A STUDY OF THE TREND ANALYSIS IN ROAD TRAFFIC ACCIDENT RATES IN KOREA

Hyeok-ku KWON
Graduate Student
Department of Transportation Engineering
Hanyang University
1271 Sa 1 dong, Ansan, Kyonggi,
425-791 Korea
Fax: +82-31-406-6290
E-mail: bivratio@ihanynag.ac.kr

Kyungwoo KANG
Professor
Department of Transportation Engineering
Hanyang University
1271 Sa 1 dong, Ansan, Kyonggi,
425-791 Korea
Fax: +82-31-406-6290
E-mail: kyungwoo@hanynag.ac.kr

Abstract: We must consider the some factors when the time series data such as road traffic accidents is used by conventional statistical methods. They make errors of estimates and tests using the existing regression methods, when the time series data is non-stationary. Although the non-correlated data is non-stationary, the regressions will show that they are correlated each other. As spurious regression, non-stationary variables have high R-square and t-value. Therefore, we use unit root tests to distinguish whether time series data is stationary or non-stationary.

This paper focuses on unit root test and cointegration using time series data, and comparison with conventional econometric analysis. We also try to compare with relationship between local and global road traffic accident rates, or if related, whether it is spurious relationship or not. In addition, we consider the global and local impacts of road traffic accident rates.

Key Words: road traffic accident, time series, unit root test, cointegration, error correction model

1. INTRODUCTION

In the industrialized countries, road traffic accident rates and fatality rates have tended to show the reverse trends. Some researchers interpret the downward trend as evidence of exponential learning(Oppe(1991)), while others treat it as a nuisance parameter that happens to be essential for model fitting(Peltzman(1975), Harvey and Durbin(1986), and Broughton(1991)).

There are many studies using time series data, such as road traffic accident but in modeling, the characteristics of time series data have not been considered. Namely, if considering, most of the analyses are the ARIMA(Autoregressive Integrated Moving Average) model using univariate time series data. Because there are many difficulties in applications of road traffic accident analysis, we have to consider other methods.

In this paper, we are concerned with the trend itself rather than the level of road safety. An intervention such as a seat-belt law or a speed limit can be expected to affect the level of the accident rates without necessarily affecting its trend. By contrast, technical change in road technology can be expected to affect the trend in road safety, and not just its level. We propose factors affecting road traffic accident trends and whether their global or local considerations. The former are common to all countries, and reflect the globalization of the automobile and road technology. The latter are country-specific and reflect the separate factors taken in individual countries. It may be, however, that some of these factors, such as seat-belt laws, are more global than local since they have been applied more or less universally.

In this paper, we use econometric techniques for handling non-stationary time series such as accident rates. In section 2, the theoretical framework of model in road traffic accidents are introduced. In section 3, unit root tests and cointegration are discussed. The data are presented in section 4 and results are discussed in section 5. In section 6, conclusions and further researches are followed.

2. THEORETICAL FRAMEWORK OF MODEL

2.1. Analysis of Existing Studies

Smeed(1949) used time series such as road traffic accidents for 20 european countries during 38 years.

Lim and Lee(1990) tried to analyze the future by using the traffic accident data from 1962 to 1989 in Korea. They regressed that dependent variables were accident rates, fatalities, and injuries, respectively, and that independent variable was registered vehicles. They also used regression as well as ARIMA and logistic model in order to develop the traffic accident model of Busan city. They assumed that trends of traffic accident would be supported by the future, so they predicted traffic accident to apply several simple methods.

Park(1995) used the reciprocal model between fatalities per vehicle and vehicles per capita in Korea, Japan, and Chungcheong-do by using Greenshields, Greenberg, Underwood, Multiplicative. But at the most, only a few referred to characteristics of time series, i.e., if it is used by conventional econometric techniques when it is a non-stationary variable of trends, this result may be wrong due to spurious relationships. For example, two trending time series, which in reality are unrelated, may be spuriously with high R^2. The high correlation can arise, either because both variables are related to time or because they are related to third variables that happen to be time-trended. The unit root test, which distinguishes between genuine and spurious correlation, is considered in this paper, and cointegration and error correction model have also apply.

2.2. Theoretical Framework

When the time series data are used by conventional statistical methods of road traffic accidents, some problems, such as autocorrelation, often happen. The theoretical framework on which the econometric analysis is based is drawn from Lave and Weber(1970) and Peltzman(1975), who argued that driving behavior is likely to be adaptive with respect to the legal and traffic environment.

