

Research Article

Integrated Model of Joint Residence-Workplace Location Choice and Commute Behavior Using Latent Class and Mixed Logit Methods

Pengpeng Jiao,¹ Meiqi Liu,² and Jin Guo³

¹Beijing Urban Transportation Infrastructure Engineering Technology Research Center, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

²Beijing Collaborative Innovation Center for Metropolitan Transportation, Beijing University of Civil Engineering and Architecture, Beijing 100044, China

³China Patent Information Center, Beijing 100088, China

Correspondence should be addressed to Pengpeng Jiao; jiaopengpeng@bucea.edu.cn

Received 11 February 2015; Accepted 7 April 2015

Academic Editor: Luca Guerrini

Copyright © 2015 Pengpeng Jiao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

With the rapid development of urbanization and motorization, urban commute trips are becoming increasingly serious due to the unbalanced distribution of residence and workplace land-use types in most Chinese cities. To explore the inherent interrelations among residence location, workplace, and commute trip, an integrated model framework of joint residence-workplace location choice and commute behavior is put forward based on the personal trip survey data of Beijing in 2005. First, to extract households' different choice characteristics, this paper presents a latent class model, clusters all households into several groups, and analyzes the conditional probability of each group. Second, the paper integrates the residence location and workplace together as the joint choice alternative, employs the socioeconomic factors, individual attributes, household attributes, and trip characteristics as explanatory variables, and formulates the joint residence-workplace location choice model using mixed logit method. Estimations of the latent class model show that four latent groups fit the data best. Further results of the joint residence-workplace location choice model indicate that there exist significantly different choice characteristics in each latent group. Generally, the integrated model framework outperforms traditional location choice methods.

1. Introduction

In most Chinese cities, with the rapid development of urbanization and motorization, the density of urban land-use is increasing very fast, and the spatial distribution of residence location and workplace is turning to be unbalanced. As a result, the urban transportation systems, especially the commute trips, are facing more and more serious problems.

During the past two decades, integrated models of urban land-use and transportation systems have been studied extensively, especially the residential location choice models using decision behavior approaches. As a competitive tool, the discrete choice model was used widely in the location choice models. Lerman (1976) [1] introduced household

car ownership, housing type, and mode to work to the residential location choice and formulated a logit model. Freedman and Kern (1997) [2] studied both workplace and residence locations in two-earner households. To make the model closer to the reality, McFadden (1978) [3], Boots and Kanaroglou (1988) [4], Gabriel and Rosenthal (1989) [5], Waddell (1993) [6], Abraham and Hunt (1997) [7], Ben-Akiva and Bowman (1998) [8], Deng et al. (2003) [9], Hunt et al. (2004) [10], Bhat and Guo (2004) [11], Waddell et al. (2007) [12], Jiao and Harata (2007) [13], Vega and Reynolds-Feighan (2009) [14], and Li et al. (2010) [15] presented different models of household residence location or workplace choices, as well as household members' activity and travel schedules. All the above work made use of discrete choice

methods. According to the modeling techniques, existing researches can be classified into the following four categories: multinomial logit model (MNL), nested logit model (NL), generalized extreme value model (GEV), and mixed logit model (ML). The MNL model is the most widespread in the pioneering work due to its simplicity; however, it assumes the independence of irrelevant alternatives (IIA), for instance, Gabriel and Rosenthal (1989) [5]. The NL model allows for the correlations between the alternatives in each “nest”; however, the alternatives in different “nests” are still independent, for example, Boots and Kanaroglou (1988) [4], Abraham and Hunt (1997) [7], Deng et al. (2003) [9], and Hunt et al. (2004) [10]. The GEV model was first developed by McFadden (1978) [3] and was applied into the job-housing balance study by Vega and Reynolds-Feighan (2009) [14] due to its flexible structure and correlated alternatives. The ML model is the most flexible in the structure, and most existing researches have used it to incorporate the random taste variations of different households, as well as the spatial correlations among different land locations.

More recently, Li et al. (2014) [16] presented a multiobjective optimization model of residential distribution using operations research method; Ibeas et al. (2013) [17] and Jiao et al. [18] also proposed some location choice models to formulate spatial interactions between residence location and workplace. Furthermore, Jiao et al. [18] combined the residence location and workplace together as the choice alternative and put forward a joint residence-workplace location choice model. The case study also indicated the advantages of the joint choice model.

