger car, 60 inches to any van, pick-up or sport utility vehicle and 90 inches to any tractor trailer, motor home or bus
X_{11} (speed)	Vehicle speed at time of collision
X_{12} (Lane#)	Number of lanes per major approach
X_{13} (grade)	Dummy variable corresponding to the grade of the roadway: Flat, $X_{12} = 0$ Graded, $X_{12} = 1$
X_{14} (alignment)	Dummy variable corresponding to the alignment of the roadway: Straight, $X_{13} = 0$ Curved, $X_{13} = 1$
X_{15} (shoulder width)	Shoulder width
X_{16} (lane width)	Lane width
X_{17} (speed limit)	Speed limit

Table 4B
Endogenous variables description.

Observed endogenous variables	Description
Y_1 (injury count)	Total number of injuries in the collision
Y_2 (fatal count)	Total number of fatalities in the collision
Y_3 (severity)	Dummy variable corresponding to the severity of the collision: Non-severe, $Y_3 = 0$ Severe, $Y_3 = 1$
Y_4 (aggressive maneuver)	Dummy variable corresponding to the aggressive nature of the maneuver made by the driver in the collision: Non-aggressive, $Y_4 = 0$ Aggressive, $Y_4 = 1$

where η, vector of latent endogenous variable; $\eta 1$, safety propensity index; Δ, matrix of structural coefficients for exogenous latent variables to endogenous latent variables. γ, vector of latent exogenous constructs; $\gamma_1, \ldots, \gamma_4$ are as previously defined. ξ, vector of measurement error terms for latent endogenous variables.

$$[\eta_1] = [\,\delta_{11} \quad \delta_{12} \quad \delta_{13} \quad \delta_{14} \quad \delta_{15}\,] * \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} + [\xi_1] \tag{4}$$

Similarly the measurement equations can be expressed as follows (Golob, 1988):

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \\ X_7 \\ X_8 \\ X_9 \\ X_{10} \\ X_{11} \\ X_{12} \\ X_{13} \\ X_{14} \\ X_{15} \\ X_{16} \\ X_{17} \end{bmatrix} = \begin{bmatrix} \Omega_{11} & 0 & 0 & 0 & 0 \\ \Omega_{21} & 0 & 0 & 0 & 0 \\ \Omega_{31} & 0 & 0 & 0 & 0 \\ \Omega_{41} & 0 & 0 & 0 & 0 \\ 0 & \Omega_{52} & 0 & 0 & 0 \\ 0 & \Omega_{62} & 0 & 0 & 0 \\ 0 & 0 & \Omega_{73} & 0 & 0 \\ 0 & 0 & \Omega_{83} & 0 & 0 \\ 0 & 0 & \Omega_{93} & 0 & 0 \\ 0 & 0 & 0 & \Omega_{10,4} & 0 \\ 0 & 0 & 0 & \Omega_{11,4} & 0 \\ 0 & 0 & 0 & 0 & \Omega_{12,5} \\ 0 & 0 & 0 & 0 & \Omega_{13,5} \\ 0 & 0 & 0 & 0 & \Omega_{14,5} \\ 0 & 0 & 0 & 0 & \Omega_{15,5} \\ 0 & 0 & 0 & 0 & \Omega_{16,5} \\ 0 & 0 & 0 & 0 & \Omega_{17,5} \end{bmatrix} * \begin{bmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \\ \gamma_5 \end{bmatrix} + \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \\ \omega_7 \\ \omega_8 \\ \omega_9 \\ \omega_{10} \\ \omega_{11} \\ \omega_{12} \\ \omega_{13} \\ \omega_{14} \\ \omega_{15} \\ \omega_{16} \\ \omega_{17} \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \\ Y_4 \\ Y_5 \end{bmatrix} = \begin{bmatrix} \lambda_{11} \\ \lambda_{12} \\ \lambda_{13} \\ \lambda_{14} \\ \lambda_{15} \end{bmatrix} * [\eta_1] + \begin{bmatrix} \tau_1 \\ \tau_2 \\ \tau_3 \\ \tau_4 \\ \tau_5 \end{bmatrix} \tag{6}$$

In addition to the three structural matrices $\Lambda X, \Lambda Y$, and Δ, the following four variance/covariance matrices need to be specified to determine a general structural equation model:

1) a VC-matrix of latent exogenous variables (Φ)
2) a VC-matrix of error terms associated with model implied structural equations (Ψ)
3) a VC-matrix of measurement errors or observed exogenous variables $(\theta\omega)$
4) a VC-matrix of measurement error terms associated with the observed endogenous variables $(\theta\tau)$

6. Application to highway segments in the Northern Virginia area

6.1. Structural equation model (SEM) development and results

Factor analysis was performed for both interrupted and uninterrupted flow scenarios using the Statistical Analysis System (SAS) software. This type of analysis establishes relationships based on a mathematical function $f(W,Z)$ connecting a variable X with the set of variables W and Z. The measurable values of Y are known, however the type of function $f(\ )$ that should be used and the variables to be included in this function are unknown. Accordingly, we assume that a set of Y variables are related to a number of functions that operate linearly:

$$X_n = \alpha_{n1}F_1 + \alpha_{n2}F_2 + \cdots + \alpha_{nm}F_m \tag{7}$$

where X is a variable with known data, α_{nj} is a constant that represents the loading, and F_j is a function $f_j(\ )$ of some unknown variables where $j = 1, \ldots, m$.

