Incident Clearance Time Analysis for Korean Freeways Using Structural Equation Model

Ju-Yeon LEE
Ph. D. Candidate
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: ljourney@yonsei.ac.kr

Jin-Hyuk CHUNG (Corresponding Author)
Associate Professor
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: jinchung@yonsei.ac.kr

Bongsoo SON
Professor
Dept. of Urban Planning & Engineering
Yonsei University
134, Shinchon-Dong, Sudaemoon-Gu,
Seoul, KOREA, 120-749
Fax : +82-2-393-6298
E-mail: sbs@yonsei.ac.kr

Abstract: Non-recurrent congestion mainly due to the traffic incident is unpredictable. However, its substantial impact on the traffic flow temporarily reduces capacity of roads and incurs enormous time loss. In order to minimize the economic loss from the congestion, the non-recurrent congestion should be dealt with properly and effectively. This study aims to understand major factors affecting the clearance time using complete 2,647 incident data occurred on freeways of Korea in 2005. Using SEM relationships among various exogenous variables and incident clearance time are explored. The model developed in this study is validated using incident data in 2004. The resulting model provides valuable information for the incident clearance time, which can be utilized to manage the incident effectively.

Keywords: Incident Clearance Time, Structural Equation Model with Observed Variables, Freeway Incident

1. INTRODUCTION

Traffic congestion is classified into non-recurrent and recurrent congestion by their periodical and location features. Since the unbalance of travel demand and transportation supply causes the recurrent congestion, it usually occurs in similar time of day and sections of roadways. Hence, the recurrent congestion is predictable so that it can be removed permanently even though it requires large-scale investment (e.g., improvement of infrastructure or/and reduction of travel demand). On the contrary, non-recurrent congestion mainly due to the traffic incident is unpredictable. However, its substantial impact on the traffic flow temporarily reduces capacity of roads and incurs enormous time loss. In order to minimize the economic loss from the congestion, the non-recurrent congestion should be dealt with properly and effectively. In Korea, freeway demand has increased rapidly during last thirty years and the total length of freeways is over 3,000km now. However, the number of accidents per one kilometer of freeways is higher than any other countries. The rapid increase of travel demand may have influenced on the high rates of traffic accident and 2,800~3,500 incidents per a year are occurred in Korean freeway. According to statistics from the Korean Expressway Cooperation, sixty percentages of traffic congestion occurred in Korean freeway are due to incidents. Hence, it is a very important issue how to manage non-recurrent traffic congestion to minimize the economic loss such as human injury (e.g., death and injury), property damaged and time loss (e.g., traffic delay) on the highways.

We have two plausible solutions to diminish such loss, one of which is to reduce accidents itself and the other is to manage accidents effectively. The first approach is a permanent solution, which can curtail the loss dramatically but accidents cannot be prevented completely. Recently ITS technology (e.g., advanced surveillance system) makes the second solution greatly applicable. Hence, the accident management system [AMS] is regarded as the one of effective tools to diminish economic loss from the accidents. Once an incident occurs, the incident should be cleared rapidly to minimize the impact of incident on traffic flows. It is believed that shortening the incident clearance duration is very effective to diminish the time loss of users from traffic accidents. Therefore, understanding of main factors affecting the clearance time is a fundamental issue, which is an objective of this study.

From the previous researches, various factors such as the number of vehicles involved in the incident, truck involvement, the number of death and/or injured persons, peak or non-peak time, day or night time and weather condition are proved to have relationships with the incident clearance and duration time. Many researchers have adopted mainly multiple regression model, non-parametric regression model and classification tree model to capture these relationships. However, all those factors inherently interact in complicate ways so that the interrelationships among the variables are not easily identified using the models. Some variables correlate to each other so that these variables have both direct effect and indirect effect on incident clearance time. It is difficult to capture the relationships using the classical regression models. In this study, a structural equation model is adopted in this study to capture the complex relationships among variables. In addition, we can understand direct/indirect effects of the variables using SEM because SEM can handle complex relationship among endogenous and exogenous variables simultaneously.

We use complete 2,647 incident data occurred on freeway of Korea in 2005 and estimate the relationships among exogenous factors and incident clearance time using structural equation models. The model is also validated using 500 incident data set Korea in 2004.

This paper consists of six sections. Next section summarizes the previous works related to incident duration and/or clearance time and structural equation models. The introduction of structural equation models and methodology are presented in section three. In the forth section description of the data in use is offered, which is incident dataset collected from the Korea Freeway Cooperation. SEM of incident clearance time, interpretation of the results, and validation of the model are shown in section five and six, respectively. Finally, the paper closes with conclusions and recommendation for future studies in section seven.

2. PREVIOUS WORKS

Many researchers have put in a great deal of effort to explain traffic accident occurrence and factors affecting incident duration and clearance time. They attempted to develop various types of models to explain the duration and clearance time of incidents. Several selected studies related to ours are summarized in this section, topics of which are accident analysis, incident duration estimation and Structural Equations Model (SEM).