The safety(S) experienced by the representative driver depends upon his own behavior(V) as well as that of other drivers($\overline{V}$). We assume that safety is determined as in Eq. (1)

$$S = S(\overset{(-)}{V}, \overset{(-)}{\overline{V}}, \overset{(+)}{C}, \overset{(+)}{R}) \tag{1}$$

where C and R denote vehicle and road quality, respectively, and where the signs of partial derivatives are indicated above the respective variables. The schedule cj in Fig. 1 plots the inverse relationship between S and V that is implied by Eq. (1).

Vehicle quality(C) is assumed to vary directly with income(W) since richer people can afford better cars, and with a technological parameter τ_c, i.e.:

$$C = \overset{(+)}{C}(\overset{(+)}{W}, \tau_c) \tag{2}$$

We assume that road quality(R) is set exogenously by the relevant authority.

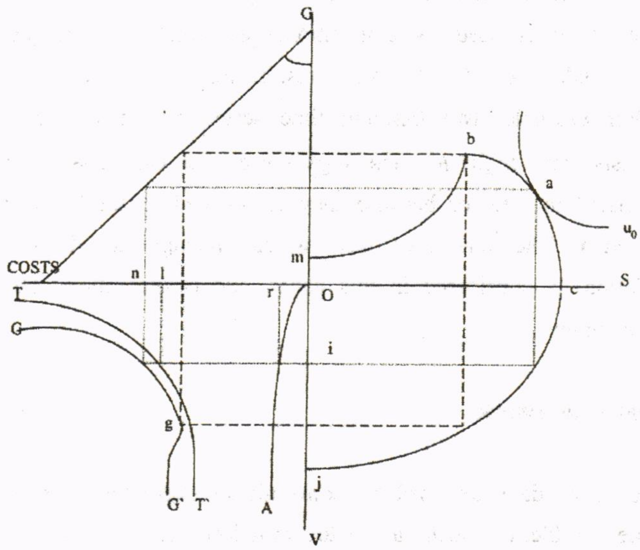

Fig. 1 Optimal driving behavior(micro)

Trip time cost is represented in Fig. 1 by the horizontal distance subtended by schedule TT', which has a natural lower asymptote. An increase in income(W) will shift this schedule to the left since time becomes more expensive.

There is a risk of punishment(π) by traffic police, which is increasing in V and policing P but it is decreasing $\overline{V}$. We denote the fine, if caught, by F, hence the expected value(A) of the fine due to apprehension may be written as;

$$A = F \cdot \pi(\overset{(+)}{P}, \overset{(+)}{V}, \overset{(-)}{\overline{V}}) \tag{3}$$

where we assume that the risk of apprehension increases in V. Schedule OA in Fig. 1

plots the relationship between V and the expected value of the fine that is implied by Eq. (3). OA will tend to rotate anti-clockwise as $\overline{V}$ increases and as enforcement decreases.

Schedule GG', which plots the horizontal sums subtended by schedules TT' and OA, measures the relationship between the total expected trip cost and V. It may have a minimum at g. The upper left quadrant of Fig. 1 simply converts travel costs into savings(G) due to increasing V.

Finally, schedule mbc plots the transformation frontier between safety and saving that is implied by schedules cj and GG'. The turning point at b corresponds to the turning point at g as indicated. The efficient part of the transformation frontier is of course bc.

The individual driver's utility function is assumed to be;

$$U = U(\overset{(+)}{S}, \overset{(+)}{G}) \tag{4}$$

Utility is maximized at a implying that the optimal driving behavior is i as indicated. This implies that in equilibrium apprehension risk is r, trip time cost is l, and total trip cost is n.

We now turn from the micro to the macro in which aggregate driving behavior is determined, i.e., where V and $\overline{V}$ are jointly determined.

If $\overline{V}$ increases, schedule cj contracts toward the origin thereby shifting schedule cb to the left. If the indifference curve is homothetic the equilibrium value of V will decrease, but safety and saving on trip costs are likely to decrease too. This implies that there is negative correlation between V and $\overline{V}$.

The principle of conjectural variation implies that if each driver optimizes his behavior in terms of the model in Fig. 1 given the behavior of other drivers, and if each driver thinks that his fellow drivers optimize in a similar way, then we may summarize the joint determination of V and $\overline{V}$ as in Fig. 2. Schedule Ⅰ is naturally flatter than schedule Ⅱ. The conjectural variation equilibrium occurs at the intersection of the two schedules.