The output derived from this analysis is useful in the following manner:

1. *Un-rotated matrix*: deals solely with uncorrelated patterns. Each pattern could potentially involve all (or nearly all) the variables, and therefore may lead to high loadings for several factor patterns.
2. *Pre-rotated matrix*: deals solely with correlated patterns.
3. *Rotated factor matrix*: the factor matrix covers both correlated and uncorrelated patterns. Using this particular case, patterns can be hypothesized and uncovered without including all (or nearly all) the variables.

The following table (Table 5) represents the factor analysis for the interrupted flow scenario. Values on the order of 0.1 were initially selected for analysis.

Determinations were made based on the factor scores of the variables (as a starting point, scores on the order of 0.1 were viewed as significant (Hamdar et al., 2008)) as well as the relevance of the variable to the dimension in question (for instance in Table 5 above the variable "lighting" could be included in Factor 4 based on its factor score, but physically it is obvious that lighting does not belong in the Traffic Dimension). The first determination was that the number of factors needed to be reduced from 5 to 4, seeing as how the driver age and sex variables needed to be dropped. Based on the pre-rotated and rotated matrices, several variables could be excluded based on their low factor scores and the dimensions could be redefined; leaving the final four dimensions as L1 (Factor 2) – infrastructure characteristics (number of lanes, surface width and lane width), L2 (Factor 3) – environmental characteristics (precipitation and lighting), L3 (Factor 1) – work zone characteristics (work zone and workers present) and L4 (Factor 4) – traffic characteristics (AADT and speed).

Several structures were then tested based on these new dimensions. due to the nature of the work zone characteristics (both are dummy variables) the dimension was dropped and seeing how surface width is simply equal to (lane width)*(number of lanes)+(shoulder width), surface width was replaced by shoulder width and lane width leaving the final structural model (computed using the LISREL software) displayed below in Fig. 4:

The results summarizing the interrupted flow model are presented below:

Table 5
Interrupted flow rotated factor matrix.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Shoulder width	0.01565	−0.09067	−0.03104	0.57401	−0.35999
Speed limit	−0.02033	0.47742	0.01813	0.15455	−0.36514
Number of lanes	0.01251	**0.96318**	0.01143	−0.08192	−0.21549
Surface width	−0.00091	**0.78369**	−0.02015	0.11129	0.37431
Lane width	−0.01624	−0.01502	−0.04276	0.2586	**0.79817**
Curve	−0.08246	0.0945	−0.01248	0.03019	−0.23815
Grade	0.08998	0.02994	−0.05587	0.50067	−0.10719
Work zone	**0.97761**	0.00201	−0.00162	−0.00556	0.0474
Workers present	**0.97372**	0.00542	0.01148	0.02562	0.01832
Control	0.00702	0.0832	−0.06712	−0.06838	0.01774
Precipitation	0.00826	0.00535	**0.93494**	0.07213	0.00987
Lighting	−0.06681	0.0505	**0.05831**	0.45311	−0.02282
Driver sex	−0.02611	−0.10793	0.05175	−0.27969	−0.01666
Friction	−0.00186	0.01006	−0.9133	−0.09865	0.00059
Speed	−0.02272	−0.08451	0.08549	**0.71535**	0.10989
Driver age	−0.01941	−0.07939	−0.08233	−0.14511	−0.08601
AADT	−0.009	0.04986	−0.03453	**0.2103**	−0.619

Endogenous measurement model
- Severe = 0.16*INDEX, Errorvar = 0.11, R^2 = 0.5
- Injtotal = 1.01*INDEX, Errorvar = 0.32, R^2 = 0.00061
- Fataltot = 0.01*INDEX, Errorvar = 0.019, R^2 = 0.93
- Agressive = −0.0045*INDEX, Errorvar = 0.1, R^2 = 0.029

Exogenous measurement model
- Shoulder = −10.71*L1, Errorvar = −103.86, R^2 = 15.05
- Numlanes = 0.01*L1, Errorvar = 0.032, R^2 = 0.00078
- Nlnwidth = 0.01*L1, Errorvar = 0.17, R^2 = 0.0016
- Precip = 0.05*L2, Errorvar = 0.13, R^2 = 0.024
- Light = 0.27*L2, Errorvar = 0.13, R^2 = 0.28
- Speed = 5.44*L3, Errorvar = 221.91, R^2 = 0.033
- AADT = 1.04*L3, Errorvar = 129.93, R^2 = 0.0022

Structural model
- INDEX = 3.26*L1 + (−15.62)*L2 + 16.32*L3, Errorvar = 1.69, R^2 = 0.18

Covariance terms
- Error covariance for AADT and speed = −4.19
- Error covariance for speed and light = 0.02
- Error covariance for speed and Precip = 0.38
- Error covariance for speed and Nlnwidth = 0.65
- Error covariance for speed and Numlanes = −0.12

t-Values
- L1/shoulder width: *−7.96*
- L1/number of lanes: *5.28*
- L1/lane width: *3.81*
- L2/precipitation: *−3.08*
- L2/lighting: *−3.78*
- L3/speed: *−0.58*
- L3/AADT: *−0.55*
- L1/index: *−0.82*
- L2/index: *−0.66*
- L3/index: 0.35
- Index/severity: *5.42*
- Index/injury total: *4.51*
- Index/aggressiveness: *−0.79*
- Index/fatal total: 0.00

For models with large sample sizes (such as the sample size used in this section: N = 1097), Chi-squared tests often encounter problems. For this reason, the goodness of fit was assessed based on the root mean square error of approximation (RMSEA) (Golob, 2001). For the Interrupted Flow Model, the RMSEA was 0.057 and the 90% confidence interval was 0.048;0.066 – both are approximately on the order of 0.05 indicating that the model is statistically significant. The *t*-values indicate that we may be more confident in some paths than in others. For an alpha of 0.05, all *t*-value less than (−1.96) or greater than (+1.96) can be viewed as significant.