Smith and Smith (2001) suggested the use of decision trees for incident duration estimation. A set of decision trees developed using standard classification techniques proposed by Breiman *et al.* (1984). The effects of factors influencing on the clearance time of freeway accidents were examined in their study, which are time of day, day of week, response of agencies like Emergency Medical Services (EMS), Fire Department, Freeway Incident Response Team (FIRT), Local Police, tow-trucks and others, number of vehicles and truck or bus involvements. Kaan Ozbay and Nebahat Noyan (2006) proposed to use Bayesian Networks. BN can estimate parameters for reasonably small number of variables in many applications. An advantage of BN is clearly to help operators to determine the incident clearance time. Data used in the study was collected from incident clearance survey forms filled out by Virginia State Police, Virginia Department of Transportation Safety Service Patrol, Fairfax Country Police Department and Fairfax County Fire and Rescue Department. The survey form is composed of total clearance time, type of incident (road hazard, property damage, personal injury, vehicle fire, type of incident vehicles, weather related, etc.), number of police vehicles, fire engines or ambulances, total number of lanes, type of roadway. Khattak *et al.* (1995) used a series of truncated regression models to explain the factors influencing on incident. They showed main factors influencing incident duration, which are incident type, number and vehicle type involved in incident, number and severity of injuries and number of lanes affected affect to incident duration time. The most popular approach to analyze incident duration is the hazard-based models, which allow the explicit analysis of duration effects. Nam and Mannering (2000) applied the hazard-based models to determine the factors affecting incident duration with 1994 and 1995 data from highway incidents in State of Washington, U. S. A. They concluded that the factors associated with incident durations are incident characteristics, environmental conditions, location factors and operational response.

Many researches attempted to understand the complex relationships among the variables using structural equations model. Hamdar and *et al*. (2007) developed a quantitative intersection aggressiveness propensity index (API) using SEM. The index was intended to capture the overall propensity for aggressive driving to be experienced at a given signalized intersection. The index was a latent quantity that could be estimated from observed environmental, situational and driving behavioral variables using SEM techniques. The exogenous variables were number of heavy vehicles, number of pedestrians, traffic volume, average queue length, percent grade, number of lanes, number of left turn lanes and so forth. Lee *et al.* (2008) analyzed relationships among the traffic accident size and the cause factors using SEM. They postulated that road factors, driver factors and environment factors are exogenous latent variables and accident size factor is an endogenous latent variable in SEM. The observed variables for latent variables were pavement type, horizontal and vertical alignment characteristics, weather condition, road surface condition, day time or night time, vehicle type, driver's gender and their age and forth. The model showed that road factors, driver factors and environment factors are strongly related to the accident size.

3. STATISTICAL METHODOLOGY

(This chapter is based on the chap. 3 of the reference Lee *et al.* (2008))

Structural Equation Model (SEM) is a technique that can handle a large number of endogenous and exogenous observed variables simultaneously. Since SEM consists of a set of equations that are specified by direct links between variables, it can be called “the simultaneous equations” from the perspective. SEM with latent variables may have a combination of the two components: 1) a measurement sub-model for the latent variables (endogenous, exogenous) and 2) a structural sub-model. A measurement sub-model for the latent variables is given by next equation.

$$\begin{aligned} x &= \Lambda_x \xi + \delta \\ y &= \Lambda_y \eta + \varepsilon \end{aligned} \tag{1}$$

where, x is a column vector of q exogenous variables, y is a column vector of p endogenous variables; ξ is a column vector of n latent exogenous variables, η is a column vector of m latent endogenous variables; Λ_X is the matrix (q×n) of coefficient relating x to ξ, Λ_Y is the matrix (p×m) of coefficients relating y to η ; δ and ε are column vectors of the error terms for x and y, respectively.

A SEM without latent variables (same as the structural sub-model) is defined in next equation.

$$y = \mathrm{B}y + \Gamma x + \zeta \tag{2}$$

Where, y is a column vector of p endogenous variables, x is a column vector of q exogenous variables, and ζ is a column vector of the error terms. The structural parameters are the elements of the three matrices: B (parameters of the B matrix denoted by β) is the matrix (p×p) of direct effects between pairs of the p endogenous variables; Γ (parameters of Γ matrix denoted by γ) the matrix (p×q) of regression effects for p endogenous variables and q exogenous variables. In this research, we use only the structural sub-model.

Parameter estimation methods

A structural equation model is applied in this research to estimate a simultaneous model that presents the interrelationships among various exogenous variables and endogenous variables. The LISREL version 8.51 and PRELIS / SIMPLIS software are used to estimate the model in this research.

To estimate parameters in SEM, LISREL offers seven different methods: instrumental variables (IV), two-stage least squares (TSLS), unweighted least squares (ULS), generalized least squares (GLS), maximum likelihood (ML), weighted least squares (WLS), and diagonally weighted least squares (DWLS). Generally, ML method is most widely used as estimator of parameters because (N -1) F_{ML} is approximately distributed in large samples (N) as χ^2 with an assumption of multivariate normality of variables. However, when distributions of variables do not have multivariate normality or when they have excessive kurtosis, it is desirable to employ the WLS estimation method. In order to determine an appropriate estimation method, the normality of the variables should be statistically tested.

