Check for updates

A Development of Trip Generation Model by Functions of Small City Based on Mobile Origin-destination Data

Kim, Kyuhyuk* ⓘ · Won, Minsu** ⓘ · Song, Taijin*** ⓘ

Abstract

It is crucial to develop a trip generation model based on urban function to estimate the optimal travel demand for multiple cities selected as the study area. However, the survey-based data's low-resolution spatial-temporal trip information impedes the identification of trip patterns by function in smaller cities. In this study, higher-resolution mobile Origin-Destination data were utilized to develop trip-generation models. The study sites were categorized based on the functions of four cities: tourism, industry, farming, and fishing villages. Various independent variables such as transportation infrastructure, land use, and socioeconomic factors, were employed in this research. A quasi-Poisson regression analysis was conducted to formulate a trip generation model. Consequently, this study revealed that the principal variables of the trip generation models varied according to the function of each city.

Keywords   Mobile Phone Data, Trip Generation Model, Quasi-Poisson Regression Analysis, Traffic Analysis Zone
주제어   모바일 데이터, 통행발생모형, 준포아송 회귀분석, 교통존

Ⅰ. Introduction

1. Research Background and Purpose

Traffic demand estimation is a process that simplifies descriptive models to recognize changes in current travel behaviors and socioeconomic conditions, and predicts future traffic demands. This process is divided into four stages: trip generation, trip distribution, mode split, and network assignment (Aboelenen et al., 2021). Among these, the first stage, trip generation estimation, provides data necessary for estimating the remaining stages (trip distribution, mode split, network assignment) (Etu and Oyedepo, 2018; Mukherjee and Kadali, 2022).

Accurately estimating the trip generation is crucial because the estimated final traffic demand values in the subsequent stages can significantly vary based on the estimated trip generation. However, the traditional traffic demand models have been typically developed using survey-based data such as household travel survey data. This data suffers from low response rates and reporting errors, and due to the high cost of conducting surveys, small collected sample size can result in biased estimates, thereby showing low reliability (Bwambale et al., 2021).

Recently, with the emergence of big data collected on a per-second basis, it has become possible to identify individuals' trip trajectories, and numerous studies have been conducted using high spatial and temporal resolution

* Ph.D. Candidate, Department of Urban Engineering, Chungbuk National University (First Author: kyuhyuk@chungbuk.ac.kr)
** Associate Research Fellow, Korea Transport Institute (minsuwon@koti.re.kr)
*** Associate Professor, Department of Urban Engineering, Chungbuk National University (Corresponding Author: tj@chungbuk.ac.kr)

personal travel data to explore human mobility. The primary data sources used in studies on human mobility, such as navigation and mobile phone data, cover all navigation users and mobile phone subscribers, thereby enabling a comprehensive description of travel behaviors and resulting in highly reliable analysis outcomes.

Lately, there has been an emerging trend in research aimed at developing trip generation models using a type of mobile phone data known as Call Detailed Record (CDR) data. CDR data is recorded daily as long as a mobile phone is used, offering the advantage of higher temporal resolution compared to household travel survey data collected over a single day. However, the data lacks information on variables such as household details, and the data are only recorded when the mobile phone is used, resulting in spatial and temporal discontinuities within the data. This makes it difficult to capture movements when the mobile phone is not in use, which is a notable drawback (Bwambale et al., 2021).

This study utilized sightings data, which is recorded at regular intervals regardless of mobile phone usage. Such data is processed into OD (Origin-Destination) form between cell tower-based TAZs (Traffic Analysis Zones), which has higher spatial resolution compared to the conventional TAZs at town, township or neighborhood levels. Using this data, we developed a trip generation model according to the functional characteristics of small cities.

At present, the most widely used KTDB data for estimating future traffic demand applies the same trip generation model uniformly across the whole country or by major metropolitan areas. However, cities are composed of different elements, and especially in small cities, trips can be generated by specific functional elements such as tourism, industry, agriculture, and transport-inducing facilities. Consequently, to enhance the reliability of future traffic demand estimates, there is a need to develop city function-specific trip generation models, particularly focusing on small cities. In this study, to develop city function-specific trip generation models, various variables not considered in current trip generation models were explored through a literature review on trip generation model development. In addition, actual travel data and various variables from cities with specific functions like tourism, industrial, agricultural, and fishing towns were considered.

‖. Theory and Literature Review

1. Past research on trip generation model development

Chang et al. (2014) utilized household travel survey data from the Seoul metropolitan area to compare the performance of various trip generation models. The model fit of different analytical models including linear regression, Tobit, Poisson, ordered logit, and category-type was compared. Variables such as household composition, household income, transportation infrastructure, demographic characteristics, regional economy, and land use were utilized in the development of the trip generation model. The analysis showed that the category-type based trip generation model exhibited the best performance, while the linear regression-based trip generation model showed an acceptable level of performance.

Usanga et al. (2020) developed a trip generation model for residential areas. The four variables were household size, household income, car ownership, and the number of employed persons per household. The results indicated that household size was the strongest factor influencing trip generation in residential areas.

Etu and Oyedepo (2018) applied both a Radial Basis Function Neural Network (RBFNN) and a multiple linear regression model to household travel survey data from the Akure region in Nigeria, comparing the fit of the two models. Independent variables used included the number of household members, number of employed household members, number of students in the household, number of household members older than age 15, and number of driver's license holders in the household. The analysis revealed that the deep learning technique, RBFNN, had a better fit compared to the multiple linear regression model.

Sulistyono et al. (2012) estimated a household-based trip generation model for residential areas in Jember. The independent variables included the number of family members, family members working, family members attending school, income, and vehicle ownership.

Roorda et al. (2010) developed a trip generation model for specific demographic groups such as low-income households, the elderly, and single-parent families. The independent variables used were age, income level, number of

household members, vehicle ownership, number of employed persons, and population density. An ordered probit model was applied to derive the model. The analysis showed that elderly travel was significantly influenced by vehicle ownership, while travel in single-parent families was affected by both vehicle ownership and employment-related indicators. Conversely, travel among low-income households was less influenced by vehicle ownership and more significantly impacted by transportation infrastructure.

Aboelenen et al. (2021) identified factors affecting trip generation in residential buildings. The independent variables included household size, income level per housing unit, vehicle ownership, nationality (Qataris, non-Qataris), number of persons with a driving license, and number of employed persons and students per housing unit. A multiple linear regression model was applied as the analytical technique.

Lafta and Ismael (2022) developed a trip generation model for the Baghdad region using various statistical approaches. The independent variables consisted of gender, number of workers, population by age group, income, type of dwelling unit, ownership of dwelling unit, area of dwelling unit, and car ownership. Artificial Neural Networks (ANN) and multiple linear regression models were used as analysis techniques. The analysis indicated that trips in the Baghdad region were associated with factors such as family size, gender, number of students in the household, number of workers in the household, and car ownership, with the ANN model performing better than the multiple linear regression model.

Masoumi (2022) developed city-specific trip generation models for cities in the Middle East and North Africa region. The independent variables included gender, age, individual driving license ownership, activity as a worker or student, car ownership, number of driving licenses in the household, household income, commute mode choice, commuting distance, node-link ratio, link length, and number of accessed facilities. The results of the analysis of variance showed that the factors influencing trip generation varied from city to city. The author of the study recommended the active use of land use variables for sustainable demand estimation.