A similar analysis was carried out for the uninterrupted flow scenario (using the SAS software) and the rotated factor chart is displayed in Table 6.

Based on this analysis, the Driver Characteristics dimension was once again dropped and several variables were excluded, leaving the final dimensions as follows: L1 (Factor 1) – infrastructure characteristics (shoulder width and lane count), L2 (Factor 4) – environmental characteristics (friction and lighting), L3 (Factor 2) – work zone characteristics (work zone and workers present) and L4 (Factor 5) – traffic characteristics (AADT and speed limit).

Several structures were once again tested based on the new dimensions. As was the case with interrupted flow, the Work Zone Characteristics dimension was dropped due to the nature of the variables. Additionally, the variable for grade was added to the Infrastructure dimension and fatality total was dropped as an output variable. The final model (computed using the LISREL software) is displayed below in Fig. 5:

The results summarizing the uninterrupted flow model are presented below:

Endogenous measurement model
Severe = 0.34*INDEX, Errorvar = 0.10, R^2 = 0.69
Injtotal = 2.30*INDEX, Errorvar = 0.85, R^2 = 0.65
Agressive = −0.01*INDEX, Errorvar = 0.24, R^2 = 0.0013

Exogenous measurement model
- Shoulder = 0.40*L1, Errorvar = 4.28, R^2 = .0013
- Lanecount = 0.72*L1, Errorvar = −0.07, R^2 = 0.76
- Nlnwidth = −0.13*L1, Errorvar = 0.19, R^2 = 0.11
- Friction = −0.13*L2, Errorvar = 0.0012, R^2 = 0.27
- Light = 0.05*L2, Errorvar = 0.19, R^2 = 0.003
- Speedlimit = 2.36*L3, Errorvar = 24.98, R^2 = 0.18
- AADT = −21.07*L3, Errorvar = 33.77, R^2 = 0.94

Structural model
- INDEX = 0.02*L1 + (−0.06)*L2 + (−0.07)*L3, Errorvar = 1.30, R^2 = 0.30

Covariance terms
- Error covariance for speed limit and friction = 0.07
- Error covariance for speed limit and shoulder = 2.59
- Error covariance for lane count and shoulder = 0.28
- Error covariance for lane count and grade = 0.11
- Error covariance for shoulder and grade = −0.09

t-Values
- L1/shoulder width: *5.05*
- L1/lane count: *27.15*
- L1/grade: *−8.72*
- L2/lighting: *2.47*
- L2/friction: *−3.75*
- L3/speed limit: *13.04*
- L3/AADT: *−31.38*
- L1/index: 0.35
- L2/index: *−1.22*
- L3/index: *−0.90*
- Index/severity: 0.00
- Index/injury total: *3.73*
- Index/aggressiveness: *−0.79*

