An analysis on the interaction of ridership and land-use: the perspective from connectivity between subway stations

Chen Qi *   Siting Chen   Xia Sun   Shichen Zhao

July 19, 2018

Abstract

The purpose of this study is to analyze the interaction of subway ridership and land-use patterns in station catchment areas. The connectivity between two stations is described by the probability of getting off at the destination station; thereby this issue is converted into a discrete choice model. Then the station areas are classified into six clusters according to the proportions of the main land-use functions. A logistic regression model is applied to the object stations selected from each cluster to estimate the passenger choice probability of getting off at the destination station. Finally, the results obtained by different types of stations are compared and discussed. The results show that the variations in the proportion of land-use in departure stations can lead to variations in the choice of destination stations.

Keywords: *Subway ridership, Land-use, Logistic regression, Station-to-station level,*

1   Introduction

Urban rail transit is a type of high-capacity public transport generally found in urban areas. Because of the rapid, punctual and environment-friendly features, the urban rail transit is becoming one of the most important travel modes in modern cities. With the popularity of the urban rail transit in modern cities and the emphasis on sustainable development, the concept of TOD (Transit-Oriented Development) is put forward, intended to build the compact city [2]. Based on the perspective of TOD, many cities around the world have adopted the policy of giving priority to the development of public transport for decreasing the share of motorized travel and increasing the willingness of using public transit. For policymakers, how to grasp the relationship between land-use and passenger volume has become an important issue that they must face.

Until now, there are many studies focusing on the relationship between various factors and transit ridership from the perspective of TOD. Most of the studies are based on regression type model and conducted from the view of station-level, the transit ridership is thought to be affected by the circumstance surrounding the station [5, 22, 25, 10, 23, 20, 13, 16]. Among them, the multiple linear regression models are the earliest and most widely used model [5, 13]. However, the data point in ordinary least squares (OLS) model is treated as a single point, which is not consistent with fact since the transit node is connected to each other. To deal with the relationship among stations in the network, the approach of spatial regression is also introduced into this issue [3, 16]. However, this relationship among stations in spatial regression models is just the expression for the distribution relationship of stations in location, it cannot reflect the real connectivity between two station areas.

---

*E-mail: chenqi.clever@gmail.com

To explore the connectivity between station areas, Choi et al. conducted a station-to-station level investigation into the effect of both origins and destinations on OD metro ridership of Seoul, Korea by using the data from the automatic fare collection system [7]. The factors considered to influence the OD ridership is divided into three groups: the factors of both origin and destination, and the impedance factors between stations. And where the variables of origin and destination are the same, representing the travel characteristics of O and D respectively. The influence of factors on OD ridership is estimated using multiplicative and Poisson regression, with the data of morning, evening peak hours, and midday hours. This station-to-station approach has connected stations by using the factors of both origin and destination. As the result of this empirical study, different land-use functions have different travel characteristics in terms of both time and space. However, this approach still cannot reflect the connectivity between two station areas because of the aggregate processing for data.

Land-use and public transit are coevolving partners in city building [14, 9]. In the urban railway transit system, the ridership between stations is thought to be related to land-use, distribution of functional regions, or travel preferences [24]. In urban planning, TOD can be viewed as a method for balancing the land-use of residences, business, and leisure within walking distance taking the station as the center, while the transit ridership between stations can be viewed as the connectivity of different TOD areas. The TOD area is generally referring to a compact residential district that includes mixing land-use to allow people to most of their daily activities within the easy walking distance of a major transit node [18]. In details, various functional buildings are the carrier for people to live, work and recreate, different functions of buildings correspond different trip purposes. When the functions of buildings within the easy walking distance of a station cannot satisfy the requirement of people's daily activities, people will choose to go to other places to conduct their business by using the transit node such as the subway. Therefore, the distribution of different functional building in a TOD area is considered to not only affect the ridership of the station where it is located but also affect the ridership of other stations connected to that station.

With the goal of explaining the variation in the ridership between stations, this study will focus on the passengers' choices for destination stations from the perspective of complementarity of building function in different TOD area, using the case of subway network in Fukuoka, Japan. The result from this study also can provide a foundation for explaining the connectivity between different TOD areas. Factors that are expected to influence the connection of stations are stated in the next section. Then the approach and model that are adopted in this study are interpreted. Based on the approach and model, the case of the subway system in Fukuoka, Japan is investigated. The discussion for the result is also presented in the last section.