Yang et al. (2020) developed a trip generation model for the Nanjing region in China using location-based social network data and mobile phone data. The independent variables included the number of POIs, number of public transportation stations, and number of highway entrances. The analysis showed that land use-related variables had a significant impact on trip generation.

2. Research on Trip generation model development using mobile phone data

bwambale et al. (2019) developed a methodology to enhance the accuracy of trip generation estimation using CDR data, socioeconomic indicators, and GSM (Global System for Mobile Communications) data. The traffic volumes derived from the CDR data and socioeconomic information were utilized to estimate future trip generation amounts, while the travel proportions derived from GSM data were used to construct trip generation models for different socioeconomic groups.

Bwambale et al. (2021) developed a methodology to improve the accuracy of trip generation estimation by integrating OD profiles that can be extracted from CDR data with census data and household travel survey data. The model from the study showed an overall improvement in the accuracy of trip generation estimation. However, the study was limited in that the variables used in the analysis were confined to basic socioeconomic indicators such as gender and age.

Colak et al. (2015) presented a methodology that uses CDR data and population density data alone to determine the purpose of trips made by mobile phone users. They classified trip generation areas based on the frequency and timing of mobile phone trips into categories such as work and leisure, and assigned purposes to trips based on the characteristics of the selected trip generation areas.

Shi and Zhu (2019) identified the residential and employment areas of users based on mobile phone data traffic and spatial distribution. They developed Traffic Analysis Zones (TAZs) based on cell tower locations and developed models for trip generation and inflow in individual TAZs through a multiple regression analysis.

3. Research distinctiveness

First, a trip generation model for small cities based on their functions was developed. Household travel survey data typically collects travel and socioeconomic information at the eup-myeon-dong level, which is suitable for deriving trip generation models for large cities, but inadequate for rural and fishing areas in small cities due to a lack of TAZ samples. In this study, trip generation models specific to small cities were derived using cell tower-based TAZs, which have higher spatial resolution compared to the eup-myeon-dong level TAZs.

Second, a PA-based trip generation model was developed through type estimation of origins and destinations in mobile origin-destination data. Since traditional mobile phone data does not include information on the purpose of travel, only OD-based trip generation models could be derived. However, a PA-based model better represents travel patterns from a behavioral perspective compared to an OD-based model and has less aggregation error, making it preferable for estimation (Kim, 1997). This study estimated the types of origins and destinations based on mobile phone users' travel behaviors to derive a PA-based trip generation model.

Third, a trip generation model was derived considering various variables including transportation infrastructure and land use. Traditional trip generation models used in policy-making are based on household travel survey data, which tends to focus predominantly on socioeconomic indicators. In this research, variables not considered in existing trip generation models but used in related studies, such as public transportation stops, node links, and land use variables, were utilized as independent variables.

Ⅲ. Analysis Process

1. Research Flow and Spatial Scope

The flow of this study is as follows. First, the spatial scope of the study was determined based on population size. Second, the concept and pre-processing of the mobile origin-destination data used in this research were explained. Third, independent variables were selected considering those used in the current KTDB and various previous stud-

ies for developing trip generation models. Fourth, analytical methods were chosen based on the characteristics of trip production and attraction in the TAZ-specific mobile origin-destination data. Fifth, key variables for the development of the trip generation model were identified through a correlation analysis, and a regression analysis was applied to develop the trip generation model.

For the development of city function-specific trip generation models, four cities were selected as spatial scopes, each differing in city functions, as shown in ⟨Figure 1⟩. It was assumed that large and medium-sized cities, having complex city functions, would present limitations in constructing function-specific models. Therefore, this study focused on developing city function-specific trip generation models centered on small cities. The demographic definition of small cities was based on Yim (2019), defining them as cities with a population of less than 200,000.

First, the city selected as a tourism city is Andong in Gyeongsangbuk-do. As of 2022, the population of Andong was approximately 150,000, and it has recently been selected as one of the five major tourism hubs by the Ministry of Culture, Sports and Tourism. Second, the city selected as an industrial city is Eumseong-gun in Chungcheongbuk-do. As of 2022, the population of Eumseong-gun was about 100,000, and it currently leads the industry in the central region with over 330 companies located in 16 industrial complexes. Third, the city selected as a rural city is Euis-
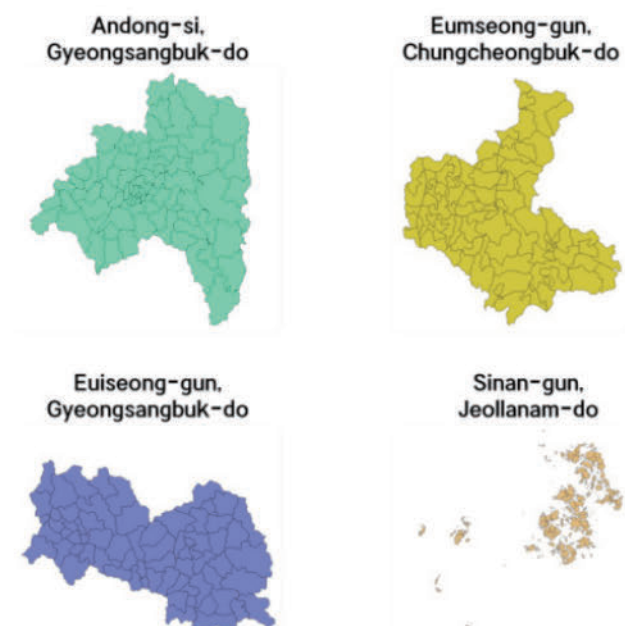

Andong-si,
Gyeongsangbuk-do

Eumseong-gun,
Chungcheongbuk-do

Euiseong-gun,
Gyeongsangbuk-do

Sinan-gun,
Jeollanam-do

Figure 1. Study sites (polygon unit)

eong-gun in Gyeongsangbuk-do. As of 2022, the population of Euiseong-gun was about 50,000, and more than 30% of the total population is engaged in agriculture. Fourth, the city selected as a fishing city is Sinan-gun in Jeollanam-do. As of 2022, the population of Sinan-gun was approximately 40,000, and it is known as a representative fishing city in Korea, based on its various fishing infrastructures.

2. Mobile origin-destination data

In this study, the mobile origin-destination data used for developing the trip generation model consists of data processed from the mobile phone trajectory data of KT subscribers into the form of origin-destination traffic volumes. It includes information such as origin and destination TAZ codes, time of departure and arrival, types of locations (residential or activity areas) of the origin and destination TAZs, gender, age, and traffic volumes.

The processing of mobile origin-destination data was carried out as follows (Kim et al., 2021). First, error data such as NULL values were removed and corrected from the mobile phone trajectory data. Second, movements and stays were distinguished based on a duration of stay. Third, to differentiate between origins and destinations, individual trips were sorted and spatially organized based on the date of data recording, start time of stay, and end time of stay. Finally, areas considered residential or potential activity locations were used as origins and destinations, and traffic volumes between TAZs were aggregated. The final aggregated mobile origin-destination data has the structure shown in 〈Table 1〉 (Kim and Song, 2022).