Once again, a Chi-squared test is not used for this model based on the large sample size (N = 911). The RMSEA of this model is 0.091 with the 90% confidence interval being 0.081;0.10. These values are on the cut-off of statistical significance (RMSEA = 0.1) so we check
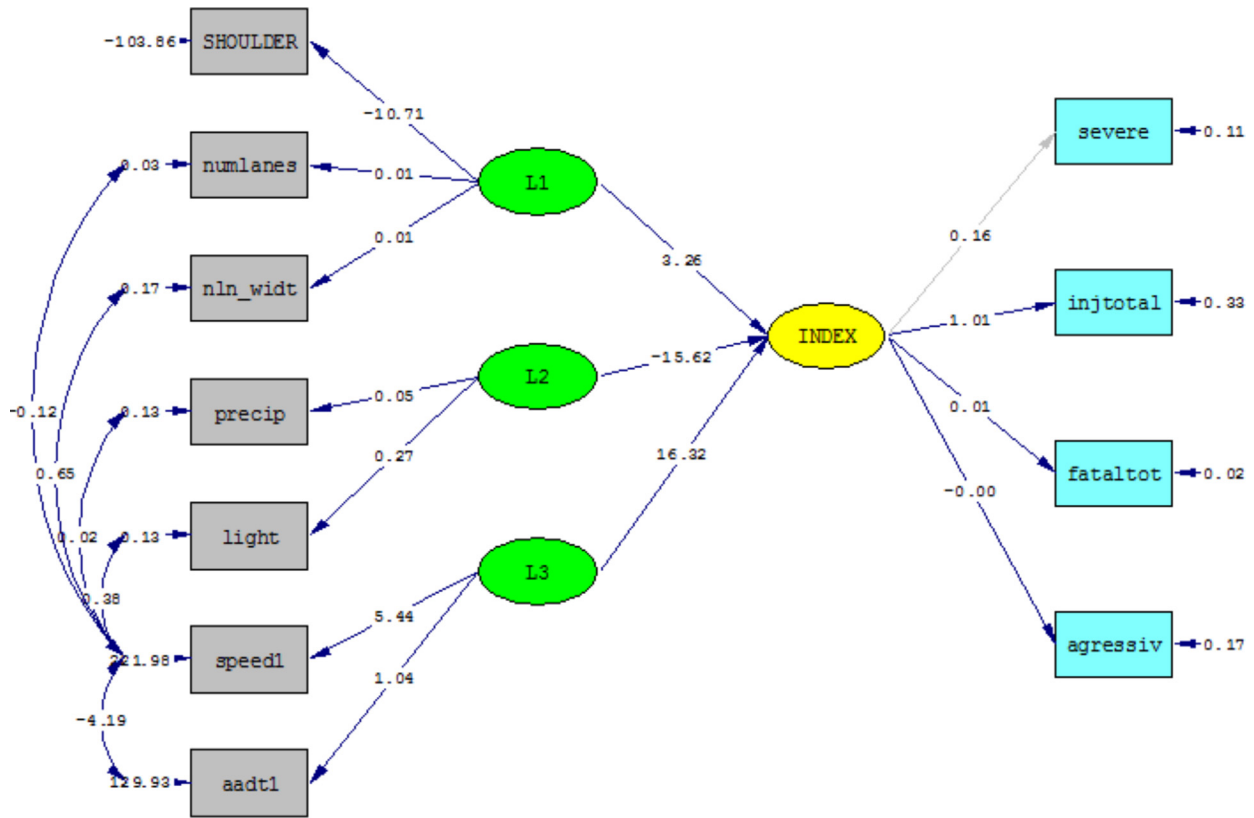
Fig. 4. Structural model for interrupted flow.

other criteria such as the standardized root mean square residual which is acceptable value of 0.072 (values less than 0.08 are considered good fit (Hu and Bentler, 1998)) and the goodness of fit index (GFI) which is an acceptable value of 0.95 (values greater than 0.9 are considered good fit). Additionally, the *t*-values may be interpreted in the same manner as for the interrupted flow model.

It should be mentioned that the exclusion of certain dimensions in both models has two main biases associated with it: over or under estimation of some dimensions' impact on the safety index and thus on the endogenous observed variables' values, as well as over or underestimation of the covariance matrix between the variables kept in the final model. In other words, the entire model structure may be altered due to the elimination of some significant dimensions. However, given the approach and the corresponding

analytical methodology adopted in this paper, several checks and procedural trials are being built to have a converging consistent model with significant variables/dimensions. Additionally, in an effort to check the internal consistency of the model and to confirm the results of the factor analysis, Cronbach's alpha was utilized (Cortina, 1993). Alpha values were calculated both for each individual model as a whole and for each dimension within both models and the results are as follows:
Interrupted flow
• Data set: 0.28• Dimension L1: 0.24• Dimension L2: 0.28• Dimension L3: 0.36
Uninterrupted flow
• Data set: 0.28• Dimension L1: 0.15• Dimension L2: 0.31• Dimension L3: 0.19

All values obtained were below the cutoff for internal consistency. This inconsistency can be attributed mainly to the varying scales within the data sets as well as within the individual

Table 6
Uninterrupted flow rotated factor matrix.

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Shoulder width	**0.10972**	−0.03671	−0.12368	0.4903	−0.26968
Lane count	**0.95146**	−0.01369	−0.16041	0.05106	−0.00611
Speed limit	−0.30856	−0.08335	−0.07204	0.48782	**0.1389**
Lane width	0.08959	−0.02462	−0.04389	−0.05957	0.44373
Curve	0.07791	−0.03357	−0.2183	−0.00251	0.04427
Grade	−0.24923	−0.01219	−0.07826	0.0098	0.35099
Work zone	−0.00128	**0.98018**	−0.01103	0.00894	−0.0044
Workers present	−0.00609	**0.97715**	−0.01275	−0.01808	−0.02034
Control	−0.03593	0.07974	0.06926	0.44318	0.27945
Precipitation	−0.08538	−0.02595	−0.34379	−0.04552	0.03189
Lighting	0.05064	0.08706	−0.27572	**0.16591**	0.01584
Driver sex	0.01293	−0.02674	−0.00061	0.05163	0.05308
Friction	0.31954	0.00999	0.55083	**0.11013**	−0.09418
Speed	0.02344	0.03543	−0.51005	0.09329	−0.10224
Driver age	−0.07041	−0.01735	0.185182	−0.09577	−0.03564
AADT	0.94531	−0.01521	0.00704	−0.04293	**0.15056**

