Journal of Southeast University (English Edition) Vol. 30, No. 1, pp. 107 – 112 Mar. 2014 ISSN 1003—7985

Influence of land use characteristics and trip attributes on commuting mode choice: a case of Nanjing

Liu Zhiwei Deng Wei Ji Yanjie

(School of Transportation, Southeast University, Nanjing 210096, China)

Abstract: The effects of socio-demographics, land use characteristics and trip attributes on the commute mode choice are studied with a nested logit (NL) model. Based on the random utility maximum theory, the NL model is formulated. The analysis is carried out in the main area of Nanjing. Direct and cross elasticities are calculated to analyze the effects of travel time and travel cost on the selection of travel mode choice. The results reveal that the non-motorized travel modes are more attractive in the areas with higher housing and employment accessibility and car owners are found to be more likely to commute to work by car. The bus and subway choice probabilities are more sensitive to changes in travel times than to changes in travel costs. In conclusion, a comprehensive public transit system and effective integration of land use and transportation policies help to relieve the traffic congestion levels caused by the increasing urban sprawl.

Key words: traffic engineering; land use; travel mode choice; nested logit model; elasticity analysis

doi: 10.3969/j.issn.1003 – 7985.2014.01.020

In the traditional "four-step" transportation demand analysis, there are two typical classes of variables: 1) Trip attributes such as travel time and travel cost; 2) Individual and household socio-demographics. The land use characteristics are often ignored based on the assumption that the impacts of land use characteristics on the travel behavior are captured in travel times and travel costs and have been considered by travelers in the long-term commuting trips. However, there are important attributes of land use that cannot be fully described by travel times and travel costs. Therefore, it is an important task to test whether the ignored land use factors affect travel mode choice and, if so, to what extent.

There is no consensus on the impact of land use on the travel behavior. The reason for this contradiction is a difference in the selected variables and the used approaches. For example, some researchers explore the impact of spatial structure of areas on travel behavior with the number of inhabitants, which are known to have little explanatory power. On the contrary, some experts regard the accessibility of facilities as one of the most important spatial variables[1]. Many studies used aggregate scale data such as census tract, block group, or zip code area, which masked the individual differences and ignored many important factors influencing the travel behavior[2].

The discrete choice modeling based on the random utility maximization (RUM) theory has been applied to analyze the choice problem of residential location and travel behavior and empirically estimate the probability choice models. The multinomial logit (MNL) model is the most widely used model because of its simple mathematical structure and ease of estimation[3]. However, the MNL model assumes that the distribution of the random error terms is independent and identical over alternatives, which leads to the independence of irrelevant alternatives (IIA) property and causes the cross-elasticities between all pairs of alternatives to be identical[4–5]. To overcome this restrictive assumption, the nested logit (NL) model, which is known as the relaxation of the MNL model and derived from Mac Fadden's generalized extreme value (GEV) model, can be used for estimation in practical applications[6–7]. The NL model allows correlation among groups of alternatives.

This paper analyzes the combined effect of socio-demographics (including individual and household attributes) and land use characteristics on commuting travel mode choice with an NL model within the theoretical framework of GEV class.

1 Model Specification

1.1 Model structure

The NL model, first proposed by Ben-Akiva[8–9], is an extension of the MNL model designed to capture some correlations among alternatives. It is based on the partitioning of the choice set C_n into M nests C_{mn} such that

$$C_n = \bigcup_{m=1}^{M} C_{mn} \tag{1}$$

and

$$C_{mn} \cap C_{m'n} = \varnothing \qquad \forall\, m \neq m' \tag{2}$$

The utility function of each alternative is composed of a

Received 2013-07-31.

Biographies: Liu Zhiwei (1987—), male, graduate; Deng Wei (corresponding author), male, doctor, professor, dengwei@seu.edu.cn.

Foundation item: The National Natural Science Foundation of China (No. 50908051).

