패널자료를 이용한 가로구간 교통사고분석
- 청주시 간선도로를 사례로 -

김준용 · 나 희* · 박병호†*
한국건설기술연구원 · *충북대학교 도시공학과
(2011. 8. 19. 접수 / 2012. 5. 2. 채택)

Traffic Accident Analysis of Link Sections Using Panel Data in the Case of Cheongju Arterial Roads

Jun-Young Kim · Hee Na* · Byung-Ho Park†*
Korea Institute of Construction Technomogy · *Department of Urban Engineering, Chungbuk National University
(Received August 19, 2011 / Accepted May 2, 2012)

Abstract : This study deals with the accident model using panel data which are composed of time series data of 2005 through 2007 and cross sectional data of link sections in Cheongju. Panel data are repeatedly collected over time from the same sample. The purpose of the study is to develop the traffic accident model using the above panel data. In pursuing the above, this study gives particular attentions to deriving the optimal models among various models including TSCSREG (Time Series Cross Section Regression). The main results are as follows. First, 8 panel data models which explained the various effects of accidents were developed. Second, R^2 values of fixed effect models were analyzed to be higher than those of random effect models. Finally, such the variables as the sum of the number of crosswalk on intersections and sum of the number of intersections were analyzed to be positive to the accidents.
Key Words : panel data, arterial link, traffic accident, EPDO (Equivalent Property Damage Only), arterial roads

1. 서 론

1.1. 연구 배경 및 목적

2010년 현재 도로교통공단에서 제공하고 있는 OECD회원국 교통사고 비교통계를 살펴보면, 회원국 평균 자동차 1만대당 인적피해 교통사고발생건수가 60.1건으로 나타나고 있으나, 우리나라의 경우 105.8건으로 회원국 중 1위를 기록하고 있는 것으로 분석되고 있다. 또한 자동차 1만 대당 교통사고 사망자수의 경우 OECD회원국 평균이 1.3명인데 반해 우리나라는 2.8명으로 나타나고 있어 두 배가 넘는 수치를 보이고 있다. 이는 우리나라의 교통사고 발생이 심각한 수준임을 의미하며, 앞으로 교통선진국으로 발돋움하기 위해서는 교통사고를 발생시키는 다양한 요인들에 대하여 연구해야 한다. 교통사고에 대한 연구논문이 매년 발간되고는 있지만 일반적으로 단기간 교통사고 자료를 이용하거나 사고가 주로 발생하는 특정장소를 연구대상으로 선정하는 경우가 많다. 그러나 이러한 연구들은 매년 변화하고 있는 다양한 도로 기하구조 및 교통시설 등에 대한 부분을 충분히 고려할 수 없다는 단점이 있다.

따라서 이 연구는 교통사고가 계속 증가하고 있는 청주시를 대상으로 시기적으로는 다소 차이가 나지만 다년간의 교통사고 자료 및 기하구조 자료를 이용하여 패널자료를 구축하고, 패널분석 방법을 이용하여 보다 현실적이고 실증적인 패널교통사고 모형을 개발하는 것이 목적이다.

패널분석은 연도별로 변화되는 자료의 특성을 모두 반영할 수 있는 교통사고모형을 개발할 수 있기 때문에 기존의 사고모형보다 사고원인에 대한 많은 정보를 알아볼 수 있는 방법이다.

이를 위해 이 연구에서는 청주시를 대상으로 간선도로의 특성 및 기하구조 실태를 조사하고, 이를 통하여 청주시 가로구간에 대한 교통사고와 간선도로의 특성 및 기하구조의 영향관계를 규명하고자 한다.

† To whom correspondence should be addressed.
bhpark@chungbuk.ac.kr

1.2. 연구 범위 및 내용

이 연구에서는 도로의 기하구조 특성이 교통사고에 미치는 영향을 분석하기 위하여 공간적 범위를 청주시 간선도로로 선정하였으며, 시간적 범위는 2005년부터 2007년까지 3년간의 청주시 가로구간 24개소의 기하구조 및 교통특성을 패널자료의 형태로 구축하였다.