2 Data

2.1 Study context

This study focuses on 35 subway stations in Fukuoka City (the sixth largest city in Japan) which has the largest population in Kyushu Island of Japan (more than 1.5 million). Figure 1 is the research area and the distribution of subway stations. Until now Fukuoka has three operating subway lines, a total of 29.8 kilometers operating mileage. The transport system carries a daily average of more than 0.4 million passengers by 2015 that accounting for more than 20% in total motorized travel (from Fukuoka City Transportation Bureau). Although the Fukuoka subway system is not a large-scale one, it plays a crucial role in public transportation in terms of the city scale and population. Moreover, one thing should be noted that the third line is not directly connected with the first and second line.
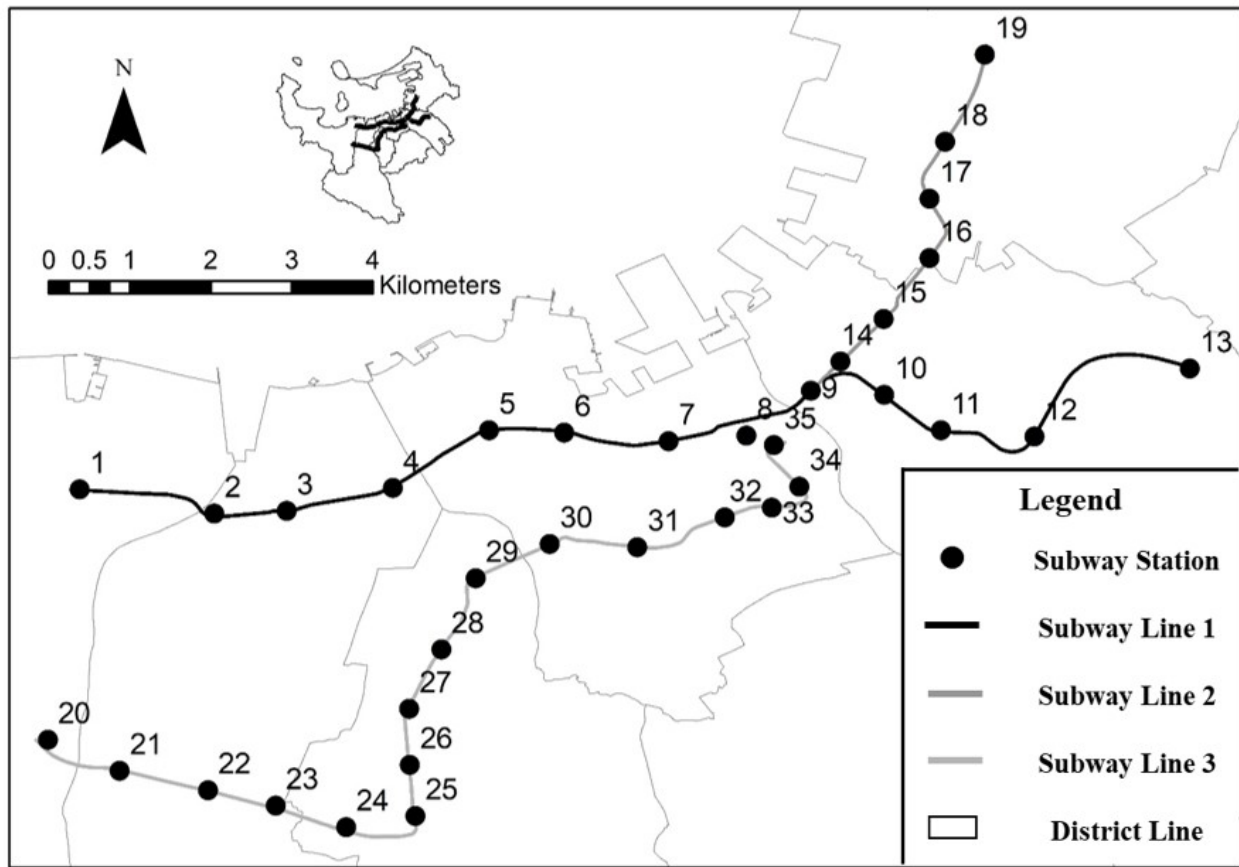
Figure 1: Research area and the distribution of subway stations

2.2 Catchment area of the station

The scale of the catchment area is an important prerequisite for this kind of issue. At present, in the United States, a half-mile-radius (800 m) circle has become the practical standard for rail transit catchment areas based on TODs [12]. The distance of 800m corresponds to the distance people can walk in 10 minutes at the speed of 4.8 km/h. A Japanese case study also supported this 800m catchment area for TOD by using the survey data of <2010 big city traffic census metropolitan area report> [21].