The methodology for estimating types of stay locations to extract PA-based trip production and attraction from mobile origin-destination data is as follows. Initially, if an individual traveler's monthly trajectory data shows characteristics of staying for more than three hours between 21:00 and 07:00 on at least three days a week and three weeks a month, then that TAZ is defined as a home (H) or non-home stay area (N). Next, if an individual traveler's monthly trajectory data shows characteristics of staying for more than three hours between 09:00 and 18:00 at least two days a week and two weeks a month, then that TAZ is defined as a workplace (C) or school (S). If the traveler is aged over 20, the location is defined as a workplace; if under

Table 1. The dataset of mobile phone OD data (Kim and Song, 2022)

Column name	Column type	Notation	Example
Origin code	String	xxxxxx	110001
Origin date	Numeric	yyyymmdd	20170101
Origin time	Numeric	00~23	12
Origin place	String	S, C, H, N, X, R	S
Destination code	String	xxxxxx	110002
Destination date	Numeric	yyyymmdd	20170101
Destination time	Numeric	00~23	13
Destination place	String	S, C, H, N, X, R	C
Age	Numeric	00~110	30
Gender	String	F, M	M
# of trip	Numeric		3

25, as a school. For those aged between 20 and 25, if these criteria are met during the vacation months of January-February and July-August, it is defined as a workplace; if not met, it is defined as a school. All other stay locations are defined as potential activity areas. Within these, TAZs where the frequency of stay occurs three or more times a week are defined as other (X) areas, and TAZs without any characteristic stay patterns are defined as religious activity areas (R).

Five types of PA travel were derived based on the types of origins and destinations. First, if either the origin or destination is a home (H) or a company (C), it is defined as household-based work commuting. Second, if either the origin or destination is a home (H) or a school (S), it is defined as household-based school commuting. Third, if either the origin or destination is a home (H) or other (X), it is defined as household-based other travel. Fourth, if either the origin or destination is a company (C) or other (X), it is defined as non-household-based work travel. Fifth, if both the origin and destination are other (X), it is defined as non-household-based other travel. Descriptive statistics for each PA travel type by city function are as shown in 〈Table 2〉.

In addition, cell tower-based TAZs were used instead of the traditional eup-myeon-dong level TAZs for mobile origin-destination data. Cell tower-based TAZs are created by applying a Voronoi diagram to the main cell towers and then removing and correcting error data (Kim et al., 2021).

Table 2. Descriptive statistics of mobile phone data (Unit: trip)

City	PA	Production			Attraction		
		Sum	Mean	S.D.	Sum	Mean	S.D.
Andong	H↔C	1,643,027	20,538	33,937	1,643,027	20,538	33,570
	H↔S	936,664	11,708	21,765	936,664	11,708	21,801
	H↔X	4,127,399	51,592	91,102	4,127,399	51,592	90,540
	C↔X	1,898,722	23,734	41,244	1,898,722	23,734	42,644
	X↔X	1,054,993	13,187	23,389	1,054,993	13,187	24,078
Eumseong	H↔C	1,328,491	16,006	19,913	1,328,491	16,006	19,773
	H↔S	358,500	4,319	7,286	358,500	4,319	7,316
	H↔X	1,860,166	22,412	29,228	1,860,166	22,412	28,936
	C↔X	1,121,348	13,510	14,897	1,121,348	13,510	15,414
	X↔X	495,052	5,964	6,746	495,052	5,964	6,863
Euiseong	H↔C	475,114	6,987	11,733	475,114	6,987	11,679
	H↔S	62,461	919	2,753	62,461	919	2,772
	H↔X	803,647	11,818	22,766	803,647	11,818	22,671
	C↔X	375,090	5,516	10,345	375,090	5,516	10,605
	X↔X	192,986	2,838	4,919	192,986	2,838	5,059
Sinan	H↔C	381,321	5,084	4,007	381,321	5,084	4,014
	H↔S	38,255	510	762	38,255	510	762
	H↔X	480,610	6,408	4,315	480,610	6,408	4,211
	C↔X	236,647	3,155	2,402	236,647	3,155	2,265
	X↔X	152,033	2,027	1,887	152,033	2,027	1,718

Compared to eup-myeon-dong level TAZs, cell tower-based TAZs offer four to six times higher spatial resolution and have the advantage of being able to match socioeconomic indicator attributes specific to cell tower units due to the matching process undergone with the cell towers.

3. Construction of variables

In the trip generation stage, a mathematical model is adopted to associate the population and socioeconomic characteristics aggregated in the TAZ, such as population, households, employment numbers, and income, with the trip production and attraction of individual TAZs (Chang et al., 2014). Traditional trip generation model development that uses survey-based data typically includes demographic characteristics such as the number of households, household members, and employees, making it straightforward to aggregate independent variables. However, mobile origin-destination data only includes personal characteristics such as age and gender. Therefore, the developed methodologies only consider such personal characteristics or combine socioeconomic indicators with mobile phone origin-destination data.

In this study, the origin-destination data between TAZs utilizes cell tower-based TAZs as the spatial units, which can be matched with targeted Si-gun-gu or eup-myeon-dong level data. This setup has the advantage of easily matching attribute information with cell tower-based TAZs if there is existing statistical data at the targeted Si-gun-gu or Eup-myeon-dong level. For the development of the trip generation model in this study, demographic variables related to population (Lafta and Ismael, 2022; Naser, 2020), socioeconomic indicators such as household size (Lafta and Ismael, 2022; Masoumi, 2022; Aboelenen et al., 2021; Usanga et al., 2020; Yang et al., 2020; Naser, 2020), transportation infrastructure variables including the number of bus stops, subway stations, and node-link density (Yang et al., 2020), and land use variables (Lee and Choo, 2021) were used as independent variables. The independent variables selected for this study are as shown in 〈Table 3〉.

Table 3. Selected Variables of trip generation models

Variables	Description	
Transport infrastructure	# of bus stop	
	# of node	
	sum of link length	m
Socio-demography	Population(~20 years)	
	Population(20~29 years)	
	Population(30~59 years)	
	Population(60~ years)	
	# of total household member	
	# of mean household member	
	# of 1st industry workers	
	# of 2nd industry workers	
	# of 3nd industry workers	
Land use	Residence	m²
	Commerce	m²
	Culture	m²
	Education	m²
	Business	m²
	Industry	m²
	Agriculture	m²

4. Selection of analytical technique

For developing a trip generation model, the most commonly used analytical technique is linear regression analysis, such as multiple regression analysis (Takyi, 1990). However, the dependent variable in this study, trip production and attraction from mobile origin-destination data, does not meet the assumptions of normality and exhibits overdispersion, making a linear regression analysis unsuitable. Accordingly, this study has chosen the quasi-Poisson regression model as the analytical technique.