Bold face values demonstrate which factor variables are considered under according to the factor analysis.
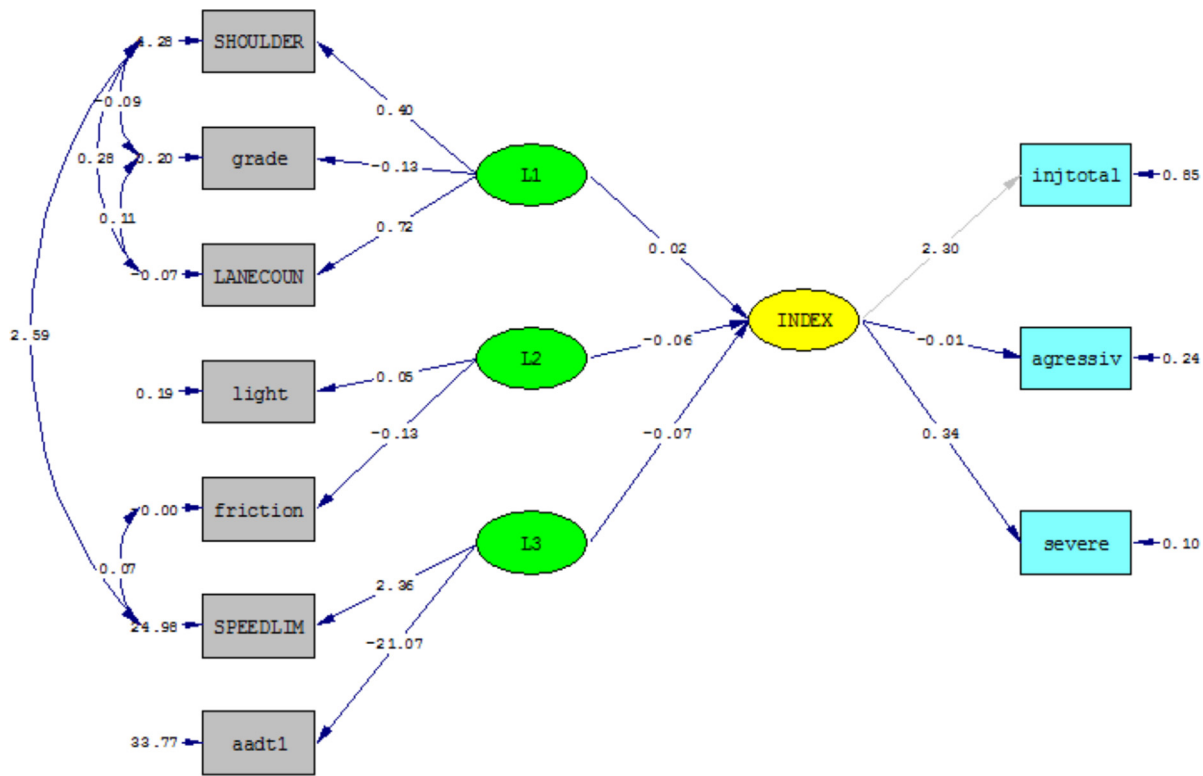
Fig. 5. Structural model for uninterrupted flow.

dimensions (for example speed limit ranges from 55 to 65, grade from 0 to 1 and lane count from 2 to 4). Alpha is typically calculated for data sets that have either dichotomous variables or variables that have the same range (i.e. a questionnaire that utilizes a Likert Scale), and even then all alpha provides is an average degree of "interrelatedness" given that no negative covariances exist (Sijtsma, 2009). Given the low alpha values alternative methods were used (including the use of alternative factor and model structures, the checking for the convergence rates and the observation of error variances and R^2 values) and the models are seen consistent. Furthermore it is important to note that this is an analysis model (not a prediction model) intended provide insight into the relationships between certain variables and roadway safety. Values are specific to the area of analysis and cannot be applied universally across all networks.

7. Interpretation and discussion of results

7.1. Interrupted flow

Table 7 displays the normalized geometric contribution to the safety index for each interrupted flow segment. They are listed in order from the safest (related to the available dimension/surrogate measures) to the least safe segment of roadway based on their geometric characteristics that were included in the model. (*Note*: the large disparity between the first seven segments and the last three is due to the fact that the last three segments all have no shoulder)

Additionally, Fig. 6 displays Google[TM] earth images of segments 4 and 5, the safest and least safe interrupted flow segments respectively.

The model for interrupted flow is characterized by a negative influence on the safety propensity index from the vehicle and environmental characteristics and a positive influence from the geometric characteristics. The absolute values of each coefficient

demonstrates that the largest value is associated with the traffic characteristic (16.32), followed by environmental (15.63) and infrastructure (3.26). As expected, speed has the largest influence on safety-increasing speed leads to decreasingly safe roadways in interrupted flow scenarios. Additionally, higher AADT (indicative of more highly traveled roadways) leads to a decrease in safety as well.

Consistent with expectations is the significant effect that geometric characteristics have on safety in these scenarios. The largest influence in terms of geometry is the high negative coefficient between shoulder width (−10.71) and the infrastructure variable L1. This negative coefficient means a lower value of the safety index, which may be indicative of a safer driving experience. Given the high variability of speeds that is inherent to interrupted flow, one would expect that the added buffer provided by a wide shoulder

Table 7
Interrupted flow – normalized geometric contribution to safety index.