Based on the personal trip survey of Fukuoka, there are nearly 80% of the passengers accessing stations by walking and about 90% of the subway passengers choose non-motorized travel mode, as shown in Table 1. The average walking distance for accessing stations is about 600m according to the average speed and time. Considering the average walking distance is a little less than the main walking distance, it can be inferred that the pedestrian catchment area (Abbreviated as PCA) in Fukuoka is consistent with the most widely accepted distance of 800m. Thereby the distance threshold of 800m is adopted in this study. All the data based on geographical information will be covered by the 800m PCA using the areal interpolation method.

Table 1: Main transportation mode accessing subway stations

	Walking		Bicycle	
	Proportion	Ave time (min)	Proportion	Ave time (min)
Average	78%	7.2	9%	9
median	85%	7	8%	9.5
Min	22%	4.8	0%	0
Max	100%	9.2	33%	20

2.3 Land-use

Land-use is generally accepted as one of the determinants for transit ridership. The indicator of floor area in different building functions can be considered as a detailed expression for land-use. Also, many researchers have verified the influence of floor area in different building functions on transit ridership by conducting empirical studies [20, 13, 6, 16]. Also, all the stations in the network are connected, the change of land-use within the PCA of a station will lead to the changes of transit ridership in all the stations in the same network. For the case of Fukuoka, about 90% of the trips accessing subway stations are non-motorized, in another word, the land-use within the PCA of stations play a crucial role in determining the trip purposes of subway users, as shown in Table 1.

The same with previous studies, this study chose several types of land-use with higher proportion to assess the indicator of land-use mix, including residential, office, commercial, education. The four main types of land-use account for about 85% of all the floor area in Fukuoka City, especially in subway PCA, reaching more than 90%, as shown in Table 2. In addition to the indicator of floor area, the index of the land-use mix is also assumed to be a crucial factor for explaining the connectivity of stations [1, 4, 11]. Different from the general definition of land-use mix, this study redefines it into the aggregation of land-use. The Euclidean Metric is used for evaluating the deviation of land-use aggregation in each subway station with respect to a reference value. The value of this indicator is ranged from 0 to 1, in which the lower value represents a higher diversity in land-use function, while the higher value means the land-use function is less diverse. This indicator of land-use aggregation is defined as for Equation 1, it is speculated to have a negative impact on ridership. To describe the mixture of land-use closer to the facts, the referenced balance proportion

of land-use types is decided by the average proportion of all subway station PCA (800m) in Fukuoka City, shown in Table 2.

$$A = \sqrt{\sum (S_i - P_i)^2} \tag{1}$$

Where:

i represents the type of land-use (respectively government, commercial, residence and education).

A is the indicator for aggregation of land-use functions.

P_i is the average proportion of the land-use with type i.

L_i is the floor area of land-use with type i within the PCA.

Table 2: Proportion of land-use in terms of the range of PCA

Range of PCA	Residence	Office	Commerce	Education	Total
600	51.80%	19.30%	13.80%	5.50%	90.50%
800	55.30%	17.50%	12.00%	6.20%	90.90%
1000	57.60%	16.50%	10.80%	6.00%	90.90%
1200	59.20%	15.50%	10.20%	5.90%	90.80%
Ave	62.80%	10.10%	7.50%	5.30%	85.70%

2.4 Impedance

Another factor generally used for representing the connectivity between stations is the remoteness, which is also widely adopted as impedance in the gravity model for explaining the connectivity between traffic zones [15, 17, 19]. Impedance is the index that represents the connectivity and cost between origin and destination, it can affect the choice of the trip in terms of both destination and travel mode. Generally, the factor of impedance can be considered from two aspects in this issue, one is the internal impedance that representing the convenience for accessing the station [8, 6]; the other is external impedance that evaluating the connectivity between stations [20].