The fundamental concept in estimating the structural equation model is that the population covariance matrix of observed variables (Σ) can be expressed in terms of unknown parameter θ, which includes all the unknown parameters in B and Γ matrices. Each element of the population covariance matrix can be written as a function of one or more model parameters, or $\Sigma = \Sigma(\theta)$. Hence, the parameters θ can be estimated by minimizing the discrepancies between the sample covariance matrix S and the population covariance matrix expressed in terms of unknown parameters $\Sigma(\theta)$.

The maximum likelihood (ML) approach will estimate θ by minimizing the fit function

$$F_{ML}(\theta) = \log|\Sigma(\theta)| + tr\left[S\,\Sigma^{-1}(\theta)\right] - \log|S| - (p+q) \tag{3}$$

This fit function assumes that the observed variables have a multinormial distribution.

The weighted least squares (WLS) approach will estimate θ by minimizing the fit function

$$F_{WLS}(\theta) = [s - \sigma(\theta)]' W^{-1} [s - \sigma(\theta)] \tag{4}$$

Values of θ are estimated so as to minimize the weighted sum of squared deviations of s from $\sigma(\theta)$ (Bollen, K. A., 1989). WLS method does not assume multivariate normality of variables and does need asymptotic covariance of variables for estimation.

4. DESCRIPTIVE STATISTICS OF DATA

The data used in this study are 2,880 accident records on Korean highway during the year 2005, which are collected by Korean Expressway Corporation. Each accident record has various and rich information such as the incident clearance time, the incident location, time zone (peak or non-peak time), the day (weekend or weekday), weather condition, day or night time, horizontal alignment, vertical alignment, vehicle type, driver's gender, driver's age, the number of deaths, the number of injured persons and the number of involved vehicles etc. After eliminating missing and erroneous data, 2,647 accident data are utilized in this search.

Table 1 Elements of Estimation Methods ML and WLS

Σ	the population covariance matrix of observed variables
$\Sigma(\theta)$	the implied covariance matrix of structural parameters
S	the sample covariance matrix
s	a vector of $\frac{1}{2}(p+q)(p+q+1)$ elements obtained by placing the nonduplicated elements of S
$\sigma(\theta)$	the corresponding same-order vector of $\Sigma(\theta)$
W^{-1}	$\frac{1}{2}(p+q)(p+q+1) \times \frac{1}{2}(p+q)(p+q+1)$ positive-definite weight matrix

Table 2 Definition of Variables

The incident clearance time	The time between arrival at the incident scene and incident clearing (minutes)
The number of death	The number of persons killed in the incident (persons)
The number of injured persons	The number of persons injured in the incident (persons)
The number of involved vehicles	The number of vehicles involved in the accident (vehicles)

Table 3 Descriptive Statistics of Accident Records

	Frequency (%)	Involved vehicles		Deaths		Injured persons		Incident clearance time	
		Mean	Var.	Mean	Var.	Mean	Var.	Mean	Var.
Time zone									
Peak	481 (18.2%)	1.39	0.78	0.06	0.07	0.55	5.51	40.53	1038.06
Non-peak	2,166 (81.8%)	1.43	0.98	0.09	0.13	0.39	0.86	45.50	1450.88
The day									
Weekdays	1,760 (66.5%)	1.44	0.74	0.08	0.10	0.44	2.17	47.09	1607.80
Weekends	887 (33.5%)	1.39	1.35	0.09	0.16	0.37	0.78	39.66	889.91
Weather condition									
Snow or fog	103 (3.9%)	1.56	1.08	0.07	0.06	0.22	0.31	53.16	1976.81
Others (clear etc.)	2,544 (96.1%)	1.42	0.94	0.08	0.12	0.42	1.76	44.25	1352.27
Day or night									
Day time	1,610 (60.8%)	1.40	0.74	0.07	0.10	0.39	1.11	42.52	1151.15
Night time	1,037 (39.2%)	1.46	1.27	0.11	0.14	0.46	2.63	47.82	1717.08
Incident location									
Toll-gate or ramp	248 (9.4%)	1.19	0.55	0.02	0.02	0.14	0.22	34.72	791.51
Main road	2,399 (90.6%)	1.45	0.98	0.09	0.13	0.44	1.85	45.62	1429.19
Horizontal alignment									
Straight (R≥500m)	2,418 (91.3%)	1.44	0.97	0.09	0.13	0.42	1.79	44.94	1396.58
Curve(R<500m)	229 (8.7%)	1.32	0.62	0.05	0.06	0.33	0.87	40.95	1185.00
Vertical alignment									
Upslope (>3%)	56 (2.1%)	1.30	0.21	0.14	0.16	0.39	0.45	42.14	1066.37
Others	2,591 (97.9%)	1.43	0.96	0.08	0.12	0.42	1.73	44.65	1386.17
Vehicles type									
Auto and/or van	1,623 (61.3%)	1.35	0.86	0.08	0.11	0.43	1.09	34.72	480.74
Truck and/or trailer	1,024 (38.7%)	1.55	1.05	0.09	0.14	0.39	2.68	60.25	2404.60