With some models focusing on location choice of specific household types (such as two-earner or single worker households), most of the above works have analyzed the location choice behavior using data of all kinds of households. However, there should exist different choice characteristics for different households. Jiao and Harata (2007) [13] clustered households into several groups according to their life styles and analyzed the residential location choice behavior of each group. Nevertheless, the simple clustering analysis method is too straightforward to extract the inherent characteristics of different households.

Fortunately, the latent class model (LCM) has been used in the category analysis widely. Lazarsfeld (1950) [19] first formulated a latent class model. Goodman (1974) [20] developed an algorithm to estimate the model parameters and made the model more applicable in practice. Haberman (1979) [21] analyzed the relations between latent class model and log-linear model. Vermunt (1997) [22] proposed a general latent class model for categorical data analysis with discrete latent variables. Actually, the LCM have been generally used in the researches about health behavior (Lanza and Rhoades, 2013 [23]) and social interactions (Harvey and Taylor, 2000 [24], Arentze and Timmermans, 2008 [25]), but it was rarely applied in the analyses of urban land-use and transportation systems.

To analyze the different residence location and workplace choice characteristics according to household types, one key feature of this paper is to formulate a latent class model and to extract the inherent household groups. Another key

feature is to further combine the residence location and workplace together as the choice alternatives and present the joint residence-workplace location choice models for each latent class using mixed logit methods.

The rest of this paper consists of the following contents. The general model framework is proposed in Section 2, including both the latent class model and the mixed logit model. The latent class model for household clustering is formulated in Section 3. The integrated model of joint residence-workplace location choice is put forward in Section 4 based on the combined choice alternatives. Both models are estimated using the personal trip survey data of Beijing in 2005, and the estimation results are reported and analyzed in Section 5. Conclusions and potential future researches are summarized in Section 6.

2. General Model Framework

The theory of LCM is based on the probability distribution principles and log-linear models, with the objective to explain the interrelations among manifest variables using the least latent categories and to achieve the local independence. The LCM is mainly used to analyze the categorical data. Compared with continuous variables, the biggest difference of categorical variables is that their values are discrete, with each value denoting different attribute or classification, for instance, gender, residence location, trip mode, and so forth.

Mixed logit (ML) model is a kind of discrete choice model. To assume the parameters subject to some random distributions, it is capable of incorporating the random taste variations of different households, as well as the spatial correlations among different land locations. Therefore, the ML model is widely used in location choice researches.

2.1. Structure of Latent Class Model. The LCM is a kind of model to transform the probabilities of categorical variables to some parameters, that is, probabilistic parameterization. There are two kinds of categorical variables in classical LCM: manifest variable and latent variable. Meanwhile, there are two groups of parameters: latent class probability and conditional probability.

The manifest variable can be observed directly, for example, time, distance, and so on. It is also called observable variable or measureable variable. However, the latent variable cannot be observed directly, for instance, psychological expectation, individual preference, and so forth.

A latent class model can be formulated as

$$\pi_{ijk}^{ABC} = \pi_t^X \pi_{it}^{\bar{A}X} \pi_{jt}^{\bar{B}X} \pi_{kt}^{\bar{C}X}, \quad (1)$$

where π_{ijk}^{ABC} is the joint probability of the LCM; A , B , and C denote three manifest variables; π_t^X is the latent class probability, which means the probability of latent variable X in class t , $t = 1, 2, \dots, T$, $\sum_{t=1}^T \pi_t^X = 1$; $\pi_{it}^{\bar{A}X}$, $\pi_{jt}^{\bar{B}X}$, and $\pi_{kt}^{\bar{C}X}$ are three conditional probabilities; $\pi_{it}^{\bar{A}X}$ shows the conditional probability of latent class t on manifest variable A in level i : that is, $P(A = i | X = t)$, $i = 1, 2, \dots, I$; $\pi_{jt}^{\bar{B}X}$ indicates

TABLE I: Variables in the latent class model.