내용적 범위로는 연구대상지 교통사고의 발생정도가 어떻게 변화해 왔는지를 규명하기 위해서 Auto CAD 및 실제 현장조사를 통해 패널자료를 수집하였으며, 사고지표로는 일반적으로 많이 쓰이는 교통사고건수와 EPDO(Equivalent Property Damage Only: 대물피해 환산계수) 계수를 이용하였다.

이러한 3년간 24개소의 도로구간 자료를 통합한 패널자료를 이용하여 TSCSREG(time series cross section regression)의 One-Way Error Component Regression Model과 Two-Way Error Component Regression Model, 각각의 고정효과모형(FEM : fixed effect model)과 임의효과모형(REM : random effect model)의 4가지 모형으로 각각 분석하고 3년간 청주시 교통사고 발생에 영향을 미친 요인을 파악한다.

2. 연구의 절차

2.1. 교통사고모형

국내에서 패널자료를 이용한 교통사고모형의 개발에 관한 논문은 찾을 수 없었기 때문에 이 연구에서는 일반적으로 많이 사용되어 온 다중선형회귀분석, 포아송 및 음이항 모형을 이용한 교통사고모형을 개발한 논문을 사례로 검토하였다.

박정순 등(2007)[1]은 4지 신호교차로를 대상으로 도로환경 요인과 교통사고와의 상관관계를 분석하고 사고추정모형을 개발하였다. 다중회귀모형의 추정결과 ADT, 주도로의 평균차로폭원, 횡단보도 수, 종단경사, 제한속도차이 등이 교통사고건수에 많은 영향을 미치는 것으로 분석되었으며, 교통사고의 98% 이상이 어느 한 가지 요인에 의해 발생하는 것이 아니라 인적요인, 차량요인 및 도로환경요인 중 두 가지 이상 요인들의 복합적인 작용에 의해서 발생하는 것으로 분석되었다.

유두선 등(2008)[2]은 청주 · 청원의 신호교차로를 대상으로 사고모형을 개발하고 지방부의 교통여건을 반영하는 사고유발요인을 파악하고, 지방부 사고의 특성 및 사고에 영향을 미치는 특정요인을 밝혀냈다. 그 결과 다중선형회귀모형과 포아송 및

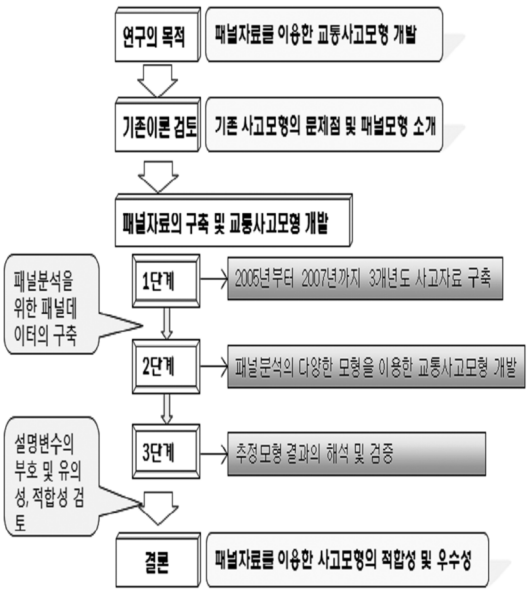

Fig. 1. Procedure to perform studies,

음이항 회귀모형을 통한 12개의 모형에 횡단보도 합계가 공통적으로 채택되어, 지방부의 교통사고에 있어 가장 영향력이 있는 변수로 분석되었다.

한상진 등(2007)[3]은 도로교통안전관리공단에서 제공하고 있는 교통사고 통계분석 자료를 이용하여 도로종류별 교통사고 추세모형을 시계열 분석기법으로 구축하였으며, 10만대 · km당 교통사고 발생건수, 사망자수 및 부상자수에서 일반국도가 가장 위험하며, 고속국도는 상대적으로 안전한 것으로 분석하였다. 또한 그들은 시계열분석모형이 가지는 한계점에 대하여 언급하였다.

김경환 등(2010)[4]은 청주시 주간선 및 보조간선도로 24개소를 대상으로 2007년의 교통사고 자료와 가로구간 특성자료를 이용하여 다중선형회귀모형, 음이항 및 포아송 회귀모형을 이용하여 구축하고 비교 분석하였다. 그 결과 차로 수와 진출입구수는 양의 관계 그리고 굴곡점 수는 음의 관계를 갖는 것으로 나타났다.