In Table 3, mean and variances of four variables such as the number of deaths, the number of injured, the number of involved vehicles, and the incident clearance time are summarized by 8 variables, which may affect on the incident clearance time. The eight variables are binary form, the threshold values of which are adopted from references and various statistical tests. For example, Lee *et al.* (2008) showed that upslope section (more than 3%) and curve section (R<500m) can help decrease the ‘accident size’, we adopted the value as the thresholds of vertical alignment and horizontal alignment. We postulates that the incident clearance time decreases as the accident size smaller. Generally, drivers decrease operating speed on toll-gate and/or ramp compared to main roads so that accident size on toll-gate and/or ramp is smaller than those on main road. Therefore, in this study ‘Incident location’ variable is segmented as ‘Toll-gate or ramp’ and ‘Main road’. In case of ‘Weather condition’ variable, it is made up of ‘snow or fog’ and ‘others’. Because the incident clearance time would be extended sharply when it is snowy and/or fog.

Comparisons of the mean values give meaningful indications to relationships between the incident clearance time and several factors. When incidents occur in ‘night time’ and ‘main road’ they have larger the number of involved vehicles, the number of deaths, the number of injured persons and incident clearance time than others. Similarly, the comparison also shows that the number of involved vehicles, the number of injuries and the incident clearance time increase as horizontal alignment is straight, vertical alignment is not upslope.

5. DEVELOPMENT OF SEM

5.1 Conceptual Model Structure

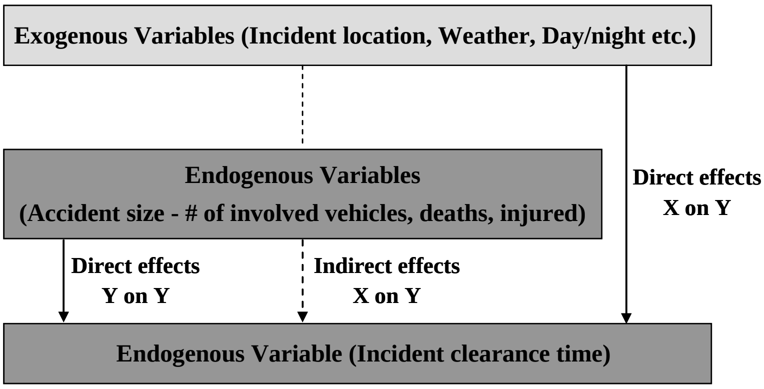

Figure 1 Conceptual Model Structure

To estimate the incident clearance and/or duration, many previous researches included exogenous variables such as the number of involved vehicles, involvement of truck, weather condition, day or night time and the number of death or injured person in their models (Shin and Kim, 2002, Nam and Mannering, 2000). Lee *et al.* (2008) also developed a SEM for 'accident size' having road factor, environment factor and driver factor as latent endogenous variables.

In this study, the three-level causal structure is employed as shown in Figure 1. Firstly, we postulate exogenous variables such as Incident location, Weather condition and Day/night etc. effect on Accident size variables such as the number of involved vehicles, deaths and injured. Secondly, the SEM model is a set that exogenous variables as well as 'Accident size' variables effect on 'Incident clearance time'.

5.2 Data Coding and Estimation Method

All categorical and nominal variables are transferred to binary variables because it is a proper way to deal with categorical and nominal variables in SEM and they allow us to identify the nonlinear influence of categorical and nominal variables on endogenous variables in SEM. Eight X observed variables (exogenous variables) are binary (0/1) variables and four Y observed variables (endogenous variables) are ordinal or continuous variables. Since the variables such as the number of deaths, injured persons and involved vehicles have less than fifteen categories, LISREL program treats them as ordinal variables. In case of the incident clearance time variable has more than fifteen categories, it is treated as continuous variables.

Table 4 Coding Values of Each Variable

Variable	Coding input value	Variable	Coding input value
Time zone	1: Peak time 0: Non-peak time	Vertical alignment	1: Upslope (>3%) 0: Others
The day	1: Weekdays 0: Weekends	Type of involved vehicles	1: Truck and/of Trailer 0: Auto and/or Van
Weather condition	1: Snow or Fog 0: Others	The number of deaths	Persons
Day or night	1: Night time 0: Day time	The number of injured persons	Persons
Incident location	1: Main road and others 0: Tollgate or Ramp	The involved vehicles	Vehicles
Horizontal alignment	1: Curve (R<500m) 0: Straight (R≥500m)	The incident clearance time	Minutes

Table 5 Test of Univariate Normality for Continuous Variables

	Skewness		Kurtosis		Skewness and Kurtosis	
	Z-Score	P-Value	Z-Score	P-Value	Chi-Square	P-Value
# of Incident Clearance Time	64.665	0.000	144.733	0.000	24840.719	0.000

Total 2,647 accident samples are analyzed and correlation matrix among 12 variables is created using PRELIS software considering the variable distribution characteristics. A fundamental principle in PRELIS is the distinction between variables of different scale types. All X observed variables are binary (two ordinal) variables and Y observed variables are continuous or ordinal variables in our model so that we use polychoric correlations matrix computed by PRELIS. This can be used to compute a weight matrix (W) for WLS in LISREL (Joreskog, K. G., Sorbom, D., 2000).