Variable	Symbol	Values
Residence location	R	Xicheng/1, Dongcheng/2, Chongwen/3, Xuanwu/4, Haidian/5, Chaoyang/6, Fengtai/7, Shijingshan/8
Workplace	W	
Commute distance (Km)	D	<4/1, 4–8/2, 8–12/3, 12–16/4, 16–20/5, >20/6
Commute mode	M	Walk/1, Bicycle/2, Bus/3, Subway/4, car/5
Household monthly income (Chinese Yuan)	I	<2500/1, 2501–3500/2, 3501–5500/3, >5500/4

In the third column, the number after the slash line “/” is value of the corresponding categorical variables.

the conditional probability of latent class t on manifest variable B in level j : that is, $P(B = j | X = t)$, $j = 1, 2, \dots, J$; π_{kt}^{CX} denotes the conditional probability of latent class t on manifest variable C in level k : that is, $P(C = k | X = t)$, $k = 1, 2, \dots, K$; $P(\cdot)$ is a probability function.

2.2. Structure of Mixed Logit Model. Similar to MNL model, NL model, and GEV model, the mixed logit model is also based on the assumption of random utility maximization. With rather flexible formulation in the structure, it mainly has the following advantages: (1) there is no IIA property in the model; (2) the error item of the utility function can be subject to any random distribution, which removes the constraint of Gumbel distribution in logit model or normal distribution in probit model; (3) the estimated parameters are subject to some kind of random distribution, which incorporates the taste variations of different decision makers.

Similar to our previous work [18], using U_{hi} to denote the utility function for decision maker h to select alternative i , then it can be divided into two items: the systematic item V_{hi} and the random item η_{hi} ; that is,

$$U_{hi} = V_{hi} + \eta_{hi}. \quad (2)$$

To incorporate random taste variations in the model, V_{hi} is further formulated as below:

$$V_{hi} = \sum_{m=1}^M \alpha_m z_{him} + \sum_{n=1}^N \beta_n x_{hin}, \quad (3)$$

where α_m and β_n are two kinds of parameters to be estimated; α_m is the fixed parameter, just like MNL model; β_n is the unfixed parameter following some random distribution to incorporate the random taste variations; z_{him} and x_{hin} are explanatory variables corresponding to α_m and β_n , respectively; M is the total number of variables corresponding to α_m ; N is the total number of variables corresponding to β_n .

Based on the fundamental theory of discrete choice analysis, the mixed logit model can be formulated as

$$L_{hi} = \frac{\exp(V_{hi})}{\sum_{j=1}^J \exp(V_{hj})} = \frac{\exp(\sum_{m=1}^M \alpha_m z_{him} + \sum_{n=1}^N \beta_n x_{hin})}{\sum_{j=1}^J \exp(\sum_{m=1}^M \alpha_m z_{hjm} + \sum_{n=1}^N \beta_n x_{hjn})}, \quad (4)$$

where L_{hi} is the probability for decision maker h to select alternative i and J is the total number of alternatives.

Therefore, the unconditional probability for decision maker h to select alternative i can be further formulated as

$$P_{hi} = \int_{-\infty}^{\infty} \left[\frac{\exp(\sum_{m=1}^M \alpha_m z_{him} + \sum_{n=1}^N \beta_n x_{hin})}{\sum_{j=1}^J \exp(\sum_{m=1}^M \alpha_m z_{hjm} + \sum_{n=1}^N \beta_n x_{hjn})} \right] \cdot f(\beta) d\beta, \quad (5)$$

where P_{hi} is the unconditional probability for decision maker h to select alternative i and $f(\cdot)$ is the density function which the unfixed parameter β_n follows.

3. Latent Class Model for Household Clustering

To explore the inherent characteristics of urban residence location choice and workplace choice, this paper further formulates the latent class model for commute trips based on the personal trip survey data of Beijing in 2005. The study area is divided into eight zones according to the urban districts: Xicheng, Dongcheng, Chongwen, Xuanwu, Haidian, Chaoyang, Fengtai, and Shijingshan. Based on the thorough analyses of influence factors of residence location and workplace choices, we introduce the following five variables into the LCM: residence location, workplace, commute distance, commute mode, and household monthly income. Here residence location, workplace, and commute mode are discrete variables; however, commute distance and household monthly income are continuous in nature. For convenience, these two continuous variables are also discretized and transformed to categorical variables. For the important commute mode, we mainly select five modes, that is, walk, bicycle, bus, subway, and car.

Variables in the latent class model are summarized in Table I.