2.2. 패널모형

외국의 경우 1900년대 중반 이후부터 패널자료를 구축하여 연구를 진행하고 있으며, 국내에서는 1990년대부터 연구기관을 중심으로 패널조사를 실시하여 사회 및 경제와 관련된 다양한 논문들이 발표되고 있다. 그러나 교통사고와 관련된 패널모형은 발견하기 어려워 패널모형을 이용한 타 분야의 논문들을

고찰하였다.

최충익(2004)[5]은 '패널모형에 의한 도시지역 수해결정요인 분석'이라는 논문에서 경기도의 31개 시군을 대상으로 각각의 인문지리적 환경자료와 1971년부터 2002년까지 32년간의 시계열자료를 실증적으로 분석하였다. 독립변수로 수해의 취약성을 증가시키는 불투수면 분석을 위하여 물리적 요소에는 도시적 토지이용면적, 자연적 요소에는 연평균강수량, 하천면적 등, 인문적 요소에는 인구밀도, 그리고 방재 요소에는 지역의 제방면적, 임야면적 등을 선택하여 도시지역 수해피해를 극대화시키는 요소가 무엇인지 분석하였다.

최열 등(2007)[6]은 도시지역에서 인구의 분포, 토지이용 특성, 도로 및 교통, 각종 산업 활동 등이 대기오염에 미치는 영향을 분석하기 위해 패널자료를 활용하였다. 그들은 서울특별시와 6개 광역시를 포함한 인구 20만 이상 22개 대도시를 공간적 범위로 선정하고 1989년부터 2003년까지 15년간의 대기오염지표 및 도시특성요소 등을 시계열자료로 구축하였다. 즉, 그들은 15년간 22개 도시 자료를

Table 1. List of variables

Number	Variables	Symbol	Definition
1	사고건수	Y_1	연도별 사고건수
2	EPDO	Y_2	EPDO
1	교차로 횡단보도	X_1	가로구간 내 교차로 횡단보도 개수
2	단일로 횡단보도	X_2	가로구간 내 단일로 횡단보도 개수
3	평균도로폭	X_3	가로구간별 평균도로폭(m)
4	진출입구수	X_4	가로구간별 진출입구 개수
5	총 교차로수	X_5	가로구간별 총 교차로 수
6	버스정류장수	X_6	가로구간별 버스정류장 수
7	교통섬수	X_7	가로구간별 교통섬 개수

주 : EPDO(대물피해 환산계수) : 사망사고, 부상사고, 대물사고 각 피해의 종류를 하나의 피해단위로 환산한 것으로 식은 다음과 같다.
[EPDO = 12*사망사고+3*부상사고+대물사고]

Table 2. Correlation of variables

variables	Y_1	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_1	1	-	-	-	-	-	-	-
X_1	.575** (.003)	1	-	-	-	-	-	-
X_2	.848** (.000)	.437* (.033)	1	-	-	-	-	-
X_3	-.053 (.806)	-.063 (.770)	.045 (.836)	1	-	-	-	-
X_4	.675** (.000)	.619** (.001)	.529** (.008)	-.353 (.091)	1	-	-	-
X_5	.137 (.523)	.174 (.416)	-.130 (.546)	-.622** (.001)	.427* (.038)	1	-	-
X_6	.794** (.000)	.551** (.005)	.609** (.002)	.019 (.929)	.550** (.005)	.226 (.288)	1	-
X_7	.841** (.000)	.715** (.000)	.849** (.000)	.074 (.730)	.673** (.000)	.027 (.901)	.706** (.000)	1
variables	Y_2	X_1	X_2	X_3	X_4	X_5	X_6	X_7
Y_2	1	-	-	-	-	-	-	-
X_1	.590** (.002)	1	-	-	-	-	-	-
X_2	.846** (.000)	.437* (.033)	1	-	-	-	-	-
X_3	-.046 (.830)	-.063 (.770)	.045 (.836)	1	-	-	-	-
X_4	.699** (.000)	.619** (.001)	.529** (.008)	-.353 (.091)	1	-	-	-
X_5	.112 (.601)	.174 (.416)	-.130 (.546)	-.622** (.001)	.427* (.033)	1	-	-
X_6	.791** (.000)	.551** (.005)	.609** (.002)	.019 (.929)	.550** (.005)	.226 (.288)	1	-
X_7	.867** (.000)	.715** (.000)	.849** (.000)	.074 (.730)	.673** (.000)	.027 (.901)	.706** (.000)	1

주 : 2007년 기준 상관관계 결과임
**. 상관계수는 0.01수준(양쪽)에서 유의함
*. 상관계수는 0.05수준(양쪽)에서 유의함

통합한 패널자료를 이용하여 다양한 패널모형 중 최적의 모형을 찾고 정책적인 시사점을 도출하고자 하였다.