In summary, the theoretical model implies that trending variables such as GDP per head have an ambiguous effect on road safety, even when preferences are homothetic.

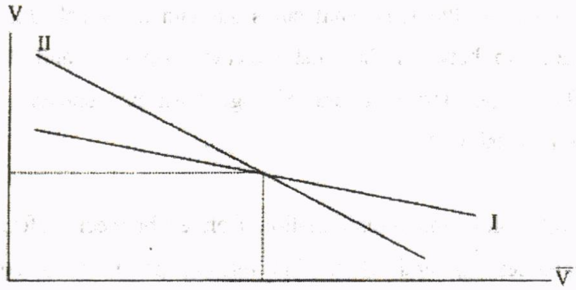

Fig. 2 Optimal driving behavior(macro)

The model implies that in equilibrium V depends on $\overline{V}, W, \tau_c, R, P$ etc. while $\overline{V}$ depends upon $\overline{W}$ etc. By the principle of conjectural variations V and $\overline{V}$ will be jointly determined implying that aggregate, or macro solution for S is;

$$S=S(V(W, \tau_c, R, P); C(W, \tau_c); R) \tag{5}$$

We may classify the variables in Eq. (5) into global and local variables. The latter include the indicators of police enforcement(P, F, π), income(W), seat-belt legislation etc. The former are embodied in the technological parameters τ_c and τ_R.

Empirical macro-representation of the variables that feature is Eq. (5) are GDP or consumption per capita(W), road capital per kilometer(R), and ratio of car and traffic police enforcement(P). In the absence of an obvious way of proxying vehicle quality (τ_c) we assume that this information is embodied in the accident rate abroad(S^*) since the latter depends, according to Eq. (5), upon W^*, τ_c^* etc. Because the vehicle is imported or exported in Korea, it is reasonable to assume that $\tau_c=\tau_c^*$, which implies that τ_c may be proxied by S^*, W^* etc. In practice we proxy it by S^* and the age of the fleet(A) since the latter is expected to vary inversely with W/W^*. But because the age of the fleet is recently surveyed in Korea, we only substitute τ_c to S^*.
In summary we estimate models of this type;

$$S_t = \alpha + \beta_1 W_t + \beta_2 R_t + \beta_3 P_t + \beta_4 S_t^* + \gamma t + u_t \tag{6}$$

where u denotes a stationary error term and t denotes time. If $\gamma=0$, there is no autonomous time trend in the accident rate; the time trend is entirely captured by the other variables in the model. Because the variables are non-stationary time series data when the they is regressed, some of problems are happened by regression assuming the normal distribution. Therefore, they need to apply other methods, so we refer to analyze unit root test and cointegration.

3. UNIT ROOT TESTS AND COINTEGRATION

3.1. Unit Root Tests

Since the variables featuring in Eq. (6) have time trends, they are non-stationary, and spurious correlation may exist between S and the right hand side variables. For example, there may be a high degree of correlation between the accident rate in Korea(S) and the accident rate abroad (S^*), when in fact these two variables are unrelated, i.e. the observed correlation is spurious. Engle and Granger(1987) were the first to argue that if u in Eq. (6) is stationary, the estimated model is not spurious. They referred to such variables as being co-integrated. Cointegration occurs when a weighted combination of two or more non-stationary variables creates a new variable (u) which happens to be stationary.

Here, we need to determine whether the variables are non-stationary and if so, the types of non-stationarity. If their first differences are stationary, we shall refer to this as difference stationarity(DS), and if their deviation from a fixed time trend are stationary we shall refer to this as trend stationarity(TS). The Dickey-Fuller(1981) test consists of running the following regression;

$$\Delta X_t = \alpha + \beta X_{t-1} + \gamma t + \sum_{i=1}^{p} \delta_i \Delta X_{t-i} + \varepsilon_t \tag{7}$$

if $\gamma=0$, the series is DS, otherwise it is TS. The augmentations that appear in Eq. (7) are specified to ensure that ε is iid(an independent and identically distributed error term). The numbers reported in table 2 are t-values for β in Eq. (7) and their critical values, as computed by Dickey and Fuller(DF, or ADF if Eq. (7) includes augmentation, i.e. $p>0$). If each variable is linearly combinated when it has unit root , this model is 'co-integrated' each other and the high correlation is not spurious.