Univariate and multivariate normality is tested to determine the estimating method. PRELIS program gives univariate and multivariate tests of normality for continuous variables and the results are showed in Table 5. According to the result, we can reject the hypothesis that the distribution of the variables has normality. Generally, WLS (weighted least squares) does not assume multivariate normality and is known as the asymptotically distribution free. Therefore, WLS estimation method is employed because distributions of variables do not have multivariate normality.

5.3 Structural Equation Model for the Incident Clearance Time

The initial SEM of the incident clearance time is depicted in Figure 2. As previously stated, the model has three-level (one exogenous and two endogenous variable groups) causal structure. It has eight X variables (Time zone, The day, Weather condition, Day or night, Incident location, Horizontal alignment and Vertical alignment and Vehicles type) and four Y variables (The number of deaths, injured persons, involved vehicles and The incident clearance time). However, the result of the initial SEM shows that the estimated parameters of several variables do not have statistical significance. All estimated parameters of direct and indirect effects with t-value are summarized in Table 6 and 7.

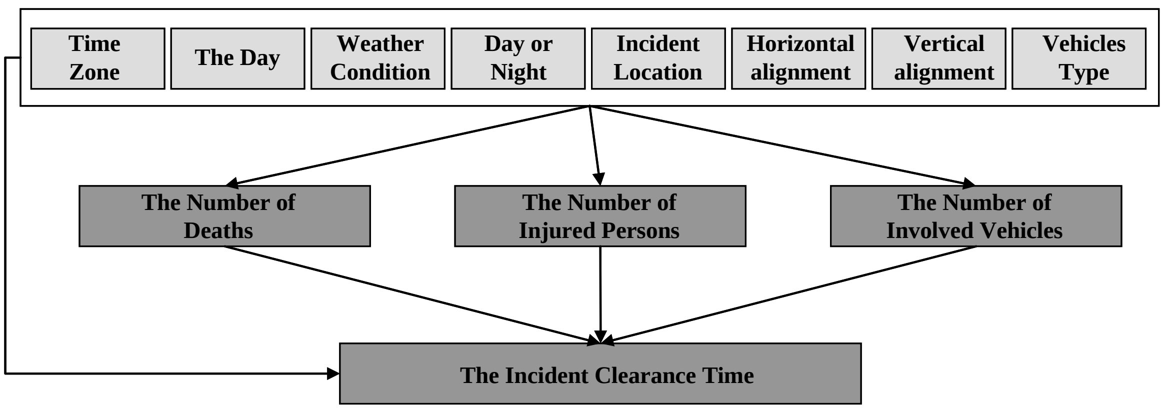

Figure 2 The Initial SEM for Incident Clearance Time

Table 6 Direct[1)], Indirect[2)] and Total Effects[3)] Between Exogenous and Endogenous Variables (X on Y) with t-value of the Initial SEM

		Time Zone	The day	Weather	Day/ Night	Location	Horizon. Align.	Vertical Align.	Type of Veh.
The number of deaths	1)	0.030 (0.302)	-0.064 (-0.824)	-0.175 (-1.457)	0.094 (1.225)	0.851 (9.395)	-0.194 (-1.851)	-0.621 (-5.905)	0.063 (0.984)
	2)	-	-	-	-	-	-	-	-
	3)	0.030 (0.302)	-0.064 (-0.824)	-0.175 (-1.457)	0.094 (1.225)	0.851 (9.395)	-0.194 (-1.851)	-0.621 (-5.905)	0.063 (0.984)
The number of injured persons	1)	0.168 (1.727)	-0.023 (-0.284)	-0.334 (-2.631)	-0.004 (-0.057)	0.933 (9.472)	-0.212 (-2.006)	-0.773 (-6.700)	-0.036 (-0.573)
	2)	-	-	-	-	-	-	-	-
	3)	0.168 (1.727)	-0.023 (-0.284)	-0.334 (-2.631)	-0.004 (-0.057)	0.933 (9.472)	-0.212 (-2.006)	-0.773 (-6.700)	-0.036 (-0.573)
The number of involved vehicles	1)	0.183 (1.357)	-0.040 (-0.357)	-0.230 (-1.329)	0.000 (0.002)	1.318 (9.027)	-0.216 (-1.523)	-1.203 (-7.650)	0.218 (2.793)
	2)	-	-	-	-	-	-	-	-
	3)	0.183 (1.357)	-0.040 (-0.357)	-0.230 (-1.329)	0.000 (0.002)	1.318 (9.027)	-0.216 (-1.523)	-1.203 (-7.650)	0.218 (2.793)
The incident clearance time	1)	-0.020 (-0.891)	0.007 (0.370)	0.088 (2.505)	0.029 (1.265)	0.112 (3.655)	- ( )	- ( )	0.378 (13.905)
	2)	0.045 (1.031)	-0.019 (-0.543)	-0.080 (-1.448)	0.015 (0.446)	0.425 (6.037)	-0.079 (-1.687)	-0.362 (-5.650)	0.056 (1.592)
	3)	0.025 (0.537)	-0.012 (-0.323)	0.009 (0.142)	0.044 (1.171)	0.537 (7.965)	-0.079 (-1.687)	-0.362 (-5.650)	0.435 (13.711)