Citation: Liu Zhiwei, Deng Wei, Ji Yanjie. Influence of land use characteristics and trip attributes on commuting mode choice: a case of Nanjing[J]. Journal of Southeast University (English Edition), 2014, 30(1): 107 – 112. [doi: 10.3969/j.issn.1003 – 7985.2014.01.020]

term specific to the alternative and a term associated with the nest. If i is an alternative from nest C_{mn}, we have

$$U_{in} = \tilde{V}_{in} + \tilde{\varepsilon}_{in} + \tilde{V}_{C_{mn}} + \tilde{\varepsilon}_{C_{mn}} \tag{3}$$

The error terms $\tilde{\varepsilon}_{in}$ and $\tilde{\varepsilon}_{C_{mn}}$ are supposed to be independent. As in the MNL model, the error terms $\tilde{\varepsilon}_{in}$ are assumed to be independent and identically Gumbel distributed with scale parameter μ_m, which can be different for each nest. The distribution of $\tilde{\varepsilon}_{C_{mn}}$ is similar to those of the error terms $\tilde{\varepsilon}_{in}$ and $\tilde{\varepsilon}_{C_{mn}}$. The random variable $\max_{j \in C_{mn}} U_{jn}$ is Gumbel distributed with scale parameter μ. Each nest within the choice set is associated with a composite utility,

$$V_{C_{mn}} = \tilde{V}_{C_{mn}} + \frac{1}{\mu_m} \ln \sum_{j \in C_{mn}} e^{\mu_m \tilde{V}_{jn}} \tag{4}$$

The second term is called the expected maximum utility, logsum, inclusive value or accessibility in the literature. The probability for individual n to choose alternative I within nest C_{mn} is given by

$$P(i \mid C_n) = P(C_{mn} \mid C_n) P(C_i \mid C_{mn}) \tag{5}$$

where

$$P(C_{mn} \mid C_n) = \frac{e^{\mu V_{C_{mn}}}}{\sum_{i=1}^{M} e^{\mu V_{C_{mn}}}} \tag{6}$$

and

$$P(C_i \mid C_{mn}) = \frac{e^{\mu \tilde{V}_{in}}}{\sum_{J \in C_{mn}} e^{\mu_m \tilde{V}_{jn}}} \tag{7}$$

Parameters μ and μ_m reflect the correlation among alternatives within the nest C_{mn}.

The covariance between the utility of two alternatives i and j in nest C_{mn} is

$$\operatorname{cov}(U_{in}, U_{jn}) = \begin{cases} \operatorname{var}(\tilde{\varepsilon}_{C_{mn}}) & i,\ j \in C_{mn} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

and the correlation is

$$\operatorname{corr}(U_{in}, U_{jn}) = \begin{cases} 1 - \dfrac{\mu^2}{\mu_m^2} & i,\ j \in C_{mn} \\ 0 & \text{otherwise} \end{cases} \tag{9}$$

Therefore, as the correlation is non-negative, we have

$$0 \leqslant \frac{\mu}{\mu_m} \leqslant 1 \tag{10}$$

and

$$\frac{\mu}{\mu_m} = 1 \Rightarrow \operatorname{corr}(U_{in}, U_{jn}) = 0 \tag{11}$$

The parameters μ and μ_m are closely related in the model. Actually, only their ratio is meaningful. It is not possible to identify them separately. A common practice is to arbitrarily constrain one of them to a specific value (usually 1). For the NL models, there are two ways to estimate the parameters: the limited information maximum likelihood (LIML) and the full information maximum likelihood (FIML). The models presented in this paper are all calibrated using the FIML estimation approach.

In this paper, direct and cross elasticities are presented according to Ref. [4]. Direct elasticities indicate the variation in a decision maker's choice probability due to a 1% change in one of the attributes affecting that alternative, while cross elasticities indicate the variation in the choice probability due to a 1% change in an attribute affecting another alternative. The direct elasticities and cross elasticities of an alternative l, which appear in one or more nests with logsum μ_m less than one are formulated as

$$E_{dl} = \frac{\sum_m P_m P_{il\,m} \left[(1 - p_i) + \left(\frac{1}{\mu_m} - 1 \right) (1 - P_{il\,m}) \right]}{p_i} \theta_l x_{il} \tag{12}$$

$$E_{cl} = -\left[P_i + \frac{\sum_m \left(\frac{1}{\mu_m} - 1 \right) P_m P_{il\,m} P_{jl\,m}}{p_j} \right] \theta_l x_{il} \tag{13}$$

In this paper, six commute travel modes are considered: walk, bicycle, bus, subway, car and taxi. The walk-bicycle nested model represents a higher level of competitiveness between walk and bicycle than that among other modes; private car and taxi represent a higher level of competitiveness among taxi and private car than that among other modes; bus and subway represents a higher level of competitiveness between bus and subway than that among other modes. Therefore, a two-level NL model is used here (see Fig. 1).