3.2. Cointegration Test and Error Correction Model

We use unit root test whether regression residual is stationary or not. In addition, we also use the same unit root test method in cointegration, but it does not use the same statistics of unit root test. In cointegration model, we show as the following equation;

$$Y_t = \alpha + \beta X_t + \gamma t + \varepsilon_t \tag{8}$$

Eq. (8) is defined by whether intercept and linear trend are or not as unit root test where independent variables(X_t) are consisted of more than 2 time series. Here, Y_t and X_t show that orders are the same.

Cointegration test is consisted of the following unit root test on regression residuals;

$$\Delta\hat{\varepsilon}_t = \gamma\ \hat{\varepsilon}_{t-1} + \sum_{i=1}^{p} \delta_i \Delta\ \hat{\varepsilon}_{t-i} + u_t \tag{9}$$

When the cointegration relationship is established, it is a representation theorem that ECM(Error Correction Model) exists, which is developed by Engle and Granger. The general equation is the following;

$$\Delta Y_t = \alpha + \gamma\ \hat{\varepsilon}_{t-1} + \sum^{n} \delta_i \Delta X_{t-i} + \sum^{n} \theta_i \Delta Y_{t-i} + \upsilon_t \tag{10}$$

In Eq. (10), when $\hat{\varepsilon}_{t-1}$ is co-integrated, it is error correction term ($\hat{\varepsilon}_{t-1} = Y_{t-1} - \hat{\beta} X_{t-1}$, $\hat{\varepsilon}_{t-1} \sim I(0)$) reflecting error between X_{t-1} and Y_{t-1}. Therefore, what Y_t changes is reflected by unbalance of two variables as well as X_t's change. These error correction model can be found long-term equilibrium and short-term dynamic coordination. As a result, long term equilibrium can be estimated by cointegration regression, while short-term coordination can be considered by ECM. This model can be estimated by these methods(Enders, 1995).

4. DATA

Fig. 3 shows that traffic accidents per 10,000 vehicles in Korea have been reduced. In comparison with 1970 and 1998, traffic accident rates are dramatically reduced about 12.6 times. However, reduction of the accident rates is slightly undulating. Similarly, comparing with the other industrialized countries such as France, Germany, Japan, Norway, and Spain(S^*), Korean traffic accident rates sharply decrease in comparison with their reduction.

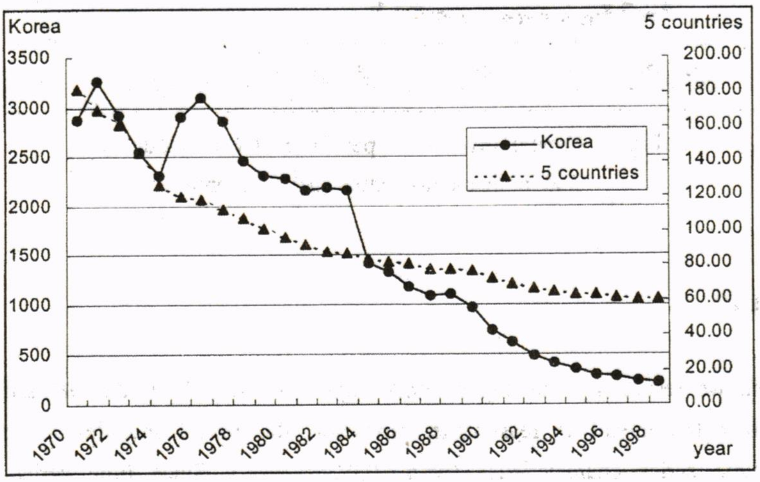

Fig. 3 Accidents per 10,000 vehicles in Korea and 5 countries
(5 countries : France, Germany, Japan, Norway, and Spain)

Nevertheless, the absolute accident rates are still high. But the opportunity reducing road traffic accidents will be high, too.

The capital stock invested in roads per kilometer of road (R: at constant 1995 prices) raises about 9 times between 1970 and 1998(Fig. 4) and includes constructions of highway, pavement, and maintenance based on government spending. The road investment is sharply increasing improvement of road quality and roads will be used next year as well as current, so accumulated road capital stock is used for analysis.