The quasi-Poisson regression model is one of the generalized regression techniques, and it is calculated as shown in Equation (1). Here, μ represents the mean of the dependent variable, x_j denotes individual independent variables, and β_j represents the regression coefficients. The quasi-Poisson regression model is more appropriate than the Poisson regression model for deriving regression models of overdispersed data (Gabriella et al., 2019).

$$\ln(\mu) = x^T\beta = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \qquad (1)$$

Generalized regression models such as the quasi-Poisson regression model typically do not use the commonly used Adjusted R-Square for assessing model fit. Instead, values such as McFadden's R-Square are used as the primary measure (McFadden, 1973). McFadden's R-Square is derived as a scalar value between 0 and 1, where values closer to 1 indicate a better fit of the model. McFadden's R-Square is calculated as shown in Equation (2), by subtracting from 1 the ratio of the log likelihood of the fitted model to the log likelihood of the null model.

$$1 - (\log likelihood_{model} / \log likelihood_{NULL}) \qquad (2)$$

IV. Analysis Results

1. Identification of key variables

In this section, prior to developing city-specific trip generation models for different population sizes using mobile origin-destination data, a Pearson correlation coefficient analysis was conducted. The key variables selected for the trip generation model analysis had a significance level of $p < 0.05$ in correlation with the dependent variables, trip production and attraction. The results of the correlation analysis are presented in Tables 1, 2, 3, and 4 in the Appendix.

For the tourism city of Andong, significant correlations were found between PA-specific traffic volumes and several variables, including the sum of link length, number of total household members, number of mean household members, number of second industry workers, number of third industry workers, population under 20 years old, population in their 20s, population in their 30s to 50s, population over 60 years old, residential area, commercial area, cultural area, business area, and agricultural area. The variable for the number of nodes showed a positive correlation with non-household-based traffic volumes.

For the industrial city of Eumseong, significant correlations were found between PA-specific traffic volumes and variables such as number of total household members, number of second industry workers, number of third industry workers, population under 20 years old, population in their 20s, population in their 30s to 50s, population over 60 years old, residential area size, commercial area size, cultural area size, educational area size, and business area size. The

number of mean household members showed a positive correlation with household-based commuting and other commuting trip volumes. Additionally, agricultural area size had a positive correlation with household-based other commuting traffic volumes. Household-based school commuting had significant correlations with only two variables: the number of second industry workers and educational area size.

For the rural city of Euiseong, significant correlations were found between PA-specific traffic volumes and variables such as number of total household members, number of mean household members, number of first industry workers, number of second industry workers, number of third industry workers, population under 20 years old, population in their 20s, population in their 30s to 50s, population over 60 years old, residential area size, commercial area size, cultural area size, educational area size, and business area size. In the case of industrial area size, there was a positive correlation with non-household-based other traffic volumes.

For the fishing city of Sinan, significant correlations were found between PA-specific traffic volumes and variables such as the number of total household members, number of second industry workers, number of third industry workers, population under 20 years old, population in their 20s, population in their 30s to 50s, population over 60 years old, and commercial area size. The variables showing significant correlations varied slightly depending on the PA purpose.

2. Development of trip generation models by city function

in this section, a quasi-Poisson regression analysis is applied to derive the trip generation model, focusing only on variables that showed a significance level of less than 0.05 in the Pearson correlation analysis. To address the issue of multicollinearity among the independent variables, the Variance Inflation Factor (VIF) values for all variables within the trip generation model were adjusted to be less than 5.

First, the results of developing PA-specific trip generation models for the tourism city Andong are shown in ⟨Table 4⟩ and ⟨Table 5⟩. For household-based commuting, the production model included variables such as the number of mean household members, population in their 20s, cultural

Table 4. The selected Andong-si trip production model

Trip purpose	Model	R^2
H ↔ C	5.891+1.430×# of mean household member+1.813E-03×Population (20~29 years)-7.377E-05×Culture+ 1.507E-04×Business	0.837
H ↔ S	4.595+1.942×# of mean household member+0.001×# of 3nd industry workers	0.491
H ↔ X	8.341-1.955E-10×Sum of link length+0.787×# of mean household member+1.404E-03×Population (20~29 years)+1.463E-05×Commerce	0.848
C ↔ X	8.428+7.708E-02×# of node-3.242E-10×Sum of link length+0.396×# of mean household member+3.605E-04×# of 3nd industry workers+1.015E-03*Population(20~29 years)	0.821
X ↔ X	8.472+5.771E-04×# of 3nd industry workers+1.068E-03×Population (20~29 years)+3.040E-05×Culture	0.809

Table 5. The selected Andong-si trip attraction model

Example	Model	R^2
H ↔ C	6.171+1.299×# of mean household member+1.239E-03×Population (20~29 years)+1.420E-05×Commerce	0.814
H ↔ S	4.618+1.930×# of mean household member+0.001×# of 3nd industry workers	0.490
H ↔ X	8.318-2.002E-10×Sum of link length+0.805×# of mean household member+1.394E-03×Population (20~29 years)+1.451E-05×Commerce	0.849
C ↔ X	8.256+8.024E-02×# of node-3.530E-10*Sum of link length+0.467×# of mean household member+3.403E-04×# of 3nd industry workers+ 1.077E-03×Population(20~29 years)	0.838
X ↔ X	7.769+7.489E-02×# of node-3.092E-10×Sum of link length+0.391×# of mean household member+3.446E-04×# of 3nd industry workers+1.157E-03×Population (20~29 years)	0.844

area size, and business area size, while the attraction model included the number of mean household members, population in their 20s, and commercial area size. For household-based school commuting, the production model included variables such as the number of mean household members and the number of third industry workers. The attraction model had the same variables.

For household-based other commuting, the production model included variables such as the sum of link length, number of mean household members, population in their 20s, and commercial area size. The attraction model had the same variables. For non-household-based work commuting, the production model included variables such as the number of nodes, sum of link length, number of mean household members, number of third industry workers, and population in their 20s. The same variables were used in the attraction model. For non-household-based other commuting, the production model included variables such as the number of third industry workers, population in their 20s, and cultural area size, while the attraction model included the number of nodes, sum of link length, number of mean household members, number of third industry workers, and population in their 20s.

Second, the results of developing PA-specific trip generation models for the industrial city Eumseong are shown in ⟨Table 6⟩ and ⟨Table 7⟩. For household-based commuting, the production model included variables such as the number of mean household members, number of secondary industry workers, population over 60 years old, and educational area size, while the attraction model included the number of mean household members, number of secondary industry workers, cultural area size, and educational area size. For household-based school commuting, the production model included variables such as population in their 20s, and the attraction model had the same variables. For household-based other commuting, the produc-

Table 6. The selected Eumseong-gun trip production model

Trip purpose	Model	R^2
H↔C	6.391+1.076×# of mean household member+8.549E-04×Population (60~ years)+7.659E-05×Education	0.525
H↔S	8.040+0.002×Population(20~29 years)	0.106
H↔X	7.277+0.877×# of mean household member+4.455E-04×# of 2nd industry workers+1.339E-03×Population (60~ years)+6.083E-05×Education-1.663e-05×Agriculture	0.586
C↔X	8.712+4.907E-04×# of 2nd industry workers+6.312E-04×Population (30~59 years)	0.568
X↔X	7.937+0.0004×# of 2nd industry workers+0.0007E+Population (30~59 years)	0.544

Table 7. The selected Eumseong-gun trip attraction model

Trip purpose	Model	R^2
H↔C	6.639+1.010×# of mean household member+4.485E-04×# of 2nd industry workers+7.277E-05×Culture+9.011E-05*Education	0.518
H↔S	8.045+0.002×Population (20~29 years)	0.103
H↔X	7.298+0.868×# of mean household member+4.416E-04×# of 2nd industry workers+1.354E-03×Population (60~ years)+5.984E-05×Education-1.638E-05×Agriculture	0.587
C↔X	8.514+5.788E-04×# of 2nd industry workers+1.210E-03×Population (60~ years)+6.045E-05×Education	0.473
X↔X	8.126+4.554E-04×# of 2nd industry workers+6.214E-04×Population (30~59 years)-1.634E-05×Agriculture	0.593

tion model included variables such as the number of mean household members, number of secondary industry workers, population over 60 years old, educational area size, and agricultural area size. The attraction model had the same variables. For non-household-based work commuting, the production model included variables such as number of secondary industry workers and population in their 30s to 50s, while the attraction model included number of secondary industry workers, population over 60 years old, and educational area size. For non-household-based other commuting, the production model included variables such as number of secondary industry workers and population in their 30s to 50s, while the attraction model included number of secondary industry workers, population in their 30s to 50s, and agricultural area size.