Segment number	Main/through	Minor/cross	Geometric contribution
4	US 50	Old Ox Road	−17.09
6	US 29	US 15	−17.09
9	State Route 234	Prince William Parkway	−17.09
7	US 17	State Route 28	−17.09
1	State Route 7	State Route 193	−12.81
8	State Route 123	Burke Lake Road	−9.60
10	State Route 123	State Route 309	−6.39
2	State Route 7	US 29	0.03
3	State Route 123	State Route 243	0.03
5	State Route 28	Liberia Drive	0.03

Fig. 6. Safest (left) and least safe (right) interrupted flow segments.

would help make the roadway safer. The dimension is dominated by the shoulder width variable, leaving very small contributions from lane count and lane width (both 0.01). Based on previous research, the large influence of shoulder width can be attributed to higher variability in the values for shoulder width than those of lane width and number of lanes throughout the data set. This finding is consistent with that of a 2009 publication by the Federal Highway Administration which found that the geometric influence on accident rates varied based on different lane, shoulder and surface width combinations. Additionally, this study found that for the values of lane width and number of lanes similar to those present in this data set, increases in shoulder width led to lower estimates of the odds of a crash (Gross et al., 2009).

Looking at the environmental characteristic dimension, results may be seen initially counter-intuitive, as lower values of precipitation (0.05) and lighting (0.27) – corresponding to clear and daylight conditions respectively – decrease the safety index of a roadway with interrupted flow. Nevertheless, considering the increased headways, lower speeds and more variations in speeds that are already present due to the nature of interrupted flow, this result may reveal itself to be consistent with expectations. When adverse conditions are present (precipitation and less lighting), already aware drivers may become more cognizant of their surroundings and may, in turn, operate their vehicles in a manner that creates a safer driving environment. Proper warning messages may lead to better awareness is such situations (Rakha et al., 2009).

Observing the endogenous side of the model, it can be seen that the safety index has the largest influence on the injury total, followed by the severity of the collision, the fatality total and the presence of an aggressive maneuver. The high coefficient (1.01) between the index and the injury total is indicative of decreasing safety as decreasing roadway injuries is a major safety goal of researchers worldwide (Nantulya et al., 2003). Further supporting the validity of the model is the positive coefficient associated with severity (0.16) and fatality total (0.01). The low value obtained for fatality total is a reflection of the extremely low number of fatal crashes in the dataset. The value of 0.00 for aggressiveness indicates that the influence is on the order of less than 0.005 and due to the "dummy" nature of the variable; this result is within the realm of expectation. This conclusion is validated by the fact that removing such variable leads to a lesser statistical significance. Accordingly, microscopic behavioral data related to aggressiveness are complementary to crash related data leading to a more robust structural model.

Looking finally at the covariances between variables, we observe that vehicle speed increases proportionally with lane width (0.65), and inversely with number of lanes (−0.12) and AADT (−4.19). This indicates that drivers will travel at higher speeds when they are provided with wider lanes, but on roadways with increased

congestion (higher AADT), the vehicle speed is decreased. The inverse proportionality with the number of lanes, is likely due to the increased congestion on the roadway, i.e. the road is designed with a higher lane count because the AADT is higher.

7.2. Uninterrupted flow

Table 8 displays the normalized geometric contribution to the safety index for each uninterrupted flow segment. They are listed in order from safest to least safe segment of roadway based on their geometric characteristics that were included in the model.

Additionally, Fig. 7 displays Google™earth images of segments 1 and 2, the safest and least safe segments on I-66 respectively.

For the uninterrupted flow situations discussed in the previous section, geometric characteristics positively influence the safety propensity index while environmental and traffic characteristics provide a negative influence. By taking the absolute value of each dimension's coefficient, it can be seen that traffic characteristics have the highest associated value (0.07) followed by environmental (0.06) and geometric (0.02). These results are consistent with previous research such as that of Karlaftis and Golias who, in their 2001 study on rural multi-lane highways, had similar findings in terms of the relative importance of the characteristics (Karlaftis and Golias, 2002).

Examining the Traffic Characteristics dimension, for interrupted flow scenarios, the largest influence comes from AADT (−21.07) which, coupled with the negative coefficient between L3 and the Safety Index, indicates that the least safe roadways are the ones that are the most highly traveled. The other influence on this dimension was the speed limit, which was the opposite of what one would expect, as higher speed limits were indicative of a safer driving experience: for interrupted flow scenarios, relatively, the flow levels rather than the speed limits govern the corresponding traffic dynamics (Section 7.3).

Moving to the Infrastructure Characteristic dimension, we see that the largest influence comes from lane count (0.72), followed

Table 8
Uninterrupted flow – normalized geometric contribution to safety index.

Segment number	Highway name	Mile markers	Geometric contribution
8	I-395 North	3–4	0.87
1	I-66 West	12–13	0.95
7	I-81 North	291–292	0.95
3	I-66 West	70–71	0.98
4	I-495 South	1.5–2.5	1.11
6	I-495 South	12–13	1.11
9	I-95 South	152–153	1.15
5	I-495 South	6–7	1.36
2	I-66 West	53–54	1.51

Fig. 7. I-66 West safest (left) and least safe (right) segments.

by shoulder width (0.40) and grade (−0.13). Positive values for lane count and shoulder width indicate that with more lanes and increased shoulder width there is a decrease in safety on the roadway: consider that in areas with wide shoulders and more lanes, drivers may perceive an opportunity for an increased margin of error and drive at higher speeds and more aggressively in general. Furthermore, areas with narrow shoulders and a limited number of lanes may lead to cautious driving due to the perceived risk. Since shoulders are typically unoccupied by vehicles, it is possible that a drivers perception of safety could cause them to alter their driving behavior. Additionally support for these findings can be inferred from the reductions in free-flow speed that are associated with narrower shoulders and less lanes on a freeway. This equation suggests that a manifestation of the lane count and shoulder width variables is that of lower driving speeds.