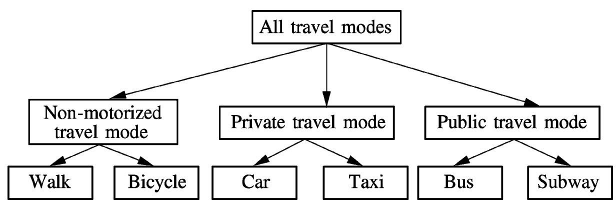

Fig. 1 Tree structure of a nested logit model for travel modes

1.2 Variable specification

The main factors which influence commute travel mode choice can be classified into three groups: socio-economic characteristics, land use characteristics, and travel related attributes.

Among these variables, socio-economic characteristics may play an even more important role in shaping the commute travel mode choice[10-12]. Land use characteristics, also known as spatial factors, are related to an individual's residential location. First, residential (i. e., population, store) density has been commonly used to study intra urban variation in commuting time or distance[13]. Generally, the areas with higher residential density lead to a smaller number of trips and a lower percentage of car use and more attractive public transit[14]. Secondly, mix or co-location of different land uses has been identified to have an important effect on the travel behavior[15]. Finally, accessibility, which refers to the ability of individuals to travel and to participate in activities at different locations in a built environment, is considered as one of the most important factors of property values[16]. Travel-related attributes include travel time and travel cost. And the travel time and the travel cost are associated with the commuting distance, i. e. the distance between residential and employment location. Tab. 1 provides descriptions for the entire set of independent variables used in the modeling process.

Tab. 1 Variables used in the simultaneous choice of residential location and travel mode model

Variable name			Description
Socio-economic characteristics	Gender		Dummy variable: male = 1; female = 0
	Income	Income 1 (reference variable)	Dummy variable indicating whether the individual's annual income is below 50 000 yuan
		Income 2	Dummy variable indicating whether the individual's annual income is between 50 001 yuan and 100 000 yuan
		Income 3	Dummy variable indicating whether the individual's annual income is over 100 000 yuan
	Age	Age 1 (reference variable)	Dummy variable indicating whether the individual is below 24 years old
		Age 2	Dummy variable indicating whether the individual is between 25 and 55 years old
		Age 3	Dummy variable indicating whether the individual is over 56 years old
	Car ownership		Dummy variable indicating whether the individual owns a car
Land use characteristics at the traffic analysis zone level	Employment accessibility		Continuous variable: the employment accessibility of the TAZ which the individual's residential location belong to
	Housing accessibility		Continuous variable: the housing accessibility of the TAZ which the individual's residential location belong to
	Population density		Continuous variable: the population density of the TAZ which the individual's residential location belong to
	Employment density		Continuous variable: the employment density of the TAZ which the individual's residential location belong to
	Point density of department stores		Continuous variable: the point density of department stores of the TAZ which the individual's residential location belong to
	Ratio of jobs to residents		Continuous variable: the ratio of jobs to residents of the TAZ which the individual's residential location belong to
Travel attributes	Travel time		Continuous variable: total time of a trip
	Travel cost		Travel costs for car and public transport fares as a function of distance

2 Data and Descriptive Statistics

2.1 Data sources

Nanjing, the capital of Jiangsu province, is selected for the case study because of its typical socio-spatial structure and deteriorating traffic condition. The study area includes 6 out of 13 administrative districts of Nanjing: Gulou, Jianye, Baixia, Qinhuai, Xuanwu and Xiaguan. There are 198 traffic analysis zones (TAZs).

The primary data of sources for the analysis is based on the 2011 Nanjing residents travel survey (NBTS) carried out by the Nanjing Institute of City and Transportation Planning (NICTP). The survey data provide socio-economic information on individuals and information on the characteristics of the households. It also contains information on disaggregate origin-destination, travel time, travel distance, travel modes as location of residence and location of work. About 5 377 individuals, above 16 years of age, in 1 661 households are involved. Land use data sources used for the GIS analysis including digital data are provided by the Planning Bureau of Nanjing.