Based on the variables in Table I, the latent class model is formulated as

$$\pi_{ijklst}^{RWD MIX} = \pi_t^X \pi_{it}^{\bar{R}X} \pi_{jt}^{\bar{W}X} \pi_{kt}^{\bar{D}X} \pi_{lt}^{\bar{M}X} \pi_{st}^{\bar{I}X}, \quad (6)$$

where X is the latent variable; T is the number of latent classes; π_t^X denotes the latent class probability; R, W, D, M , and I show the manifest variables; i, j, k, l , and s are levels of manifest variables, respectively; $\pi_{it}^{\bar{R}X}$, $\pi_{jt}^{\bar{W}X}$, $\pi_{kt}^{\bar{D}X}$, $\pi_{lt}^{\bar{M}X}$,

TABLE 2: Variables in the mixed logit model.

Variable	Symbol	Remark
House renting price (Hundred Chinese Yuan/Month)	FJ	Continuous
Commute distance (Km)	DI	Continuous
Commute time (Min)	TT	Continuous
Household monthly income (Chinese Yuan)	INC	Discrete: <2500/1, 2501–3500/2, 3501–5500/3, >5500/4
Population density (Thousand/Km ²)	POP	Continuous
Number of employment opportunities (Thousand)	EMP	Continuous
GDP of workplace (Ten Billion Chinese Yuan)	GDP	Continuous

In the third column, the number after the slash line “/” is value of the corresponding categorical variables.

and $\pi_{st}^{\bar{I}X}$ are conditional probabilities; $\pi_{it}^{\bar{R}X}$ is the conditional probability of latent class t on manifest variable R in level i : that is, $P(R = i | X = t)$; $\pi_{jt}^{\bar{W}X}$ is the conditional probability of latent class t on manifest variable W in level j : that is, $P(W = j | X = t)$; $\pi_{kt}^{\bar{D}X}$ is the conditional probability of latent class t on manifest variable D in level k : that is, $P(D = k | X = t)$; $\pi_{lt}^{\bar{M}X}$ is the conditional probability of latent class t on manifest variable M in level l : that is, $P(M = l | X = t)$; $\pi_{st}^{\bar{I}X}$ is the conditional probability of latent class t on manifest variable I in level s : that is, $P(I = s | X = t)$; $\pi_{ijklst}^{RWD MIX}$ is the joint probability of the LCM.

Based on the above symbols and equations, we can further formulate the conditional probability of latent variable X as

$$\pi_{ijklst}^{\bar{R}\bar{W}\bar{D}\bar{M}\bar{I}X} = \frac{\pi_{ijklst}^{RWD MIX}}{\sum_{t=1}^T \pi_{ijklst}^{RWD MIX}}. \quad (7)$$

Using (7), we can obtain the conditional probability of latent variable. Every observation is assigned to the corresponding latent class according to the largest magnitude in the value of $\pi_{ijklst}^{\bar{R}\bar{W}\bar{D}\bar{M}\bar{I}X}$. Therefore, for urban residence location and workplace choice problems, households will be clustered into several groups logically using the above LCM method, and each group will have different choice characteristics.

4. Integrated Model of Joint Residence-Workplace Location Choice

Based on the clustered household groups from the LCM, we can further formulate residence location and workplace choice models for each kind of household using mixed logit method, just like our previous work [18]. Explanatory variables are summarized in Table 2, including house renting price, commute distance, commute time, household monthly income, population density, number of employment opportunities, and GDP of workplace.

In this model, we assume that households make residence location and workplace choices simultaneously, that is, the joint location choice. Since there are eight zones in the study area, totally we have 64 choice alternatives; that is, each residence-workplace location pair denotes one alternative.

Based on many estimation experiments, commute distance and commute time are assumed to be corresponding to the unfixed parameters β_n . Here β_n is assumed to follow a logarithmic normal distribution. Therefore, the integrated model of joint residence-workplace location choice is presented as

$$P_{hi} = \int_{-\infty}^{\infty} \left[\left(\exp(\beta_1 DI_i + \beta_2 TT_i + \alpha_1 FJ_i + \alpha_2 INC_i + \alpha_3 POP_i + \alpha_4 EMP_i + \alpha_5 GDP_i) \cdot \left(\sum_{j=1}^J \exp(\beta_1 DI_j + \beta_2 TT_j + \alpha_1 FJ_j + \alpha_2 INC_j + \alpha_3 POP_j + \alpha_4 EMP_j + \alpha_5 GDP_j) \right)^{-1} \right) \right] \cdot f(\beta) d\beta, \quad (8)$$

where J is the total number of choice alternatives, β_n follows the logarithmic normal distribution in (9), and definitions of other variables are the same as those stated before:

$$\beta_n(\gamma_k, \xi_n, \theta_n) = \frac{1}{2\sqrt{\pi}\theta_n} \exp \left[- \left(\frac{\ln(\gamma_k) - \xi_n}{\theta_n} \right)^2 \right], \quad (9)$$

where γ_k is the random variable and ξ_n and θ_n are the expectation and variance of $\ln(\gamma_k)$, respectively.