Police enforcement(P) is approximated by the ratio of the number of vehicles to police reports for offenses. This will tend to decrease, if the rate of growth of the vehicles keeps up with the rate of growth of police enforcement. That relates police enforcement to road **traffic accident reduction**

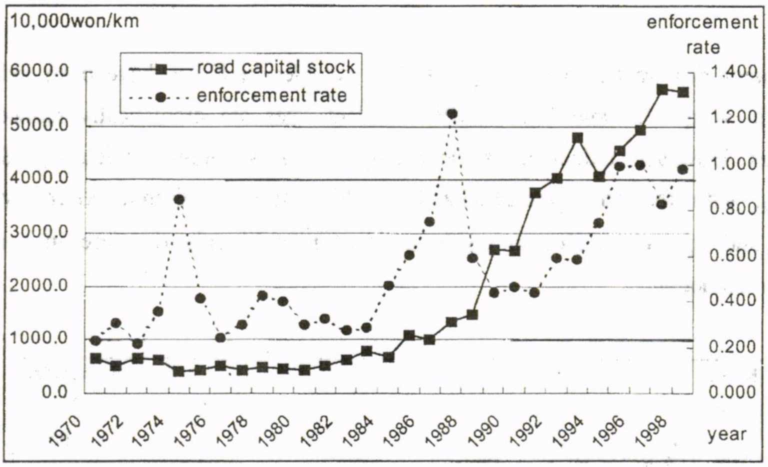

Fig. 4 Capital investment per kilometer of road(R) and police enforcement(P) in Korea

5. RESULTS

Table 1 shows results by the conventional regression. Based on the R^2 and t-values revealed that the model is reasonable. However, the model includes autocorrelation such as time series characteristics (with $D.W.:0.2565$). Since time series data are non- stationary, spurious regression may happen. Therefore, we must consider other methods and one of the methods is the unit root test.

Table 1. Conventional regression(local)

coefficient*		statistics
$\hat{\alpha}$	12.373(15.841)	$D.W.=0.2565$
$\widehat{\beta_2}$	-0.608(-8.120)	
$\widehat{\beta_3}$	-0.335(-1.912)	$R^2=0.88$

* $\ln S=\alpha+\beta_2\ln R+\beta_3\ln P$ (t-value in parenthesis)

Table 2 shows the unit root test referred in the section 3. The variables are stationary in first difference, which show the characteristics of time series data. In addition, all variables are stationary after first difference.

Table 2. Unit root test

	DF		ADF*	
	DS**	TS**	DS	TS
lnS	**-3.461**	**-3.545**	**-3.046**	-2.258
lnP	-1.736	**-3.763**	-1.864	**-3.546**
lnS*	**-7.531**	-1.003	**-3.195**	-2.188
lnR	**-4.447**	**-14.364**	-0.945	-2.142

* p=3 in ADF test

** in critical value of sig. level 5%, DS : -2.95, TS : -3.54

If we consider the two non-nested models which are co-integrated, these models are non-nested because neither model is a special case of its rival in global and local model.

We refer to these models as the global and local models because the former attaches importance to the international transmission of road safety technology via S^*, whereas the latter comprises of local variables. Results for the global and local models are presented in table 3 and table 4, respectively.

Table 3. Global impact

coefficient*		statistics**
$\hat{\alpha}$	-0.0822 (-2.074)	$DF = -4.553$, $ADF_1 = -3.557$
$\hat{\beta}_4$	0.9414 (0.850)	$D.W. = 2.1696$, $R^2 = 0.096$

* $\ln S = \alpha + \beta_4 \ln S^*$(t-value in parenthesis)

** critical value of DF and ADF : -3.55(in sig. level 5%)

In the global model(table 3), R^2 is not high enough as the conventional regression criteria. However, because the variables are non-stationary, the criteria of conventional regression are not appropriate. We show that cointegration test is not spurious in the relationship of accident reduction between in Korea and abroad. This shows that the

Korean accident reduction is explained by global safety trend(S^*). When considering the lack of explanatory variable, we need to add another variable in model. But explanatory power may slightly fall because of difficulty obtaining time series data.