3. 분석 틀의 설정

3.1. 변수의 구성 및 설명

도로의 기하구조 요소가 교통사고에 미치는 영향을 패널자료 분석을 위하여 3년간의 시계열 자료와 청주시의 가로구간 총 24개소의 연도별 횡단면 자료를 이용하여 패널자료를 구축하였다. 일반적으로 교통사고에 가장 많은 영향을 미치는 일평균교통량의 경우 가로구간 전체에 대한 자료를 취득하기란 불가능하여 독립변수에 반영하지 못하였으며 이 연구에 최종 선정된 종속변수 및 독립변수는 다음 Table 1과 같다.

분석에 앞서 변수의 상관관계를 분석해 본 결과 다음과 같이 사고건수(Y_1)는 평균도로 폭과 총 교차로수를 제외한 모든 변수와 상관관계가 높은 것으로 나타났고 EPDO(Y_2)와의 상관관계도 같은 결과로 나타났다.

3.2. 연도별 사고 현황

청주시에서 2005년부터 2007년까지 발생한 3년간의 전체 가로구간의 총 사고건수는 다음 표 2와 같이 2005년 1,950건에서 2007년에는 다소 줄어든 1,895건으로 나타났다.

연구의 대상지인 간선도로에서 발생한 사고건수는 연도별로 2005년 717건에서 2007년 393건으로 점차 줄어들고 있다. 청주시는 주간선도로 13개소와 보조간선도로 11개소로 총 24개소의 주요도로가 있다.

Fig. 2. Arterial roads list of cheongju.

Table 3. Number of accidents

Classification	2005	2006	2007
가로구간 총 사고건수	1,950	1,786	1,895
연구 대상지 사고건수	717	609	393

3.3. 분석방법론 설정

패널모형이란 패널자료를 이용한 계량경제 분석으로 시계열 분석과 횡단면 분석을 동시에 수행하는 분석모형을 의미한다. 이것은 다양한 자료의 유형 중에서 가장 정보가 많고 유용하여 연구자들이 가장 선호하는 데이터 형태라고 볼 수 있다(최충익, 2004)[5].

패널자료 분석은 시계열 과정에서 발생하는 추정오차와 지역 또는 장소별 단위의 자료에서 발생하는 추정오차를 통제할 수 있는 장점을 가지고 있기 때문에 횡단면 또는 시계열 각각의 자료에 비해 현실을 보다 제대로 분석할 수 있는 장점을 가지고 있다(Baltagi, 2005)[7]. 패널모형의 추정 식은 오차항의 구조가 일반 회귀추정식과 다르며, 추정식은 다음과 같다.

$$Y\text{it} = \alpha + \beta X\text{it} + \cdots + \epsilon\text{it}$$

$$\epsilon\text{it}(\text{오차항}) = \text{ui} + \lambda\text{t} + \nu\text{it}$$

여기서, i(장소): 1, 2, 3, ..., N, t(시간): 1, 2, 3, ..., T,

μi : 관찰되지 않은 지역(장소)특성 효과

λt : 관찰되지 않은 시간효과

νit : 확률적 교란항

패널모형은 오차항의 고려방식에 따라 One-Way Error Component Regression Mode(OWECRM : 일방오차요소회귀모형)과 Two-Way Error Component Regression Model(TWECRM : 양방오차요소회귀모형)로 나누어지며, 오차항의 가정에 따라 Fixed Effect Model(FEM : 고정효과모형), Random Effect Model (REM : 확률효과모형)로 나누어진다(Baltagi, 2005).