Using (8) and (9), we can formulate every household group's joint residence-workplace location choice behavior.

TABLE 3: Fit criteria of the LCM.

Number of latent classes	Likelihood ratio chi-square		AIC	BIC
	Value	P-value		
$T = 1$	13148.723	<0.001	64326.826	64490.171
$T = 2$	11145.600	<0.001	62425.088	62758.060
$T = 3$	9404.705	0.053	60731.082	61233.680
$T = 4$	8127.792	0.097	57506.073	58178.298
$T = 5$	8845.764	0.057	60255.852	61097.705

5. Estimation Results

Based on the personal trip survey data of Beijing, we can estimate the integrated model of joint residence-workplace location choice and commute behavior.

5.1. Estimation Results of Latent Class Model. Parameter estimation of latent class model is usually implemented using two kinds of iterative algorithms based on maximum likelihood method: expectation maximization algorithm and Newton-Raphson algorithm. The iterative process consists of two steps: the first step is to achieve the maximized value from a starting number, which is taken as the initial estimation value in the algorithm and the second step is to estimate again from the result of the first step, until the process arrives at the accuracy requirement.

To obtain the LCM with best-fit ratio, this paper makes use of the maximum likelihood method for estimation. The progress of estimation begins by fitting a complete independence model with $T = 1$ and then iteratively increases the number of latent classes one by one until an appropriate T is achieved. The likelihood ratio chi-square value will increase with the increase of differences between observed values and expected values. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) are employed to evaluate the model and to determine the appropriate number of latent classes T . For the LCM, the fit ratio increases with the decrease of AIC and BIC.

Based on several estimation experiments, the fit criteria of the proposed LCM are summarized in Table 3.

From Table 3, one can find out that the fit criterion of the model is the best while $T = 4$; that is, we totally get 4 latent classes.

The detailed parameter estimation results are further summarized in Table 4.

From Table 4, we can find out the following results.

Class 1. Households mainly reside in Haidian and Chaoyang districts and work in the same zones. Their commute distances are rather long, with 55.3% within 8 kilometers, and 29.6% between 8 and 16 kilometers. The frequently used commute modes are bicycle, bus, and car. Household income is rather high. Furthermore, from the latent class probability, the proportion of this latent class is the biggest.

Class 2. Households mainly reside in Xicheng and Dongcheng districts. Some of them work in zones close to home, while some others tend to work in Haidian and Chaoyang districts.

The possible reasons are that there are many scientific, technological, and cultural zones in Haidian district, and the central business district (CBD) locates in Chaoyang district. With higher income than class 1, their commute distances are also longer than the first household group. Meanwhile, the most frequently used commute mode is car, and the second one is bus.

Class 3. Most households reside in Dongcheng and Chaoyang districts and work in the same area; that is, these households mainly live and work in the east area of Beijing. Therefore, the commute distance is the shortest, with 87.7% within 8 kilometers. Meanwhile, the frequently used commute modes are walk and bicycle, which are most appropriate for short distance trips. Almost no household uses subway; however, the percentage of car is about 20%. The possible reasons are that there was few subway lines under operation when the survey was carried out, and some people tend to use car due to rather high income.

Class 4. Most households reside in Fengtai and Shijingshan districts and work in the same zones. Besides, some households tend to live and work in Chongwen and Xuanwu districts. The percentage is rather small, because there are a large number of historical and cultural protection areas in these two districts, which are not very appropriate for residence. Commute distances of this latent class are rather short, with 85.5% within 8 kilometers. Therefore, corresponding to the distance, the most frequently used commute mode is the bicycle, covering 52.4%. Furthermore, since the economic level of Fengtai and Shijingshan districts are not very high, household income of this group is the lowest in all latent classes.