( ): t-value, -:not identified

Table 7 Direct[1)] and Total Effects[2)] Among Endogenous Variables (Y on Y) of the Initial SEM

		# of deaths	# of injured	# of involved vehicles	Clearance time
The number of deaths	1), 2)	-	-	-	-
The number of injured persons	1), 2)	-	-	-	-
The number of involved vehicles	1), 2)	-	-	-	-
The incident clearance time	1)	0.162 (2.090)	0.008 (0.204)	0.212 (2.829)	-
	2)	0.162 (2.090)	0.008 (0.204)	0.212 (2.829)	

The final SEM for incident clearance time is developed after removing the parameters which were insignificant in the initial SEM. It has six X variables (Time zone, Weather condition, Incident location, Horizontal alignment, Vertical alignment and Vehicles type) and four Y variables (The number of deaths, injured persons, involved vehicles and The incident clearance time). The estimated parameters with t-value of the final SEM are summarized in Table 8 and 9.

6. EMPIRICAL RESULTS

6.1 Overview of Model Fit and Validation

Various types of measure of goodness of fit are available for this type of model, including General Fitting Index (GFI), Adjusted General Index (AGFI), Root Mean square Residual

(RMR), Root Mean Square Error of Approximation (RMSEA) and Chi-square (Joreskog and Sorbom, 1995). The indices have ranges from zero to one meaning a perfect fit. The RMR and the RMSEA near zero indicates a “good” model. The GFI and AGFI obtained from the finalized model are 0.999 and 0.996, respectively. The RMR for the final model is 0.054 which indicates a good correspondence between the replicated and original variance-covariance matrices. It is generally accepted the value of RMSEA for a good model should be less than 0.05 and the RMSEA of our model is 0.029. The Chi-square value of the model is 41.011 with 13 degrees of freedom. It is known that Chi-square value is so sensitive of sample size that P-value has low value along by increasing sample size (more than 200). The sample size of our model is so large (2,647) that Chi-square value is high. In SEM approach, therefore, the goodness of fit is generally performed by using other criteria such as RMSEA, RMR, AGFI, CFI and NFI. CFI and NFI measures how much the model better fits as compared to the baseline model, and these indices are supposed to lie between 0 and 1.

Table 8 Direct[1)], Indirect[2)] and Total Effects[3)] Between Exogenous and Endogenous Variables (X on Y) of the Final SEM

		Time Zone	The day	Weather	Day/ Night	Location	Horizon. Align.	Vertical Align.	Type of Veh.
The number of deaths	1)		-	-	-	0.796	-0.204	-0.587	-
			( )	( )	( )	(11.296)	(-2.144)	(-6.817)	( )
	2)		-	-	-	-	-	-	-
	3)	-	-	-	-	0.796	-0.204	-0.587	-
		( )	( )	( )	( )	(11.296)	(-2.144)	(-6.817)	( )
The number of injured persons	1)	0.083	-	-0.211	-	0.871	-0.210	-0.669	-
		(2.023)	( )	(-3.557)	( )	(13.886)	(-2.268)	(-8.533)	( )
	2)	-	-	-	-	-	-	-	-
	3)	0.083	-	-0.211	-	0.871	-0.210	-0.669	-
		(2.023)	( )	(-3.557)	( )	(13.886)	(-2.268)	(-8.533)	( )
The number of involved vehicles	1)	-	-	-	-	1.288	-0.180	-1.124	0.229
		( )	( )	( )	( )	(11.955)	(-1.397)	(-10.492)	(5.209)
	2)	-	-	-	-	-	-	-	-
	3)	-	-	-	-	1.288	-0.180	-1.124	0.229
		( )	( )	( )	( )	(11.955)	(-1.397)	(-10.492)	(5.209)
The incident clearance time	1)	-	-	0.046	-	0.111	-	-	0.380
		( )	( )	(1.493)	( )	(4.574)	( )	( )	(16.542)
	2)	-	-	-	-	0.454	-0.078	-0.379	0.059
		( )	( )	( )	( )	(8.509)	(-1.678)	(-8.435)	(3.104)
	3)	-	-	0.046	-	0.565	-0.078	-0.379	0.439
		( )	( )	(1.493)	( )	(11.758)	(-1.678)	(-8.435)	(19.472)

( ): t-value, -: either not indentified or insignicant

Table 9 Direct[1)] and Total Effects[2)] Among Endogenous Variables (Y on Y) of the Final SEM

		# of deaths	# of injured	# of involved vehicles	Clearance time
The number of deaths	1), 2)	-	-	-	-
The number of injured persons	1), 2)	-	-	-	-
The number of involved vehicles	1), 2)	-	-	-	-
The incident clearance time	1)	0.156 (2.161)	-	0.256 (4.713)	-
	2)	0.156 (2.161)		0.256 (4.713)	