2.2 Descriptive analysis

We conduct some descriptive analysis to explore the association between the socio-economic characteristics and commuting travel mode. The modal split of different groups with respective socio-economic characteristics is shown in Tab. 2. As shown in Tab. 2, the travel mode choices of different people groups are quite different. As expected, the younger people are found to be more likely to commute to work by bus probably due to low income. The male has a higher proportion in going to work by car while the female has a higher proportion in travelling by bike. About a half of the people whose annual income above 100 000 yuan and almost 60% of car owners choose cars to commute to work, which indicates that travelling by car is more convenient and it becomes the main travel mode of wealthy people and car owners. Tab. 2 illustrates the land use characteristics of the main area of Nanjing at the TAZ level. The distribution of point density of department stores, employment density, and the accessibility measure for jobs by car are presented in Figs. 2 to 4.

Tab. 2 Travel mode choice of different population groups

Population group		Travel mode frequencies/%					
		Walk	Bicycle	Car	Taxi	Bus	Subway
Age	16 to 24	7.69	48.72	12.82	2.56	23.08	5.13
	25 to 49	10.65	58.71	17.88	0.70	11.52	0.54
	>50	15.94	62.32	2.90	2.90	11.59	4.35
Gender	Female	12.64	67.03	6.04	0.55	11.54	2.20
	Male	10.56	58.00	18.15	0.66	11.96	0.67
Annual income/yuan	<50 000	11.38	66.70	8.94	0.32	12.02	0.64
	50 000 to 100 000	10.75	48.36	28.06	0.90	10.75	1.18
	>100 000	6.72	26.89	47.90	2.52	14.29	1.68
Car ownership	No	12.34	69.62	3.29	0.77	13.79	0.19
	Yes	6.44	29.97	56.30	0.56	6.45	0.28