Third, the results of developing PA-specific trip generation models for the rural city Euiseong are shown in ⟨Table 8⟩ and ⟨Table 9⟩. For household-based commuting, the production model included variables such as the number of mean household members and residential area size, while the attraction model included the number of first industry workers and cultural area size. For household-based school commuting, the production model included variables such as the number of mean household members, number of first industry workers, and cultural area size. The attraction model used the same variables. For household-based other commuting, the production model also included variables such as the number of mean household members, number

Table 8. The selected Euiseong-gun trip production model

Trip purpose	Model	R^2
H↔C	1.935+3.125×# of mean household member+Residence	0.675
H↔S	3.760×# of mean household member+2.080E-02×# of 1st industry workers+1.425E-04×Culture	0.705
H↔X	2.607*# of mean household member+1.688E-02×# of 1st industry workers+1.232E-04×Culture	0.666
C↔X	4.148×# of mean household member+0.020×# of 1st industry workers	0.508
X↔X	2.177×# of mean household member+1.645E-02×# of 1st industry workers+1.257E-04×Culture+ 2.634E-05×Industry	0.653

Table 9. The selected Euiseong-gun trip production model

Trip purpose	Model	R^2
H↔C	8.415+1.994E-02×# of 1st industry workers+1.884E-04×Culture	0.575
H↔S	3.658×# of mean household member+2.114E-02×# of 1st industry workers+1.443E-04×Culture	0.692
H↔X	2.554×# of mean household member+1.674E-02×# of 1st industry workers+1.249E-04×Culture	0.671
C↔X	4.295×# of mean household member+0.020×# of 1st industry workers	0.529
X↔X	2.308×# of mean household member+1.603E-02×# of 1st industry workers+1.258E-04×Culture+ 2.626E-05×Industry	0.674

of first industry workers, and cultural area size. The attraction model had the same variables. For non-household-based work commuting, the production model included variables such as the number of mean household members and number of first industry workers. The attraction model also used these variables. For non-household-based other commuting, the production model included variables such as the number of mean household members, number of first industry workers, cultural area size, and industrial area size. The attraction model had the same variables.

Fourth, the results of developing PA-specific trip generation models for the fishing city Sinan are shown in ⟨Table 10⟩ and ⟨Table 11⟩. For household-based commut-

ing, the production model included variables such as the number of bus stops and population under 20 years old. The attraction model used the same variables. For household-based school commuting, the production model included variables such as the number of third industry workers, population under 20 years old, and cultural area size, while the attraction model included the number of third industry workers, number of first industry workers, and cultural area size. For other household-based commuting, the production model included variables such as the number of bus stops, number of third industry workers, population under 20 years old, and residential area size, while the attraction model included the number of bus stops, population under 20 years old, population in their 20s, and residential area size. For non-household-based work commuting, both the production and attraction models included variables such as population in their 20s. Similarly, for non-household-based other commuting, both the

Table 10. The selected Sinan-gun trip production model

Trip purpose	Model	R^2
H↔C	7.624+0.027×# of bus stop+0.011×Population(~20 years)	0.403
H↔S	5.011+0.003×# of 3nd industry workers+0.011×Population(~20 years)- 0.0004×Culture	0.377
H↔X	8.121+4.190E-02×# of bus stop+1.362E-03×# of 3nd industry workers+9.350E-03×Population(~20 years)-2.916E-05×Residence	0.430
C↔X	7.429+0.021×Population(~20 years)	0.247
X↔X	7.060+0.019×Population(~20 years)	0.154

Table 11. The selected Sinan-gun trip attraction model

Trip purpose	Model	R^2
H↔C	7.627+0.027×# of bus stop+0.011×Population(~20 years)	0.399
H↔S	5.013+0.002×# of 3nd industry workers+0.011×# of 1st industry workers+0.0004×Culture	0.372
H↔X	8.047+3.471E-02×# of bus stop+ 7.195E-03×Population(~20 years)+ 1.156E-02×Population(20~29 years)- 2.665E-05×Residence	0.416
C↔X	7.459+0.020×Population(~20 years)	0.239
X↔X	7.108+0.018×Population(~20 years)	0.141

production and attraction models included variables including population in their 20s.

3. Summary of analysis results

in the case of tourism cities, the main variables influencing traffic were identified as the number of mean household members, number of third industry workers, population in their 20s, and the sizes of cultural and commercial areas. For industrial cities, the primary variables included the number of mean household members, number of secondary industry workers, population over 30 years old, and educational area size. In rural cities, the key variables were the number of mean household members, number of first industry workers, and cultural area size. In fishing cities, the population under 30 years old was the only notable variable influencing travel.

The analysis results suggest that the variables influencing trip production and attraction vary depending on the function of the city. When conducting economic feasibility studies for road and railway projects, the area under influence is defined for estimating traffic demand, and this area usually covers multiple cities and counties. If the models developed in this study, which take into account city functions, are applied, more reliable trip generation values can be obtained.

However, in the case of tourism and industrial cities, the explanatory power of the models for household-based school commuting was found to be very low, between 0.1 and 0.2, compared to other models. When the explanatory power of a model is low, its utility for actual demand estimation is reduced, indicating a need to develop additional factors that can enhance the explanatory power.

Moreover, rural and fishing cities generally exhibited lower explanatory power in their models compared to other types of cities. This result may be due to the populations of rural and fishing areas being predominantly composed of older adults, leading to less active trip production and attraction. Indeed, the analysis of data from this study showed that mobile phone traffic participation among middle-aged and older populations in Sinan and Euiseong was higher than in Andong and Eumseong, as depicted in ⟨Figure 2⟩. Conversely, the proportion of mobile phone traffic among the younger population, specifically those under
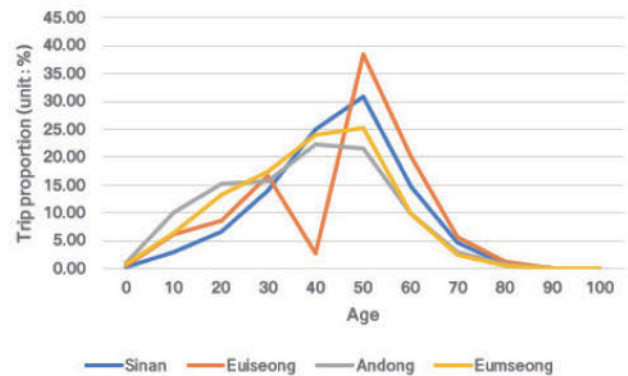

Figure 2. Trip proportion by age by city

30, was higher in Andong and Eumseong.