For Fukuoka, there are about 90% of all the trip accessing stations by walking and bicycles, as shown in Table 1. Because the PCA is set by the pedestrian distance, the internal road impedance has been already considered in the phase of setting the PCA. It can be thought that there is no need to create a new indicator to represent the internal impedance in this case study.

Since the focus of this study is the connectivity between stations, it can be assumed that the interconnection between stations in a subway network is an influential factor in determining the ridership between stations. Two types of external impedance are considered in this study. The operation distance of subway line between two stations is the direct indicator representing the spatial connectivity between two stations. And the impedance of competing modes is also considered in this study. Two indicators of bus accessibility and bus capacity are proposed as shown in Equation 2 and Equation 3 respectively. The former one is used for representing the convenience for accessing the station which is speculated as a positive factor for ridership, while the later one is the transport capacity of the bus within the PCA of the station which is considered as a negative factor for ridership because it may share part of ridership from the subway.

5

$$BC = \sum_{k}^{K} \sum_{r}^{R} f_r^k \tag{2}$$

$$BA = \sum_{k}^{K} R_k \tag{3}$$

Where:

K is the numbers of bus stations within PCA.

R is the number of lines at one bus station.

f_r^k is the frequency of the rth line at the kth station

R_k is the number of bus lines passing through the kth bus station.

3 Model

It is difficult to analyze the connectivity of all the stations simultaneously. To simplify this issue, the study will be started by one single station, and then the connectivity between this station and all the other stations connected to it will be investigated. For a passenger, the behavior of going to some places by subway can be viewed as a procedure of choice. In this selection process, the choice is decided by the trip purpose of the passenger. As stated in the introduction, the type of building functions can be mainly categorized into residence, office, education, and commerce, which represent the trip purposes of going home, business, commute, and leisure respectively. Therefore, this is-sue can be converted into a discrete choice model, in which the dependent variables are the building environment in the station catchment area, and the independent variable is the choice of the destination station.

Taking one station as the destination station, it will correspond to the other departure stations in the subway system. The passengers from the all the departure stations have two options at that destination station, that is getting off and not getting off at this destination station. This probability of getting off at that destination station can be viewed as the representation of the connectivity between the departure station and the destination station. At this point, this issue has been converted into a binary choice problem that investigating the probability of getting off at one subway station. Thus, the Binary-Logistic Regression is introduced into this study to deal with this issue. Equation 4 shows the Binary-Logistic Regression that will be adopted in the next section.

$$p_i(y_i = 1 \mid X_i) = \frac{1}{1 + e^{-(\alpha + X_i)}} \tag{4}$$

Where:

p_i is the probability of getting off the subway for the ith passenger.

y is the choice of passengers.

X_i is the attribute vector of the ith passenger.

α is the residual item

For the estimation of a regression model, the model of logistic regression built in this study is a non-linear regression model, which can be directly estimated by using maximum likelihood estimation (MLE). Thus, the MLE is adopted to estimate the coefficient. In this study, the sample size is the passenger volume of all the departure stations connected to the investigated destination station, which is marked as N. The same with Equation 4, p_i is the probability for choosing to get off, thus, $1 - p_i$ represent the probability for choosing not to get off. Then the occurrence probability $L(\theta)$ in all observation sample can be expressed as Equation 5:

$$L(\theta) = \prod_{i=1}^{N} p_i{}^{y_i} (1 - p_i)^{(1-y_i)}$$

(5)

4   Result

The characteristic of land-use within the PCA of a station can represent the characteristic of passengers departing from that station. Since a city is the collection of different functional regions, the land-use characteristics of each PCA are also different. To explore the connectivity between stations with different types of land-use characteristics, this study firstly classifies all the stations in terms of the proportion of different function of land-use and population density. Then the typical stations are chosen from each group of different land-use type, thereby estimating the connectivity between stations by conducting the logistic regression.