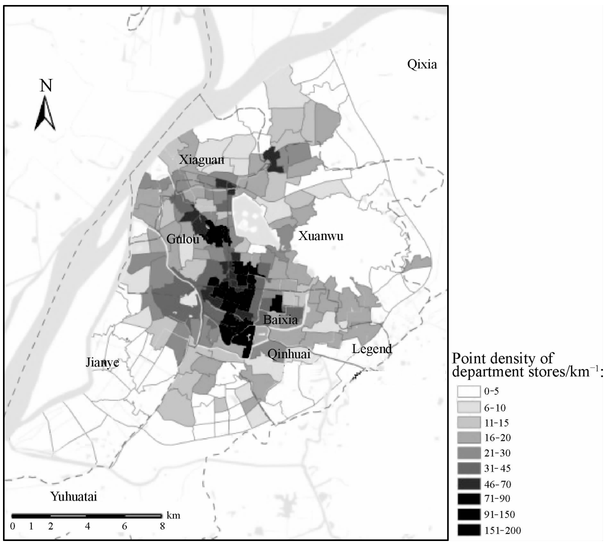

Fig. 2 Spatial distribution of the point density of department stores

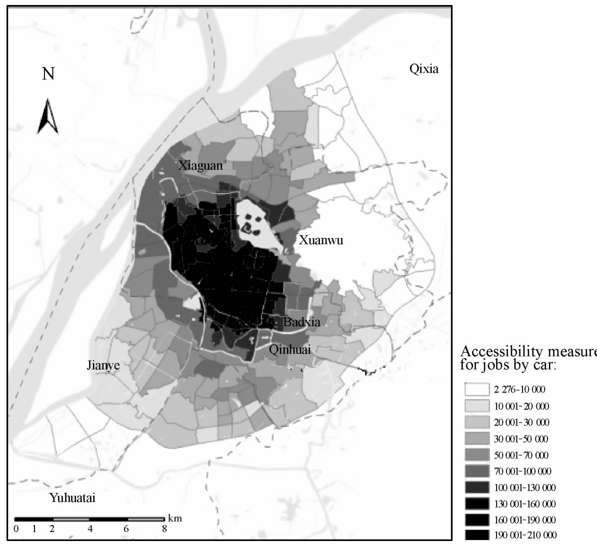

Fig. 4 Spatial distribution of the accessibility measure for jobs by car

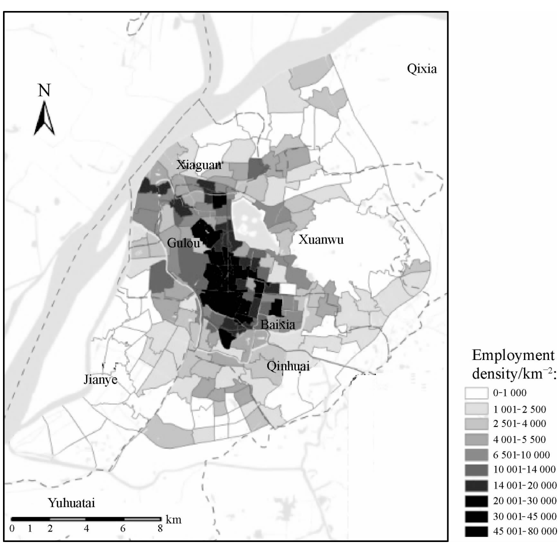

Fig. 3 Spatial distribution of the employment density

3 Model Estimation

In order to study the effect of socio-economic variables and land use characteristics on the choice, this paper analyzes the commute travel mode choice of the people who live in the main area of Nanjing. The NL model results obtained with the freely available optimization package Biogeme are reported in Tab. 3[17].

As expected, coefficients for travel time and travel cost have negative signs and are statistically significant at levels of confidence well above the usual 95% limit, which is consistent with the random utility theory. The age above 25 years and annual income over 100 000 yuan have negative effects on the systematic utility of a certain alternative. Expected signs are obtained from socio-economic characteristics. The car owners are found to be more likely to commute to work by car. When it comes to land use characteristics, the variables for housing and employment accessibility have positive signs while the coefficients for the employment density and ratio of jobs to residents have negative signs, which indicate that the non-motorized travel modes are more attractive in the areas with higher housing and employment accessibility.

Tab. 3 Estimation results for the NL model

Variable	Coefficient values	*T*-statistics
Travel time	−1.92	−8.60
Travel cost	−0.058 2	−3.68
Income 1	−0.164	−1.28
Income 2	−8.04	−1.31
Income 3	−2.44	−2.00
Gender	0.002	0.34
Age 1	0.514	1.41
Age 2	−3.11	−2.94
Age 3	−3.37	−5.14
Car ownership	0.667	6.58
Employment accessibility	0.418	7.70
Housing accessibility	0.234	7.86
Population density	0.308	1.14
Employment density	−2.18	−2.12
Point density of department stores	−2.17	−1.37
Ratio of jobs to residents	−2.59	−4.26
Nest A	1	9.30
Nest B	1	7.60
Nest C	2	2.10
Number of observations	5 327	
Adjusted rho-square	0.626	

The variables for gender, population density and the point density of department stores in the TAZ where the individuals live are found to be non-significant estimates. Dissimilarity parameters are estimated for each nest, which implies that the alternatives in the nest of the travel distance have high correlations.

Disaggregate direct and cross elasticities concerning the travel time and travel costs are provided in Tab. 4. In this case, the elasticities respond to a change in travel time and travel cost of the alternative commute travel mode. For example, if alternative m corresponds to commute to work by car, the car travel time cross elasticity for alternative m shows the change in the travel mode choice probability of alternative m, due to a 1% change in travel times for taxi at the same distance to work. The direct and cross elasticties are computed for all alternatives presented in Tab. 4 for a random individual traveler.

Tab. 4 Direct and cross elasticities for a random individual

Travel mode	Travel time		Travel cost	
	Direct elasticity	Cross elasticity	Direct elasticity	Cross elasticity
Walk	−2.813 2	0.359 1	0	0
Bicycle	−0.573 2	0.855 1	0	0
Car	−0.311 7	0.046 0	−0.190 6	0.028 1
Taxi	−0.355 6	0.002 1	−1.390 4	0.008 1
Bus	−0.675 5	0.118 9	−0.099 0	0.017 4
Subway	−0.406 4	0.001 7	−0.115 9	0.000 5

Direct time elasticities for walk are the greatest, which manifests that walk travelers are more sensitive to travel time than any other travel mode choice and walk is suitable for short-distance travel. On the contrary, direct time elasticities for car are the smallest, showing that commuting to work by car is the most convenient way. The effects of travel costs on taxi choice probabilities are greater than that on car. In addition, taxi choice probabilities are more sensitive to changes in travel costs than to changes in travel times, while the travel times have more effects on bus and subway choice probabilities than travel costs. This means that measures such as public transport priority help to improve the operating speed of bus and attract more people to commute to work by bus. Cross elasticities for travel cost are so small for all travel mode choices that they have null effect on choice probabilities. Changes in bicycle travel times are found to have the largest cross effects.