In particular, the explanatory power of non-household-based commuting models in fishing cities was very low, below 0.3, suggesting that research on developing trip generation models for such cities should be conducted separately. Additionally, PA-based models, which reflect characteristics of both origin and destination TAZs, have shown challenges in incorporating land use or transportation infrastructure characteristics (Song et al., 2011). Therefore, to develop more refined PA-based trip generation models, it will be necessary to focus on incorporating a variety of socio-economic indicators and demographic characteristics.

V. Conclusion

The aim of this study is to develop PA-specific trip generation models for small cities using mobile origin-destination data that contains actual traffic information. Based on a literature review, the variables for developing the trip generation model were categorized into three groups: transportation infrastructure-related variables, population and socio-economic variables, and land use-related variables. Due to the violation of the normality assumption and the presence of overdispersion in city-specific trip production and attraction data, this study employed the quasi-Poisson regression model, a type of generalized regression model analysis technique, to develop trip generation models. To prevent issues of multicollinearity among variables that may arise during analysis, all variables with a VIF of 5 or above were removed before performing the analysis.

This study moves beyond traditional survey-based trip generation model development, utilizing big data that

records actual travel and commuting. This approach allows us to develop trip generation models that reflect real-world behaviors. The research holds academic significance as it confirms conventional wisdom, such as rural cities being significantly influenced by the number of first industry workers, industrial cities by the number of secondary industry workers, and tourist cities by the number of third industry workers. It also discovered that the variables composing trip generation models differ according to city functions. There is policy significance in that some of the models developed in this study can be applied to actual demand forecasting. However, future research should consider and address the following points for further development.

First, despite increasing the spatial resolution of the analysis, relatively low model fit was observed in county-level units with populations less than 100,000. In this study, cell tower-based TAZs were used to maximize the spatial resolution of TAZs in the analysis; however, in smaller cities like Sinan, where the number of TAZs in the city is around 100, there are limitations in deriving reliable trip generation models. Therefore, for future development of trip generation models for small cities, it might be beneficial to consider using spatial units with even higher spatial resolution.

Second, while this study developed trip generation models based on primary urban functions such as tourism, industry, rural, and fishing, it is somewhat premature to assert that the developed models represent all cities with these functions. However, the results of this study demonstrate that the factors influencing traffic vary depending on the city's function, thereby highlighting the need to develop region-specific trip generation models.

Third, since this study was conducted using TAZ units with higher spatial resolution than Si-gun-gu or eup-myeon-dong levels, it could not incorporate commonly included variables in existing trip generation models, such as income and student numbers. Specifically, because small cities were set as the spatial scope, it was challenging to obtain educational status data, such as schools and universities, with spatial resolution higher than eup-myeon-dong level, resulting in generally low model fits for school commuting. In addition, while PA-based models typically require employment data, employment data matchable at the TAZ level was not available, and thus data on the number of workers had to be used instead. If additional socioeconomic indicators such as income, student numbers, and employment figures can be obtained at the cell tower-based TAZ level, the authors plan to develop trip generation models that consider such variables.

References

1. Aboelenen, K.E., Mohammad, A.N., Elgaar, M.I., and Choe, P., 2021. "Trip Generation Rates Using Household Surveys in the State of Qatar", *Journal of Traffic and Logistics Engineering*, 9(1): 10-19.

2. Bwambale, A., Choudhury, C.F., Hess, S., and Iqbal, M.S., 2021. "Getting the Best of Both Worlds: A Framework for Combining Disaggregate Travel Survey Data and Aggregate Mobile Phone Data for Trip Generation Modelling", *Transportation*, 48: 2287-2314.

3. Bwambale, A., Choudhury, C.F., and Hess, S., 2019. "Modelling Trip Generation Using Mobile Phone Data: A Latent Demographics Approach", *Journal of Transport Geography*, 76: 276-286.

4. Chang, J.S., Jung, D., Kim, J., and Kang, T., 2014. "Comparative Analysis of Trip Generation Models: Results Using Home-based Work Trips in the Seoul Metropolitan Area", *Transportation Letters*, 6(2): 78-88.

5. Colak, S., Alexander, L.P., Alvim, B.G., Mehndiratta, S.R., and González, M.C., 2015. "Analyzing Cell Phone Location Data for Urban Travel: Current Methods, Limitations, and Opportunities", *Transportation Research Record*, 2526(1): 126-135.

6. Etu, J.E. and Oyedepo, O.J., 2018. "Forecasting Trip Generation for High Density Residential Zones of Akure, Nigeria: Comparability of Artificial Neural Network and Regression Models", *Journal of Civil Engineering, Science and Technology*, 9(2): 76-86.

7. Gabriella, A., Abdullah, S., and Soemartojo, S.M., 2019. "Comparison of Poisson and Quasi-Poisson Regression: A Simulation Study", Paper presented at the Proceeding of International Conference on Statistics and Analytics-2019, Swiss: Belhotel Bogor.

8. Kim, I.K., 1997. "Theoretical Comparison of O-D Trips and P-A Trips in Travel Demand Analysis", *Journal of Korean Society of Transportation*, 15(1): 45-62.
김익기, 1997. "교통수요분석에서 통행목적별 O-D 접근방법과 P-A 접근방법의 이론적 비교연구", 「대한교통학회지」, 15(1): 45-62.

9. Kim, K.H. and Song, T.J., 2022. "A Data-driven Approach for Adjusting Taxi Business Zone based on Travel Boundary Using Mobile Phone Network Data", *Journal of Korea Planning Association*, 57(6): 61-73.
김규혁·송태진, 2022. "모바일폰 네트워크 데이터 기반 생활통행권 도출을 통한 택시사업구역 재설정 연구", 「국토계획」, 57(6):

61-73.

10. Kim, K.H., Lee, D.Y., Kim, D.H., Won, M.S., Hong, S.M., and Song, T.J., 2021. "A Study on the Classification and Understanding of Travel Boundary by City Population Scale Based on Mobile Travel Data", *Journal of Korean Society of Transportation*, 39(5): 662-679.
김규혁·이동엽·김동호·원민수·홍성민·송태진, 2021. "모바일 생활통행데이터 기반 도시 인구 규모별 생활권 분류 및 특성 파악", 「대한교통학회지」, 39(5): 662-679.

11. Lafta, S.A. and Ismael, M.Q., 2022. "Trip Generation Modeling for a Selected Sector in Baghdad City Using the Artificial Neural Network", *Journal of Intelligent Systems*, 31(1): 356-369.

12. Lee, S. and Choo, S., 2021. "Development of Trip Generation and Trip Distribution Models for Car-sharing Considering Land Use and Transport Characteristics", *Journal of Korea Planning Association*, 56(5): 69-82.

13. Masoumi, H., 2022. "Home-based Urban Commute and Non-commute Trip Generation in Less-studied Contexts: Evidence from Cairo, Istanbul, and Tehran", *Case Studies on Transport Policy*, 10(1): 130-144.

14. McFadden, D., 1973. "Conditional Logit Analysis of Qualitative Choice Behavior", *Frontiers in Econometrics*, 105-142.

15. Mukherjee, J. and Kadali, B.R., 2022. "A Comprehensive Review of Trip Generation Models Based on Land Use Characteristics", *Transportation Research Part D: Transport and Environment*, 109: 103340.