Analysis of the environmental characteristic demonstrates that there is a less safe driving environment in daylight conditions and in instances with increased friction. As far as lighting goes, the same logic applies that was used to explain the similar influence on interrupted flow roadways. For the friction, it is important to observe that the error value for the friction is less than 0.005. The small influence of the friction variable, coupled with the limited variability of the variable itself may explain why this value is so low (Rakha et al., 2008).

The right side of the model illustrates that the safety index has the strongest influence on the injury total, followed by the severity of the collision and drivers aggressive maneuvers. The high coefficient (2.30) between the safety index and the injury total is not only expected, but nearly *required* in the sense that many other results would be indicative of a model that is incorrect. Similarly, another high coefficient (0.34) with severity (which is 1 for a severe collision and 0 otherwise) is additional evidence that the model used in this paper is creating a proper framework for measuring safety. For the aggressive maneuver variable, the coefficient itself is small – especially compared to that of the other endogenous variables. However, this variable contributes positively to the significance of the model.

Finally, looking once again at the covariance, we observe the expected result of speed limit increasing proportionally with shoulder width (2.59) (this should be the case as wider shoulders mean that the design speed limit can be higher). Lane count also increases proportionally with shoulder width (0.28), which is once again an expected design feature. A minimal covariance is observed between the friction value and the speed limit.

7.3. Model comparison

Individually, both structural models presented in this paper produce results that are consistent with previous research and intuitive expectations. However, some of the most significant observations about roadway safety are arrived at when comparing the two models to one another.

The fact that the same model would not converge for the two different flow situations is a major conclusion in itself. The models constructed are unique for the flow situations they describe; demonstrating that while safety is a major concern on all roads, it needs to be approached from different angles. An interesting difference between the models was that vehicle speed played a role for in the interrupted flow model, but it was speed limit that was included in the uninterrupted model. This distinction indicates the major role of differential speeds in forming traffic dynamics in uninterrupted flow conditions versus the major role the flow level plays in governing the perception of drivers to safety during interrupted flow conditions; for example, with higher speed limits on uninterrupted roadways, vehicles may be nearly uniformly traveling at speeds where if accidents do occur, there are most likely going to be injuries. For the interrupted flow scenarios, the lower speed limits may indicate that drivers put themselves at an increased risk by traveling at speeds exceeding the posted limit.

Commonalities between the models also provide a basis for conclusions. As expected, speed (or the manifestation thereof) has one of the most significant effects on safety for both flow situations. This natural product of the basic laws of physics is a welcomed similarity between the two models. Another similarity is demonstrated by the fact that AADT plays a role in both models, meaning that more traffic creates more of an opportunity for injury. The following qualitative table provides a summary of the results for common variables between the two models (Table 9).

In both models one of the largest influences in terms of infrastructure characteristics came from shoulder width. Interpretation of the influence of shoulder width in each model is outlined above, and can readily be explained by taking into account the flow scenario that is present.

It should be noted that the above results are not intended to offer definitive conclusions, especially when the corresponding coefficient or covariance term are small in relative magnitude. However, they illustrate how the technique proposed in this paper can be an effective tool for roadway segments performance assessment in support of safety policy analysis (identification of critical safety contributor(s) and the direction of the corresponding contribution: magnitude and sign of $\mathbf{\Lambda X}$, Eq. (1)). Another important aspect of the contribution of this study is the ability to obtain numerical values of each segment's SPI either by considering an aggregate value for each dimension (averaged variables' values for a given segment) or by taking a dimension's contribution to the safety index (Δ matrices in Eq. (3)); such ranking exercise becomes straightforward especially if the corresponding exogenous variables are constant for a given roadway segment. The results are illustrated

Table 9
Qualitative results – common exogenous variables.

Exogenous variable	Effect on safety for flow condition	
	Interrupted	Uninterrupted
Shoulder width	Increased shoulder width means increased safety	Increased shoulder width means decreased safety
Lance count	Increased number of lanes means decreased safety	Increased number of lanes means decreased safety
Lighting	Daylight means decreased safety	Daylight means decreased safety
AADT	Increased AADT means decreased safety	Increased AADT means decreased safety

in Tables 7 and 8 for the geometric dimension. The segments that exhibit the greatest propensity for "unsafe" endogenous measures can be considered as a mitigation priority by the corresponding transportation agency (Federal Highway Administration–NHTSA, local governments, State Departments of Transportation – DOTs).

8. Concluding comments

In this paper, a safety propensity index was developed for both interrupted and uninterrupted flow scenarios utilizing the structural equation modeling approach. This index helps to define not only which characteristics contribute to roadway safety, but also how they do so. The strength of the relationships is captured by the safety propensity index.

Large data sets were combined and vetted in order to create usable inputs for both interrupted and uninterrupted flow scenarios. Pertinent variables were identified and an initial structural model was proposed based on the data available. Based on these equations, an exploratory data analysis was conducted for 9 uninterrupted and 10 interrupted flow scenarios. Through factor analysis, the initial model was improved so that it produced a quantitative output. This output was then analyzed in an effort to identify and explain the characteristics that effect safety in each flow scenario. The models were then compared to one another to observe how the factors effecting safety vary with flow scenario.