Table 10 Goodness of Fit Statistics

Fit index		Fit index	
Chi-Square	41.011 (P=0.000)	AGFI (adjusted goodness of fit index)	0.996 (0.9 and more)
RMSEA (root means square error of approximation)	0.029 (0.05 and less)	CFI (comparative fit index)	0.998 (0.9 and more)
RMR (root mean square residual)	0.054 (0.05 and less)	NFI (normed fit index)	0.997 (0.9 and more)

The validation of the final SEM is conducted using 500 incident data sets, which are collected by the Korean Expressway Cooperation in 2004. The data sets in 2004 have same form with the data sets in 2005. *MSPR* and *MSE* values are compared to validate the model developed in this study. *MSPR* (mean of the squared prediction errors) stands for mean squared prediction error as shown in Eq. (3).

$$MSPR = \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{n} \tag{3}$$

'If the *MSPR* is fairly close to *MSE* (mean of the squared errors) based on the SEM fit to the model-building data se, then the *MSE* for the selected regression model is not seriously biased and gives an appropriate indication of the predictive ability of the model. (John Neter *et al.*, 1990)' *MSPR* of 2004 data is 52.3 and *MSE* of 2005 data is 57.3 so that the difference is fairly small. Hence, the final SEM can be used as confirmative model and predictable models.

6.2 Direct Effects

The estimated parameters of the B and the Γ matrices which represent all direct causal effects are summarized in Table 8 and 9. Most parameters among endogenous variables are significant at the p=0.05 level. Variables such as time zone, weather condition, incident location, horizontal alignment, vertical alignment and type of vehicles have significant effects on endogenous variables such as the number of deaths, injured persons, involved vehicles and incident clearance time. These results suggest that bad weather condition (snow and fog) leads to increase (γ = 0.046, t-value = 1.493) the incident clearance time as it leads to decrease the accident size (the number of death (γ = -0.211, t-value = -3.557). The slower operating speed in bad weather can explain the result (Choi *et al.* (1999).) The incident location have positive effects on both accident size and the incident clearance time (γ = 0.111, t-value = 4.574). If the accident occurs at main road or others (bridge and tunnel etc.) the location variable is coded as '1', and if the accident occurs at tollgate or ramp it is coded as '0'. Hence, we can interpret that accidents at tollgate or ramp reduce the incident clearance time and accident size. These relationships make sense because the locations are easy to access by the incident response team.

An interesting result is that the accident size is negatively correlated to the poor road section (sharp curve or steep upslope) because of its slower operating speeds, which is a similar finding of previous studies (Lee *et al.* (2008)). The type of involved vehicles variable has significant effect on both the number of involved vehicles (γ = 0.229, t-value = 5.209) and the incident clearance time (γ = 0.380, t-value = 16.542). As vehicle type is either truck or trailer, the number of involved vehicles and the incident clearance time increase. In general, accidents including truck or trailer require more clearance time as the descriptive statistics of data in Table 3.

In case of the endogenous variables, the estimated coefficients of the number of deaths (β = 0.156, t-value = 2.161) and the number of involved vehicles (β = 0.256, t-value = 4.713) show a positive relationship with the incident clearance time. This result is reasonable and consistent with the findings of previous researches (Nam and Mannering (2000)). The direct effect between the number of injured persons and the incident clearance time is not estimated because it does not have significant relationship in the initial SEM. In summary, the incident clearance time has significant relationship with the number of death, the number of involved vehicles and several exogenous variables (weather condition, incident location, horizontal/vertical alignment and type of involved vehicles).

6.3 Indirect and Total Effects

The total effect of one variable on another variable might be different from the direct effect of the first variable on the second if the first variable also affects other variables that in turn, directly or indirectly, affect the second variable. Since model estimation results include a set of equations with parameters representing interrelationships, the SEM approach allows us to examine not only direct effects but also indirect effects among variables. The indirect and total effects are as shown in Table 8 and 9, respectively.

In case of incident location variable, total effect (β = 0.565, t-value = 11.758) is the sum of direct effect (β = 0.111, t-value = 4.574) and indirect effect (β = 0.454, t-value = 8.509). In case of type of vehicles variable, total effect (β = 0.439, t-value = 19.472) is the sum of direct effect (β = 0.380, t-value = 16.52) and indirect effect (β = 0.059, t-value = 3.104).

Therefore, we can conclude that 1) accidents in main road need more incident clearance time because of poor accessibility and bigger accident size; 2) accident including truck or trailer need more incident clearance time because of its bigger accident size; and 3) incidents occurred in sharp curve (R<500m) or steep upslope (more than 3%) have less incident clearance time because accident size is smaller due to the lower operating speed.

The endogenous variables for ‘accident size’ have same amount of the total effects on incident clearance time as direct effects because no indirect effect is postulated. ‘Accident size’ variables such as the number of death, injured persons and involved vehicles variances only affected by exogenous variables do not have no direct and/or indirect effect of other endogenous variables. As previously stated, the incident clearance time has a positive relationship with the number of deaths (β = 0.156, t-value = 2.161) and the number of involved vehicles (β = 0.256, t-value = 4.713).