16. Naser, I.H., 2020. "Studying of Household Characteristic and Their Impact on a Trip Generation Model by MLR", *International Journal of Civil & Environmental Engineering*, 20(1): 1-7.

17. Roorda, M.J., Páez, A., Morency, C., Mercado, R., and Farber, S., 2010. "Trip Generation of Vulnerable Populations in Three Canadian Cities: A Spatial Ordered Probit Approach", *Transportation*, 37: 525-548.

18. Shi, F. and Zhu, L., 2019. "Analysis of Trip Generation Rates in Residential Commuting Based on Mobile Phone Signaling Data", *Journal of Transport and Land Use*, 12(1): 201-220.

19. Song, J.I., Na, S.W., Choo, S.H., 2011. "Developing Trip Generation Models Considering Land Use Characteristics", *The Journal of the Korea Institute of Intelligent Transport Systems*, 10(6): 126-139.
송재인·나승원·추상호, 2011. "토지이용 특성을 반영한 통행발생모형 추정 연구", 「한국ITS학회 논문지」, 10(6): 126-139.

20. Song, T.J. and Lee, H.S., 2018. "The Promise of Mobile Phone Signaling Data for Inferring Mobility Patterns", *International Journal of Highway Engineering*, 20(6): 199-209.
송태진·이해선, 2018. "모빌리티 분석을 위한 모바일 빅데이터 활용성 검토", 「한국도로학회논문집」, 20(6): 199-209.

21. Sulistyono, S., Hasanuddin, A., and Adrisyanti, Y.O., 2012. "Trip Generation Analysis Using Multiple Linear Regression Method on Bumi Estate Muktisari and Taman Gading Housing Jember Regency", Paper presented at The 15th FSTPT International Symposium, Bekasi: Indonesia.

22. Takyi, I.K., 1990. "Trip Generation Analysis in a Developing Country Context", *Transportation Research Record*, 1285: 9-21.

23. Usanga, I.N., Etim, R.K., and Umoren, V., 2020. "Estimation of Trip Generation Rates for Residential Land Use in Uyo, Nigeria", *Nigerian Journal of Technology*, 39(4): 972-980.

24. Yang, F., Li, L., Ding, F., Tan, H., and Ran, B., 2020. "A Data-driven Approach to Trip Generation Modeling for Urban Residents and Non-local Travelers", *Sustainability*, 12(18): 7688.

25. Yim, S. H., 2019. "Local Small Cities's Population Reduction and the Characteristics of Their Growth and Decline", *Journal of the Korean Geographical Society*, 54(3): 365-386.
임석회, 2019. "지방소도시의 인구감소 및 성장과 쇠퇴의 특성", 「대한지리학회지」, 54(3): 365-386.

Appendix

Appendix _ Table 1. The result of Pearson's correlation analysis in Andong-si

	Production					Attraction				
	H↔C	H↔S	H↔X	C↔X	X↔X	H↔C	H↔S	H↔X	C↔X	X↔X
# of bus stop	0.029	-0.088	0.038	0.066	0.068	0.033	-0.087	0.039	0.062	0.065
# of node	0.195.	0.128	0.212.	0.288**	0.282*	0.206.	0.129	0.211.	0.277*	0.269*
sum of link length	-0.285*	-0.23*	-0.282*	-0.224*	-0.215.	-0.282*	-0.229*	-0.285*	-0.229*	-0.221*
Population(~20 years)	0.944***	0.583***	0.925***	0.803***	0.803***	0.934***	0.582***	0.927***	0.822***	0.819***
Population(20~29 years)	0.484***	0.457***	0.368**	0.291**	0.261*	0.477***	0.455***	0.375***	0.307**	0.282*
Population(30~59 years)	0.187.	0.099	0.156	0.159	0.144	0.189.	0.1	0.157	0.151	0.135
Population(60~ years)	0.788***	0.669***	0.812***	0.885***	0.869***	0.804***	0.67***	0.811***	0.875***	0.857***
# of total household member	0.766***	0.662***	0.815***	0.893***	0.884***	0.783***	0.664***	0.812***	0.88***	0.868***
# of mean household member	0.905***	0.606***	0.869***	0.747***	0.746***	0.895***	0.605***	0.872***	0.765***	0.761***
# of 1st industry workers	0.804***	0.428***	0.846***	0.751***	0.767***	0.799***	0.428***	0.847***	0.761***	0.773***
# of 2nd industry workers	0.934***	0.581***	0.89***	0.758***	0.755***	0.921***	0.579***	0.894***	0.78***	0.775***
# of 3nd industry workers	0.878***	0.589***	0.856***	0.784***	0.768***	0.875***	0.585***	0.858***	0.789***	0.772***
Residence	0.64***	0.5***	0.65***	0.648***	0.632***	0.644***	0.496***	0.651***	0.644***	0.626***
Commerce	0.715***	0.575***	0.752***	0.862***	0.843***	0.736***	0.576***	0.751***	0.842***	0.821***
Culture	0.469***	0.493***	0.464***	0.462***	0.463***	0.47***	0.496***	0.468***	0.463***	0.46***
Education	0.028	-0.046	0.104	0.128	0.141	0.031	-0.046	0.103	0.111	0.12
Business	0.504***	0.62***	0.508***	0.586***	0.561***	0.521***	0.62***	0.509***	0.569***	0.541***
Industry	-0.083	-0.082	-0.097	-0.043	-0.061	-0.078	-0.082	-0.099	-0.053	-0.069
Agriculture	-0.311***	-0.276*	-0.319**	-0.285*	-0.285*	-0.309**	-0.274*	-0.32**	-0.289**	-0.288**

***: p-value < 0.001, **: p-value < 0.01, *: p-value < 0.05, .: p-value < 0.1

Appendix _ Table 2. The result of Pearson's correlation analysis in Eumseong-gun

	Production					Attraction				
	H↔C	H↔S	H↔X	C↔X	X↔X	H↔C	H↔S	H↔X	C↔X	X↔X
# of bus stop	0.136	0.086	0.008	-0.012	-0.01	0.141	0.089	0.009	-0.013	-0.008
# of node	0.116	-0.129	0.096	0.071	0.072	0.115	-0.132	0.097	0.077	0.078
sum of link length	-0.12	-0.146	-0.142	-0.147	-0.143	-0.118	-0.144	-0.142	-0.144	-0.142
Population(~20 years)	0.818***	0.206.	0.824***	0.708***	0.743***	0.811***	0.2.	0.83***	0.723***	0.757***
Population(20~29 years)	0.304**	-0.163	0.293**	0.208.	0.172.	0.298**	-0.165	0.294**	0.218*	0.178
Population(30~59 years)	-0.055	-0.04	-0.077	-0.05	-0.07	-0.055	-0.039	-0.076	-0.049	-0.063
Population(60~ years)	0.3**	0.214.	0.234*	0.436***	0.342**	0.306**	0.213.	0.234*	0.416***	0.335**
# of total household member	0.764***	0.124	0.767***	0.726***	0.773***	0.764***	0.119	0.769***	0.737***	0.781***
# of mean household member	0.825***	0.182.	0.808***	0.703***	0.731***	0.818***	0.176	0.813***	0.717***	0.743***
# of 1st industry workers	0.775***	0.299**	0.798***	0.707***	0.758***	0.77***	0.293**	0.802***	0.718***	0.767***
# of 2nd industry workers	0.831***	0.164	0.818***	0.706***	0.727***	0.824***	0.158	0.824***	0.721***	0.741***
# of 3nd industry workers	0.476***	0.209.	0.611***	0.449***	0.518***	0.468***	0.204.	0.619***	0.467***	0.534***
Residence	0.588***	0.201.	0.647***	0.53***	0.554***	0.584***	0.196.	0.651***	0.538***	0.564***
Commerce	0.631***	0.062	0.635***	0.614***	0.612***	0.629***	0.06	0.64***	0.619***	0.62***
Culture	0.279*	0.092	0.322**	0.246*	0.268*	0.274*	0.086	0.324**	0.249*	0.272*
Education	0.557***	0.224*	0.564***	0.504***	0.558***	0.558***	0.221*	0.565***	0.511***	0.564***
Business	0.689***	0.03	0.639***	0.571***	0.596***	0.69***	0.027	0.643***	0.581***	0.609***
Industry	0.034	0.103	-0.057	0.131	0.036	0.041	0.103	-0.055	0.11	0.03
Agriculture	-0.162	-0.101	-0.216*	-0.197.	-0.216.	-0.16	-0.099	-0.217*	0.202.	-0.218*