The major limitation for this project was the availability and quality of data collected by the state of Virginia; creating major difficulties in developing converging and statically significant models. By collecting more in-depth and higher quality data, the state of Virginia would allow for a more complete analysis of roadway safety throughout the state. Better data collection would allow future research to investigate how driver behavior can be included in safety models. Out of necessity, one of the major contributions of this study became the inspection and manipulation of these datasets.

While the models proposed by this study produced results for the roadway segments outlined above, it remains to be seen if the same models would produce results if applied universally to different roadways across the nation. Additionally, if the models did produce results, it is unknown whether or not the weighting factors for the different characteristics would be similar to those achieved by this analysis. Universal application is an empirical matter and can only be discovered through further research. Given the nature of the consistency between the results of this study and expectation, it can be anticipated that the framework developed and presented above can form the basis for a systematic and common conceptual and quantitative framework for identifying and understanding the factors that affect safety on roadways with both interrupted and uninterrupted flow.

Acknowledgements

References

Baker, J.S., Fricke, L.B., 1990. Drag Factor and Coefficient of Friction in Traffic Accident Reconstruction, Topic 862 of the Traffic Accident Investigation Manual (1 ed., Vol. 2). Northwestern University Traffic Institute, Evanston, IL.

Cortina, J., 1993. What is coefficient alpha? An examination of theory and application. Journal of Applied Psychology 78 (1), 98–104.

Golob, T.F., 1988. Structural Equation Modeling of Travel Choice Dynamics. The University of California Transportation Center, Berkeley, CA.

Golob, T.F., 2001. Structural Equation Modeling for Travel Behavior Research. Institute of Transportation Studies, University of California, Irvine, Irvine, CA.

Golob, T.F., Meurs, H., 1986a. A Structural Model of Temporal Change in Multimodal Travel Demand. Institute of Transportation Studies, University of California, Irvine, Irvine, CA.

Golob, T.F., Meurs, H., 1986b. Development of Structural Equation Models of the Dynamics of Passenger Travel Demand. Institute of Transportation Studies, University of California, Irvine, Irvine, CA.

Gross, F., Jovanis, P.P., Eccles, K., Chen, K.-Y., 2009. Evaluation of Lane and Shoulder Width Combinations on Rural, Two-Lane, Undivided Roads. Federal Highway Administration, Washington, DC.

Hamdar, S.H., Mahmassani, H.S., Chen, R.B., 2008. Aggressiveness propensity index for driving behavior at signalized intersections. Accident Analysis and Prevention 40 (1), 315–326.

Hu, L.-T., Bentler, P.M., 1998. Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification. Psychological Methods 3 (4), 424–453.

Jones, I.S., Whitfield, R.A., 1988. Predicting injury risk with new car assessment program crashworthiness ratings. Accident Analysis and Prevention 20 (6), 411–419.

Joshua, S., Garber, N., 1990. Estimating truck accident rate and involvements using linear Poisson regression models. Transportation Planning and Technology 15 (1), 41–58.

Karlaftis, M., Golias, I., 2002. Effects of road geometry and traffic volumes on rural roadway accident rates. Accident Analysis and Prevention 34 (3), 357–365.

Lee, J.-Y., Chung, J.-H., Son, B., 2008. Analysis of traffic accident size for Korean highway using structural equation models. Accident Analysis and Prevention 40 (6), 1955–1963.

Li, J., Abdelwahab, W., Brown, G., 1994. Join effects of access and geometry on two-lane rural highway safety in British Columbia. Canadian Journal of Civil Engineering 21 (6), 1012–1024.

Nantulya, V., Sleet, D., Reich, M., Rosenberg, M., Peden, M., Waxweiler, R., 2003. Introduction: the global challenge of road traffic injuries: can we achieve equity in safety? Injury Control and Safety Promotion 10 (1–2), 3–7.

National Highway Traffic Safety Administration (NHTSA), 2010. Traffic Safety Facts 2009. A Compelation of Motor Vehicle Crash Data from the Fatality Analysis Reporting System and the General Estimates System. U.S. Department of Transportation, National Center for Statistics and Analysis, Washington, DC.

Paleti, R., Eluru, N., Bhat, C., 2010. Examining the influence of aggressive driving behavior on driver injury severity in traffic crashes. Accident Analysis and Prevention 6, 1839–1854.

Rakha, H., Farzaneh, M., Arafeh, M., Sterzen, M., 2008. Inclement weather impacts on freeway traffic stream behavior. Transportation Research Record 2071, 8–18.

Rakha, H., Arafeh, M., Park, S. (2009). Modeling Inclement Weather Impacts on Traffic Stream Behavior Transportation Research Board 88th Annual Meeting, 2009. Washington, DC.

Sijtsma, K., 2009. On the use, the misuse, and the very limited usefulness of Cronbach's alpha. Psychometrika, 107–120.

Talebpour, A., Mahmassani, H. S., & Hamdar, S. H. (2012). Safety First: Assessing Congestion Effects on Experienced Driver Risk Using a Microsimulation Approach. Transportation Research Board 91st Annual Meeting.

Treiber, M., Hennecke, K., Helbing, D., 2000. Congested traffic states in empirical observations and microscopic simulations. Physical Review E 2 (2), 1805–1824.