7. CONCLUSIONS

In this research, SEM for the incident clearance time is constructed and the interrelationships among various factors are explored. Exogenous variables in this model are peak/non-peak time, day of week, weather condition, day/night time, incident location, horizontal alignment, vertical alignment and type of vehicles involved. Endogenous variables include the number of deaths, the number of injured persons, the number of involved vehicles, and the incident clearance time. In our three-level causal model structure, we postulate that exogenous variables effect on 'accident size' variables, and both exogenous variables and 'accident size' variables effect on 'incident clearance time'. The SEM illustrates positive or negative effects of each variable on the incident clearance time. The SEM approach allows us to examine not only direct effects but also indirect effects among variables and the total effects of each variable are sum of the direct effect and indirect effect.

The results can be interpreted as 1) accidents in main road need more incident clearance time, 2) accident including truck or trailer need more incident clearance time, and 3) incidents occurred in sharp curve (R<500m) or steep upslope (more than 3%) have less incident clearance time. These relationships are explained by difficulty to access main roads. Incident response team can easily access to the incident location where either on tollgate or ramp than on the main roads. In addition, the incident clearance time decreases in tollgate/ramp and curve/slope because drivers tend to decrease operating speed so that accident size decreases in these sections. If incident vehicle type is either truck or trailer, the number of involved vehicles and the incident clearance time increase. In general, accidents including truck or trailer require more clearance time. The accident size is negatively correlated to the poor road section (sharp curve or steep upslope) because of its slower operating speeds, which is a similar finding of previous studies (Lee *et al.* (2008)).

The results in this study can help to estimate the incident clearance time of various incident types and to response the incident effectively. This study can be extended to various directions in the future. The most important thing is to search for more exogenous variables to express the relationships with the incident clearance time and the incident recovery time should be also considered in the model.

REFERENCES

Breiman, Leo, Jerome H. Friedman, Richard A. Olshen, and Charles J. Stone (1984) Classification and Regression Trees, *Pacific Grove*, CA: Wadsworth

Bollen, K. A. (1989) Structural equations with latent variable, *Wiley, New York*

Choi, J. S., and Bongsoo Son (1999) The Effect of Rain on Traffic Flows in Urban Freeway Basic Segments, **Journal of Korean Society of Transportation, Vol. 17, No. 1**, 29-39

Choi, Y. S., and Jin-Hyuk Chung (2003) Multilevel and multivariate structural equation models for activity participation and travel behavior, **Journal of Korean Society of Transportation, Vol. 21, No. 4**, 145-154

Chung, J. H., and Yongsung Ahn (2002) Structural equation models of day-to-day activity participation and travel behaviour in a developing country, **Transportation Research Record, No. 1807**, 109-118

Chung, J. H., and Dongkyu Lee (2002) Structural model of automobile demand in Korea, **Transportation Research Record, No. 1807**, 87-91

Fitzpatrick, K., and J. M. Collins (1999) Speed profile model for two-lane rural highway, **Transportation Research Record, No. 1737**, 7-15

Golob, T. F. (2000) A simultaneous model of household activity participation and trip chain generation, **Transportation Research Part B, Vol. 34**, 355-376

Golob, T. F. (1996) A model of an activity participation and travel interactions between household heads, **Transportation Research Part B, Vol. 31**, 177-194

Golob, T. F. (1987) A structural model of temporal change in multi-modal travel demand, **Transportation Research Part A**, 391-400

John Neter, William Wasserman, Michael H. Kutner (1990) Applied Linear Statistical Models, *Richard D. IRWIN, INC.*

Joreskog, K. G., and Sorbom, D.(2000) LISREL 8: User's Guide, Scientific Software International, Chicago, Ill.

Khattak, A., Schofer, J., Wang, M.-H. (1995) A Simple time sequential procedure for predicting freeway incident duration, **IVHS Journal. Vol. 2, No.2**, 1995, 113-138.

Kim, J. S., Kim, G. F. Ulfarsson, and L. Porrello (2007) Bicyclist injury severities in bicycle-motor vehicle accidents, **Accident Analysis and Prevention, Vol. 39**, 238-251

Kelvin K.W., Yau (2004) Risk factors affecting the severity of single vehicle traffic accidents in Hong Kong, **Accident Analysis and Prevention, Vol. 36**, 333-340

Lee, Jeom-Ho, Hong, Da-Hui and Lee, Su-Beom (2006) Development of Predicting Models of the Operating Speed Considering on Traffic Operation Characteristics and Road Alignment Factors In Express Highways, **Journal of Korean Society of Transportation, Vol. 24, No. 5**, 109-121

Lee, Ju-Yeon, Chung, Jin-Hyuk and Son, Bongsoo (2008) Analysis of Traffic Accident size for Korean Highway Using Structural Equation Models, **Accident Analysis & Prevention, Vol. 40, No. 6**, 1955-1963

Lu, X., and E. I. Pas (1999) Socio-demographics, activity participation and travel behavior, **Transportation Research Part A, Vol. 33**, 1-18

Nam, Doohee, and Mannering Fred (2000) An Exploratory Hazard-Based Analysis of Highway Incident Duration, **Transportation Research Part A, Vol. 34,** 85-102

Ozbay, Kaan, and Noyan, Nebahat (2006) Estimation of incident clearance times using Bayesian Networks approach, **Accident Analysis and Prevention, Vol. 38**, 542-555