***: p-value < 0.001, **: p-value < 0.01, *: p-value < 0.05, .: p-value < 0.1

Appendix _ Table 3. The result of Pearson's correlation analysis in Andong-si

	Production					Attraction				
	H↔C	H↔S	H↔X	C↔X	X↔X	H↔C	H↔S	H↔X	C↔X	X↔X
# of bus stop	-0.078	-0.109	-0.125	-0.123	-0.136	-0.079	-0.113	-0.122	-0.123	-0.133
# of node	0.184	0.122	0.171	0.22.	0.211.	0.187	0.122	0.171	0.212.	0.202.
sum of link length	-0.079	-0.137	-0.117	-0.098	-0.107	-0.079	-0.139	-0.115	-0.103	-0.111
Population(~20 years)	0.931***	0.965***	0.938***	0.904***	0.878***	0.931***	0.964***	0.939***	0.912***	0.893***
Population(20~29 years)	0.499***	0.485***	0.514***	0.473***	0.48***	0.497***	0.479***	0.515***	0.479***	0.489***
Population(30~59 years)	0.582***	0.615***	0.59***	0.562***	0.543***	0.582***	0.612***	0.592***	0.572***	0.56***
Population(60~ years)	0.895***	0.891***	0.911***	0.929***	0.93***	0.895***	0.893***	0.91***	0.929***	0.932***
# of total household member	0.919***	0.957***	0.939***	0.917***	0.89***	0.919***	0.957***	0.939***	0.924***	0.903***
# of mean household member	0.937***	0.975***	0.959***	0.923***	0.9***	0.936***	0.975***	0.96***	0.932***	0.914***
# of 1st industry workers	0.93***	0.97***	0.944***	0.906***	0.881***	0.929***	0.968***	0.945***	0.914***	0.896***
# of 2nd industry workers	0.938***	0.975***	0.951***	0.914***	0.89***	0.937***	0.974***	0.953***	0.923***	0.904***
# of 3nd industry workers	0.912***	0.936***	0.909***	0.88***	0.854***	0.912***	0.934***	0.91***	0.888***	0.868***
Residence	0.77***	0.789***	0.754***	0.722***	0.71***	0.77***	0.788***	0.756***	0.733***	0.728***
Commerce	0.884***	0.898***	0.896***	0.879***	0.866***	0.884***	0.898***	0.896***	0.886***	0.878***
Culture	0.678***	0.688***	0.69***	0.605***	0.601***	0.674***	0.682***	0.693***	0.616***	0.615***
Education	0.847***	0.863***	0.845***	0.797***	0.784***	0.845***	0.859***	0.846***	0.809***	0.802***
Business	0.857***	0.9***	0.87***	0.846***	0.82***	0.856***	0.902***	0.872***	0.85***	0.829***
Industry	0.142	0.073	0.126	0.201	0.264*	0.142	0.074	0.125	0.187	0.249*
Agriculture	-0.11	-0.151	-0.156	-0.15	-0.131	-0.11	-0.152	-0.155	-0.153	-0.134

***: p-value < 0.001, **: p-value < 0.01, *: p-value < 0.05, .: p-value < 0.1

Appendix _ Table 4. The result of Pearson's correlation analysis in Sinan-gun

	Production					Attraction				
	H↔C	H↔S	H↔X	C↔X	X↔X	H↔C	H↔S	H↔X	C↔X	X↔X
# of bus stop	0.384**	0.16	0.362**	0.189	0.078	0.384***	0.163	0.357**	0.179	0.061
# of node	0.13	0.098	0.13	0.013	-0.072	0.131	0.092	0.142	0.03	-0.059
sum of link length	0.295*	0.189	0.263*	0.182	0.031	0.292*	0.191	0.246*	0.163	0.001
Population(~20 years)	0.59***	0.349**	0.579***	0.391***	0.267*	0.589***	0.354**	0.571***	0.392***	0.262*
Population(20~29 years)	0.004	0.285*	0.001	0.065	0.069	0.005	0.278*	-0.003	0.072	0.076
Population(30~59 years)	-0.049	-0.085	-0.018	0.025	0.038	-0.048	-0.087	-0.072	-0.045	-0.041
Population(60~ years)	0.43**	0.328**	0.32**	0.262*	0.115	0.431***	0.332**	0.293*	0.244*	0.085
# of total household member	0.316**	0.449***	0.378***	0.299**	0.178	0.315**	0.437***	0.374***	0.321**	0.195.
# of mean household member	0.617***	0.457***	0.592***	0.395***	0.244*	0.614***	0.455***	0.584***	0.397***	0.238*
# of 1st industry workers	0.543***	0.47***	0.577***	0.499***	0.359**	0.543***	0.47***	0.562***	0.504***	0.36**
# of 2nd industry workers	0.648***	0.389***	0.67***	0.529***	0.392***	0.646***	0.392***	0.648***	0.519***	0.377***
# of 3nd industry workers	0.498***	0.261*	0.455***	0.263*	0.142	0.497***	0.268*	0.45***	0.257*	0.128
Residence	0.403***	0.134	0.286*	0.086	-0.045	0.402***	0.138	0.289*	0.087	-0.058
Commerce	0.369**	0.347**	0.367**	0.273*	0.163	0.369**	0.35**	0.398***	0.336**	0.233
Culture	0.169	0.344**	0.2.	0.211.	0.085	0.17	0.345**	0.194.	0.224.	0.091
Education	0.28*	0.181	0.233*	0.092	-0.02	0.278*	0.185	0.249*	0.132	0.014
Business	-0.055	0.04	-0.047	-0.055	-0.046	-0.055	0.036	-0.048	-0.059	-0.053
Industry	0.003	-0.038	0.042	0.099	0.093	0.004	-0.036	-0.029	0.011	-0.006
Agriculture	0.06	-0.002	-0.047	-0.016	-0.109	0.057	0.004	-0.065	-0.039	-0.146

***: p-value < 0.001, **: p-value < 0.01, *: p-value < 0.05, .: p-value < 0.1