두 모형을 구분하는 가장 탁월한 방법은 시간불변의 개별특성효과가 독립변수들과 관련이 되어 있는 지로 살펴볼 수 있으며, 관련이 있는 경우에는 Fixed Effect Model을 사용하고, 관련이 없을 경우에는 Random Effect Model을 선택한다.

4. 패널 분석

패널분석을 위한 프로그램에서는 일반적으로 SAS와 STATA를 많이 쓰며, 본 연구에서는 SAS 9.2 PROC

TSCSREG[8] 툴을 이용하여 연구를 수행하였습니다.

일방 및 양방 오차요소회귀모형의 각각에 대한 FEM과 REM의 4가지 모형을 이용하여 Y_1(사고건수)과 Y_2(EPDO)로 나누어 총 8개의 모형을 추정하였다.

4.1. 사고건수 모형

Table 4의 사고건수 추정모형(OWECRM)의 경우 FEM과 REM의 변수들의 부호가 상이하게 나타났으나, 공통적으로 교차로의 횡단보도수와 총 교차로의 수가 정(+)의 효과를 보이며, 교통섬 수는 부(-)의 효과를 보이는 것으로 분석되었다. 이는 일반적으로 교통사고의 경우 차량의 상충이 많은 지점에서 발생하는 경향이 있기 때문으로 여겨지며, 교통섬 수는 차량의 상충을 줄이는 역할을 하기 때문으로 여겨진다.

Table 4. Model estimates the number of traffic accidents (OWECRM)

변수	Time Series Cross Section Regression			
	One-way fixed effects		One-way random effects	
	Estimates	t-value	Estimates	t-value
Intercept	371.396	3.15	-67.126	-1.24
교차로 횡단보도수	3.783	1.00	0.955674	1.37
단일로 횡단보도수	-	-	18.741	2.31
평균도로폭	-	-	16.260	1.20
진출입구수	-3.353	-1.34	0.311	0.59
총교차로수	0.216	1.39	0.151	1.08
버스정류장수	-8.680	-2.35	0.540	1.06
교통섬수	-14.288	-1.63	-3.060	-0.87
R2	0.8516		0.3466	

Table 5. Model estimates the number of traffic accidents (TWECRM)

변수	Time Series Cross Section Regression			
	Two-way fixed effects		Two-way random effects	
	Estimates	t-value	Estimates	t-value
Intercept	330.813	2.16	-67.356	-1.37
교차로 횡단보도수	4.531	1.19	0.837	1.32
단일로 횡단보도수	-	-	16.249	2.19
평균도로폭	-	-	15.709	1.27
진출입구수	-2.743	-1.01	0.337	0.70
총교차로수	0.182	1.14	0.141	1.05
버스정류장수	-9.064	-2.37	0.574	1.24
교통섬수	-13.572	-1.57	-2.072	-0.64
R2	0.8627		0.3952	

Table 5의 TWECRM의 경우에도 FEM과 REM의 변수부호가 상이하게 나타났으나, 공통적으로 교차로 횡단보도수와 총 교차로 수는 정(+)의 효과, 교통섬 수는 부(-)의 효과가 나타나는 것으로 분석되었다.

4.2. EPDO 모형

EPDO 계수를 종속변수로 한 모형에 있어서 사고건수를 종속변수로 한 모형과 비슷한 결과가 도출되었다. 변수에 있어서도 공통적인 정(+)의 변수는 교차로 횡단보도수와 총 교차로 수, 부(-)의 변수에는 교통섬 수로 분석되었다. 그러한 결과는 다음 Table 6와 Table 7에 나타나 있다.

Table 6. Estimation model EPDO (OWECRM)

변수	Time Series Cross Section Regression			
	One-way fixed effects		One-way random effects	
	Estimates	t-value	Estimates	t-value
Intercept	1062.82	0.007	-219.065	-1.40
교차로 횡단보도수	10.306	0.396	2.604	1.29
단일로 횡단보도수	-	-	47.431	2.01
평균도로폭	-	-	53.123	1.35
진출입구수	-8.712	0.282	0.873	0.57
총교차로수	0.670	0.184	0.4	0.94
버스정류장수	-26.202	0.031	1.473	1.00
교통섬수	-37.773	0.183	-6.377	-0.62
R2	0.8228		0.3399	

Table 7. Estimation model EPDO (TWECRM)