Table 4. Local impact

coefficient*		statistics**
$\hat{\alpha}$	14.370 (26.968)	$DF = -6.405$, $ADF_1 = -3.428$
$\widehat{\beta_2}$	-0.8130 (-14.615)	
$\widehat{\beta_3}$	-0.2876 (-5.01)	$D.W. = 2.16$, $R^2 = 0.52$

* $\ln S = \alpha + \beta_2 \ln R + \beta_3 \ln P$(t-value in parenthesis)

** critical value of DF and ADF : -3.55(in sig. level 5%)

By contrast, in the local model(table 4) all the variables show characteristics of countrty; road quality(R) and police enforcement(P). Road traffic accident rates vary according to road quality and police enforcement. As a result, we find that each variable is cointegrated by Error Correction Model(ECM) proposed by Engle and Granger(1987).

6. CONCLUSIONS AND FURTHER RESEARCHES

When we consider conventional regression and related problems, we introduce the unit root test and cointegration. On the conventional econometric analysis not to consider characteristics of time series data, it can happen spurious regression. To correct these spurious problems, we consider ECM which introduces the unit root test and cointegration test in non-stationary time series data.

Introducing new approaches such as ECM is required time series data, so selection of variables slightly limits in difficulty obtaining data. As fleet age has surveyed recently, it is excluded. We need further research how fleet age and traffic accident are related. We widely focus on the road traffic accident rate. Because any variable may have different effect in relation to accident rate and death rate, further research needs the expended analysis of severity and death rate. For example, the seat-belt law does not reflect traffic accident rates but reduces its severity. Nevertheless, we do not consider since we only focus on traffic accident rates.

In addition, regional impact may be considered the factors of road traffic accidents in Korea. Therefore, panel data which add time series data and cross-sectional data, may be expended.

ACKNOWLEDGEMENTS

This research has been performed as a part of Advanced Highway Research Center Project funded by Korea Ministry of Science and Technology, Korea Science and Engineering Foundation.

REFERENCES

Banerjee, A.(1994) **Co-integration, Error Correction, and the Econometric Analysis of Non-stationary Data,** Oxford University Press, New York.

Dickey, D.A. And Fuller, W.A.(1981) Likelihood ratio statistics for autoregressive time series with unit roots, **Econometrica 49,** 1057-1972.

Enders, W.(1995) **Applied Econometric Time Series,** John Wiley, New York.

Engle, R. F. And Granger, C.W.J.(1987) Co-integration and error correction: Representation, estimation and testing, **Econometrica 55,** 251-276.

Green, W.H.(2000) **Econometric Analysis**, Prentice Hall, New Jersey.

Gregory, A.W., And Nason, J.M., Watt, D.G.(1996) Testing for structural breaks in co-integrated relationships, **Journal Of Econometrics 30,** 321-341.

Gregroy, A.W. And Hansen, B.E.(1996) Residual-based tests for co-Integration in models with regime shifts, **Journal Of Econometrics 30,** 99-126.

Griffiths, W.E., Hill, R.C., and. Judge, G.G.(1993) **Learning and Practicing Econometrics,** Wiley, New York.

Hatanaka, M.(1996) **Time-Series-Based Econometrics: Unit Roots and Co-integrations,** Oxford University Press, New York.

Harvey, A.C.(1990) **The Econometric Analysis of Time Series,** MIT Press.

Harvey, A.C. And Durbin, J.(1986) The effects of seat belt legislation on British road casualties: A case study in structural time series modelling, **Journal of The Royal Statical Society 149,** 187-210.

Lave, C.A.(1985) Speeding, coordination, and the 55 mph limit, **American Economic Review 75,** 1159-1164.

Lee, I. B. and Lim, H. Y.(1990) A study on development of forecasting model for traffic accident in Korea, **Journal of Transportation Research Society of Korea Vol. 8, No. 1,** 73-88.

Lim, H. Y. and Lee, I. B.(1992) Development of traffic accident forecasting model in Pusan, **Journal of Korea Transportation Society, Vol. 10, No. 3,** 103 - 121.

Loeb, P.D.(1987) The determinants of automobile fatalities with special consideration to policy variables, **Journal of Transportation Economics And Policy 21,** 279-287.

Michael, B. and Dalit, G.(2000) Globalization in road safety: Explaining the downward trend in road accident rates in a single country, **Accident Analysis and Prevention 32,** 71-84.

Mills, T.C.(1992) **Time Series Techniques for Economists,** Cambridge University Press.

Park, B. H.(1995) A study on development of forecasting model traffic accident in Chungchong region, **Journal of Korea Transportation Research Society Vol. 13, No. 1,** 63-82.

Peltzman, S.(1975) The effects of automobile safety regulation, **Journal of Political Economy 83,** 677-725.