As the result, all the stations are fallen into six categories, they are medium-density residence, high-density residence, education, office, commerce, and airport. As shown in Table 3, the average values of attributions in each group are represented, furthermore, the factors of bus service and land-use aggregation are also put in this table to help to observe the difference among groups. As is shown, both the group of medium-density residence and the group of high-density residence have a higher proportion of residence floor area, but there is a significant difference in the population density in the two groups. It is clearly that the education type has the highest proportion of education floor area. Both the office type and the commerce type have a higher proportion in office than other types, but the commerce proportion in commerce type is also higher than other types. Additionally, the station of Fukuoka airport does not belong to any group, which mainly takes the feeder traffic from the airport to urban area. It also can be noted that the residence and office type have a relatively lower land-use aggregation and higher population density.

To investigate the stations with significant differences in land-use characteristic, the most typical object station is chosen from each group in terms of the guidelines as follow. 1) Avoid selecting the stations with the highest value of feature indicator. Because the high value may be caused by the particularity of its distribution of land-use and road network, it does not have generality in terms of the type it is. 2) Avoid selecting the transferable stations. Because the transit ridership of this kind of stations generally relates the other stations, which cannot reflect the distribution of land-use within the PCA. 3) Avoid selecting the stations with much lower or higher population density than other stations. Because the low or high population density may lead to different travel preference due to the difference in the distribution of facilities (such as hospitals, schools, and transportation facilities etc.). Finally, six stations are selected from six groups, the factor of each station is shown in Table 4.

All the estimation showed in in Table 4 are statistically significant at the 0.05 or 0.01 level. The index of *Exp(B)*, which means the odds ratio, represents the extent to the effect on the probability of choice if the coefficient *B* changes one unit. Here giving a brief explanation of the index of *Exp(B)*, since the unit and meaning of the variables entered the models are different. For the variables of land-use, the index of *Exp(B)* represents that the 1 percent (0.01 unit) change in the variables will lead to an *Exp(B)* multiples of increases in the probability of getting off at the object station. For

Table 3: Station classification

Type	Population density	Commerce	Office	Residence	Education	Land-use Aggregation	Bus Capacity	Bus Accessibility
Medium-density residence	98	4%	3%	83%	4%	0.34	18	28
High-density residence	173	5%	7%	76%	6%	0.26	51	80
Education	83	5%	6%	51%	22%	0.3	45	52
Office	135	8%	31%	49%	3%	0.18	83	131
Commerce	69	34%	32%	24%	1%	0.47	132	213
Airport	36	1%	3%	45%	2%	0.23	32	56

*Note: the highlighted cells refer to the representative values that having much higher percentage than other land-use types.

Table 4: Result of Logistic Regression

| Destination Station | | | Variables in Boarding Station | | | | | | | |
| | | | Land-use | | | | | Impedance | | |
Type	Station	Statistical index	Commerce	Office	Residence	Education	Land-use Aggregation	Distance	Bus Capacity	Bus Accessibility
Medium-density residence	Kamo	B	-0.05	-0.06	-0.06	-0.07	-0.03	0.07	0.01	-0.01
		P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.04
		Exp(B)	0.95	0.94	0.94	0.93	0.97	1.07	1.01	1.00
High-density residence	Fujisaki	B	0.01		-0.01	0.01	-0.02	-0.06	0.00	0.00
		P-value	0.00		0.00	0.00	0.00	0.00	0.00	0.00
		Exp(B)	1.01		0.99	1.01	0.98	0.95	1.00	1.00
Education	Hakozakikyudai	B	0.02		0.02	0.03	-0.01	-0.13	0.00	0.00
		P-value	0.00		0.00	0.00	0.00	0.00	0.04	0.00
		Exp(B)	1.02		1.02	1.03	0.99	0.88	1.00	1.00
Office	Gofukumachi	B	0.02	-0.02	0.02	0.05	-0.01		0.00	0.00
		P-value	0.00	0.00	0.00	0.00	0.01		0.03	0.00
		Exp(B)	1.02	0.98	1.02	1.06	0.99		1.00	1.00
Commerce	Tenjinn	B	-0.01	-0.01	0.01	0.03	0.02	0.06		0.00
		P-value	0.00	0.00	0.00	0.00	0.00	0.00		0.00
		Exp(B)	1	0.99	1.01	1.03	1.02	1.06		1.00
Airport	Fukuoka Airport	B	-0.03	0.03	-0.01	-0.02	0.03	-0.08	0.00	0.00
		P-value	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
		Exp(B)	0.97	1.03	0.99	0.99	1.03	0.92	1.00	1.00

Note: *B* is the coefficient, *Exp(B)* is the odds ratio. The blank means the estimated value is not significant at the 0.05 level.

the variables of impedance, it means the 1 unit change in the variables will lead to an *Exp(B)* multiples of increases in the probability of choosing the object station as a destination.