4 Conclusion

This paper studies the commute travel mode choice for the main area of Nanjing, with a closed-form discrete choice model specified as an NL model. The estimation results show that the non-motorized travel modes are more attractive in the areas with higher housing and employment accessibility. The results indicate that effective integration of land use and transportation policies help to balance the traffic demand. For example, appropriate land use policies such as improving the housing and employment accessibilities and mixing land use and other measures can help to reduce people's dependence on cars, change travel behaviors and increase the non-motorized trips.

Future research will need to not only extend the research area to citywide and consider the effect of residential location choice on the commute travel mode choice, but also use a more suitable model structure.

References

[1] Handy S, Niemeier D. Measuring accessibility: an exploration of issues and alternatives [J]. *Environ Planning A*, 1997, **29**(7): 1175 − 1194.

[2] Boarnet M G, Sarmiento S. Can land-use policy really affect travel behavior? A study of the link between non-work travel and land-use characteristics [J]. *Urban Studies*, 1998, **35**(7): 1155 − 1169.

[3] Macfadden D. *Conditional logit analysis of qualitative choice behavior* [M]. New York: Academic Press, 1973.

[4] Wen C H, Koppelman F S. The generalized nested logit model [J]. *Transportation*, 2001, **35**(7): 627 − 641.

[5] Yang L Y, Zheng G, Zhu X N. Cross-nested logit model for the joint choice of residential location, travel mode, and departure time [J]. *Habitat International*, 2013, **38**(4): 157 − 166.

[6] Guadagni P M, Little J D C. When and what to buy: a nested logit model of coffee purchase [J]. *Journal of Forecasting*, 1998, **17**(3): 303 − 326.

[7] Ortúzar D J. On the development of the nested logit model [J]. *Transportation Research Part B*, 2001, **35**(2): 213 − 216.

[8] Ben-Akiva M E. Structure of passenger travel demand models [D]. Cambridge, MA, USA: Massachusetts Institute of Technology, 1973.
[9] Ben-Akiva M E. Structure of passenger travel demand models [J]. *Transportation Research Record*, 1974, **526**: 26 - 42.
[10] Sultana S. Effects of married-couple dual-earner households on metropolitan commuting: evidence from the Atlanta metropolitan area [J]. *Urban Geography*, 2005, **26** (4): 328 - 352.
[11] Limtanakool N, Dijst M, Schwanen A T. The influence of socioeconomic characteristics, land use and travel time consideration on mode choice for medium-and longer-distance trips [J]. *Journal of Transport Geography*, 2006, **14** (5): 327 - 341.
[12] Cao X, Mokhtarian P, Handy A S. Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership [J]. *Environment and Planning A*, 2007, **39**(4): 830 - 847.
[13] Peng Z. The jobs-housing balance and urban commuting [J]. *Urban Studies*, 1997, **34**(8): 1215 - 1235.
[14] Litman T. Evaluating accessibility for transportation planning [R]. Victoria, Canada: Victoria Transport Policy Institute, 2008.
[15] Cervero R, Duncan M. Transit's added value [J]. *Urban Land*, 2002, **1805**(2): 77 - 84.
[16] Des Rosiers F, Thériault M, Villeneuve P. Sorting out access and neighbourhood factors in hedonic price modelling [J]. *The Journal of Property Investment and Finance*, 2000, **18**(3): 291 - 315.
[17] Bierlaire M. Biogeme: a free package for the estimation of discrete choice models [C]//*3rd Swiss Transportation Research Conference*. Ascona, Switzerland, 2003: 1 - 24.

土地利用特征与出行属性对交通方式选择的影响：以南京主城区为例

刘志伟　邓　卫　季彦婕

（东南大学交通学院，南京 210096）

摘要：以南京市主城区为例，基于随机效用最大化理论，建立了居民的出行方式选择分层 logit 模型. 研究了个人社会经济属性、城市土地利用特征以及出行属性等因素对居民出行方式选择的影响. 同时，通过计算出行时间与出行费用的直接弹性和交叉弹性分析它们对出行方式选择的敏感程度. 研究结果表明：非机动化的出行方式在居住可达性和就业可达性高的区域更加具有吸引力，小汽车拥有者更加偏向于选择小汽车出行；相对出行费用而言，选择公交与地铁的居民对出行时间更加敏感；完善的公共交通系统以及城市土地利用与交通政策的有效结合有助于缓解由于城市的不断扩张引起的交通拥堵水平.

关键词：交通工程；土地利用；出行方式选择；分层 logit 模型；弹性分析
中图分类号：U121