변수	Time Series Cross Section Regression			
	Two-way fixed effects		Two-way random effects	
	Estimates	t-value	Estimates	t-value
Intercept	735.803	1.54	-221.05	-1.60
교차로 횡단보도수	14.562	1.23	2.231	1.26
단일로 횡단보도수	-	-	39.001	1.87
평균도로폭	-	-	51.467	1.49
진출입구수	-4.440	-0.52	0.971	0.72
총교차로수	0.607	1.22	0.364	0.93
버스정류장수	-25.726	-2.16	1.574	1.22
교통섬수	-34.037	-1.26	-3.075	-0.34
R2	0.8439		0.4061	

5. 결 론

패널자료를 활용하여 청주시 간선도로를 대상

으로 한 가로구간의 사고모형을 개발한 이 연구의 주요 결과는 다음과 같다.

첫째, 가로구간에서 발생하는 교통사고의 다양한 요인들을 설명하는 통계적으로 의미 있는 8개의 패널모형이 개발되었다. 즉, 사고건수와 EPDO, 일방오차모형과 이방오차모형 그리고 FEM 및 REM의 총 8개의 모형이 개발되었다.

둘째, 고정효과모형의 R^2 값이 임의효과모형에 비하여 다소 높은 것으로 분석되었다. 마지막으로 교차로의 횡단보도수와 총 교차로수가 사고에 정(+)의 효과, 그리고 교통섬 수는 부(-)의 효과를 보이는 것으로 분석되었다. 앞서 사례연구의 청주시 가로구간 교통사고 모형의 결과에서는 차로수와 진출입구수가 많을수록 굴곡점수가 적을수록 교통사고가 많이 발생할 것으로 예측결과를 내놓고 있지만 본 연구는 다년간 자료를 이용한 패널분석으로 2007년 1개년도만을 분석한 기존 논문과는 다른 결과가 도출되어 다년간 사고자료와 기하구조 자료를 확보하는 것이 매우 중요하다는 것을 알 수 있다.

본 연구는 공간적 전이성, 시간적인 변화 까지 모두 반영할 수 있는 패널교통사고모형의 초기단계로서 모형에 대한 세부적인 차이나 검증통계량에 대한 해석이 부족한 것은 사실이다. 하지만 지금까지 시도되지 않았던 패널자료를 활용했다는 데에 의의가 있어 유의성이 없는 변수라 하더라도 일단 다 넣어서 해석을 시도했다. 향후 연구과제로는 우선 모형구축에 필요한 패널자료를 보다 많이(최소 10년) 확보하여 패널분석에서 제시하고 있는 모형으로 분석을 시도하고, 아울러 기하구조 측면뿐만 아니라, 인적요인, 환경적 요인 등을 포함한 다양한 해석과 검증이 필요할 것으로 판단된다. 특히 교통량의 경우 경찰청, 지자체 등에서 신호운영을 위해 조사한 자료를 활용할 수 있는 방안의 모색이 필요하다.

참고문헌

1) 박정순, 김태영, 유두선 “도로환경요인과 교통사고의 상관분석 및 사고추정모형 개발(청주시 4지 신호교차로를 중심으로)”, 대한교통학회지, 2007.
2) 유두선 “청주 · 청원 지방부 신호교차로의 특성을 감안한 사고모형”, 충북대학교 대학원 석사학위 논문, 2008.
3) 한상진, 김근정, “도로종류별 교통사고 추세분석 및 시계열 분석모형 개발”, 한국도로학회 논문집, 2007.
4) 김경환, 박병호, “청주시 가로구간의 교통사고모형 개발”, 대한국토 · 도시계획학회, 국토계획, 제45권, 제5호, 2010.
5) 최충익 “패널모형에 의한 도시지역 수해결정요인 분석”. 대한국토 · 도시계획학회, 『국토계획』 제39권 제7호, 2004.
6) 최열, 문설희, 임하경, “도시특성 요소가 대기오염에 미치는 영향에 관한 연구”, 대한국토··도시계획학회, 국토계획, 제42권 제3호, 2007.
7) Baltagi, B. H, Economic Analysis of Panel Data, Third Edition, Wiley, 2005.
8) SAS Institute Inc., SAS/ETS 9.1 User's Guide, Cary, NC : SAS Institute Inc, 2004.