Here takes the Tenjin station as an example to make a further explanation. From the point of land-use, the increase of commerce and office floor area in the catchment area of the boarding stations can lead to a decrease in the probability of choosing the Tenjin station as the destination station; while the increase of residence and education floor area in the catchment area of boarding stations can raise the probability of choosing to go the Tenjin station. If interpreted in terms of connectivity, the business type of Tenjin Station is weakly connected to stations with the same commercial type, and the stations of office type; while the stations of business type have relatively strong connectivity with the stations of residence and education type.

5 Discussion

From the perspective of TOD, this study investigated the connectivity between stations with different land-use characteristic by examining the impact of land-use and impedance on subway passengers' choice of the destination station. As shown in Table 3, the PCA of the station in Fukuoka shows a significant characteristic of land-use distribution, and all the subway stations are categorized into six major types. The result of logistic regression shows the coefficient of factor influencing the choice of the destination station, which can be explained as the connectivity between different PCAs.

Some findings of the land-use factors can be known from the result in Table 4. The increase in the proportion of residence area of the departure station can lead to an increase in the choice of taking the office and education type station as destination stations. This can be explained as the result of commuter traffic since there is a relatively stronger connectivity between the residential area and work area [1]. But the situation may change in the medium-residence type due to the different travel preferences in the area with different population density. Another interesting finding is that all kinds of land-use have positive connectivity with the education type station. One speculation is that students tend to take public transit because of the low income. Overall, one kind of land-use generally shows a rejection effect on the station which has a similar consist of land-use type. Such as the Tenjin station located in the CBD area, the proportion of commerce area in the PCA of departure station causes a negative effect on the choice of getting off at Tenjin station. Moreover, the variable of land-use aggregation shows that there is a positive connectivity between the areas with an unbalanced distribution of land-use.

The factors of impedance also showed a good statistical significance, however, the results do not seem to show a certain regularity in different types of stations. Here some possible speculations of the reasons are given for helping to find the limitation of this study and explore the direction for the next study. First, the share of different transportation modes is not considered in this study. The distance between two stations may also affect the variation in the share of different transportation modes, thus it is not stable in representing the impedance. Second, the variables of bus service describe the features of one single station, but not the features that reflecting the connectivity between two stations. If considering two stations, one is in the downtown where the transportation hub locates, and one is in suburban where there are few public transportation facilities; it can be inferred that even though the bus service is rich in the downtown area, it affects very little on the station located in the suburban area.

6 Conclusion

This study investigated the effect of land-use and impedance between stations on the OD subway ridership from the perspective of the connectivity. The effect of factors on the choice of destination station was estimated using the logistic regression model. The result showed that the influence of land-use on the ridership between stations was effective, and this result could be explained consistently with fact; while the factor of impedance was still difficult to get explained.

From the results obtained in this article, in the urban rail transit, the land-use within the station catchment area is an important factor in determining the destination of passengers. Correspondingly, the change in the choice of destination can be reflected to the transit ridership. Furthermore, this change in the transit ridership of each station, which is caused by the change in the choice of destination, is the reflect of the connectivity between the station and the station in terms of land-use in each station catchment area. The results are clearly shown in Table 4.

Based on the results in this study, the next stage of this research will focus on the share ratio in different transportation modes to make a deeper exploration of subway ridership. The factor of impedance will also be rebuilt to help to describe the connectivity between stations more accurately.

References

[1] D a Badoe and E J Miller. Transportation - Land Use Interaction: Empirical Findings in North America and their Implications for Modeling. *Transportation Research - D*, 5:235–263, 2000.

[2] Peter Calthorpe. *The next American metropolis: Ecology, community, and the American dream*. Princeton Architectural Press, 1993.

[3] Osvaldo Daniel Cardozo, Juan Carlos García-Palomares, and Javier Gutiérrez. Application of geographically weighted regression to the direct forecasting of transit ridership at station-level. *Applied Geography*, 34(4):548–558, 2012.

[4] Robert Cervero. *Transit-Oriented Development in the United States: Experiences, Challenges, and Prospects*. Transportation Research Board, 2004.

[5] Robert Cervero and Kara Kockelman. Travel demand and the 3Ds: Density, diversity, and design. *Transportation Research Part D: Transport and Environment*, 2(3):199–219, 1997.

[6] Arnab Chakraborty and Sabyasachee Mishra. Land use and transit ridership connections: Implications for state-level planning agencies, 2013.

[7] Jinkyung Choi, Yong Jae Lee, Taewan Kim, and Keemin Sohn. An analysis of Metro ridership at the station-to-station level in Seoul. *Transportation*, 39(3):705–722, 2012.

[8] Xuehao Chu. Ridership models at the stop level. Technical report, 2004.

[9] Hank Dittmar and Gloria Ohland. *The new transit town: best practices in transit-oriented development*. Island Press, 2012.

[10] N Estupinan and D A Rodriguez. The relationship between urban form and station boardings for Bogota's BRT, 2008.

[11] Lawrence D. Frank, Martin A. Andresen, and Thomas L. Schmid. Obesity relationships with community design, physical activity, and time spent in cars. *American Journal of Preventive Medicine*, 27(2):87–96, 2004.

[12] Erick Guerra and Robert Cervero. Is a Half-Mile Circle the Right Standard for TODs? *ACCESS Magazine*, 2013.

[13] Javier Gutiérrez, Osvaldo Daniel Cardozo, and Juan Carlos García-Palomares. Transit ridership forecasting at station level: An approach based on distance-decay weighted regression. *Journal of Transport Geography*, 19(6):1081–1092, 2011.

[14] Susan Handy. Smart growth and the transportation-land use connection: What does the research tell us? *International Regional Science Review*, 28(2):146–167, 2005.

[15] Tomasz Iwanow and Colin Kirkpatrick. Trade facilitation, regulatory quality and export performance. *Journal of International Development*, 19(6):735–753, 2007.

[16] Myung-Jin Jun, Keechoo Choi, Ji-Eun Jeong, Ki-Hyun Kwon, and Hee-Jae Kim. Land use characteristics of subway catchment areas and their influence on subway ridership in Seoul. *Journal of Transport Geography*, 48:30–40, 2015.

[17] Konstantinos Kepaptsoglou, Matthew G Karlaftis, and Dimitrios Tsamboulas. The Gravity Model Specification for Modeling International Trade Flows and Free Trade Agreement Effects: A 10-Year Review of Empirical Studies. *The Open Economics Journal*, 3:1–13, 2010.

[18] Hollie M Lund, Robert Cervero, and Richard W Wilson. *Travel characteristics of transit-oriented development in California*. 2004.

[19] Volker Nitsch. National borders and international trade: evidence from the European Union. *Canadian Journal of Economics/Revue canadienne d'économique*, 33(4):1091–1105, 2000.

[20] Keemin Sohn and Hyunjin Shim. Factors generating boardings at Metro stations in the Seoul metropolitan area, 2010.

[21] Tadakatsu Nakamura. An Empirical Analysis of Situations around Station as a Factor Affecting the Number of Station Users. *Regional Policy Research*, 17(3):15–26, 2015.

[22] Brian D. Taylor, Douglas Miller, Hiroyuki Iseki, and Camille Fink. Analyzing the determinants of transit ridership using a two-stage least squares regression on a national sample of urbanized areas. *University of California Transportation Center*, 2003.

[23] Brian D. Taylor, Douglas Miller, Hiroyuki Iseki, and Camille Fink. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas, 2009.

[24] Gregory Thompson. Achieving Suburban Transit Potential: Sacramento Revisited. *Transportation Research Record: Journal of the Transportation Research Board*, 1571:151–160, 1997.

[25] Fang Zhao, Lee-Fang Chow, Min-Tang Li, and Xuemei Liu. A Transit Ridership Model Based on Geographically Weighted Regression and Service Quality Variables. *Lehman Center for Transportation Research, Florida International University, Miami, Florida. http://lctr. eng. fiu. edu/re-project-link/finalDO97591_BW. pdf (accessed December 12, 2010)*, 2005.