From the above analyses, we can further find out that the differences among four latent classes are distinct. Moreover, the clustered results from the LCM method are much more logical than those from traditional simple cluster analysis methods [13].

5.2. Estimation Results of Mixed Logit Model. The joint residence-workplace location choice model based on mixed logit is estimated using maximum simulated likelihood (MSL) method, which was proved to be rather effective and efficient by Bhat and Guo (2004) [11]. Furthermore, to implement the MSL algorithm, we integrate randomly scrambled Halton draws [26] into the estimation algorithm. Similar to our previous work, we also code the algorithm using GAUSS platform.

TABLE 4: Estimation results of the LCM.

Latent class	1	2	3	4	Variable values
Residence location					
1	0.057	0.477	0.018	0.000	Xicheng
2	0.052	0.486	0.140	0.000	Dongcheng
3	0.001	0.000	0.019	0.131	Chongwen
4	0.004	0.018	0.000	0.079	Xuanwu
5	0.411	0.002	0.000	0.000	Haidian
6	0.364	0.005	0.821	0.001	Chaoyang
7	0.051	0.004	0.002	0.480	Fengtai
8	0.060	0.008	0.000	0.309	Shijingshan
Workplace					
1	0.092	0.218	0.000	0.024	Xicheng
2	0.092	0.225	0.205	0.004	Dongcheng
3	0.015	0.014	0.011	0.094	Chongwen
4	0.016	0.037	0.000	0.106	Xuanwu
5	0.306	0.167	0.000	0.002	Haidian
6	0.440	0.278	0.784	0.000	Chaoyang
7	0.030	0.048	0.000	0.469	Fengtai
8	0.009	0.013	0.000	0.301	Shijingshan
Commute distance					
1	0.215	0.083	0.377	0.332	<4
2	0.338	0.406	0.500	0.523	4–8
3	0.158	0.256	0.086	0.111	8–12
4	0.138	0.125	0.023	0.012	12–16
5	0.070	0.068	0.014	0.017	16–20
6	0.081	0.062	0.000	0.005	>20
Commute mode					
1	0.097	0.031	0.198	0.133	Walk
2	0.307	0.214	0.439	0.524	Bicycle
3	0.235	0.286	0.165	0.114	Bus
4	0.061	0.063	0.000	0.000	Subway
5	0.300	0.406	0.199	0.229	Car
Household monthly income					
1	0.156	0.068	0.158	0.308	<2500
2	0.250	0.198	0.269	0.330	2501–3500
3	0.313	0.406	0.269	0.341	3501–5500
4	0.281	0.328	0.305	0.021	>5500
Latent class probability	47.55%	11.00%	21.30%	20.15%	

Using the above method, the joint residence-workplace location choice model is estimated as though that all households tend to choose residence location and workplace simultaneously; that is, these two kinds of land-use types influence each other. The estimated mean values of all parameters are

reported in Table 5, with the t -statistics shown in parentheses to indicate the significances of explanatory variables. As stated before, for unfixed commute distance and commute time, estimations of their mean values and standard deviations are both reported. Furthermore, to compare different household groups, the estimations of four latent classes are summarized in four columns, respectively.

From Table 5 we can find out that all estimated parameters have expected signs and significance, which generally proves the effectiveness of the integrated model of joint residence-workplace location choice and commute behavior.

For all latent classes, we can get the following results from the signs of parameters.

- (1) The expected negative sign of house renting price shows that with other conditions fixed, households tend to live in areas with rather low housing price.
- (2) Both commute distance and commute time between residence location and workplace have negative signs as expected, which indicates that households tend to job-housing balance when they consider their residence location and workplace choices; that is, proximity to workplace is very important for households to choose residence location; at the same time, proximity to residence location is also very important for households to choose workplace.
- (3) The positive sign of household monthly income is also consistent with expectation, which means that households are more likely to reside or work in places which could bring them higher income.
- (4) The sign of number of employment opportunities is positive, showing that job opportunities are a rather important factor influencing households' residence location and workplace choices. It means that people tend to live and work in locations with more opportunities.
- (5) GDP of workplace has the expected positive sign, indicating that households are more inclined to work in places with good economic environment.
- (6) A very interesting thing is that there is an exception in population density; that is, in latent class 1, the parameter is positive, while in latent classes 2, 3, and 4, the parameters are negative. Characteristics of each latent class could explain such exception. In class 1, most households live in Haidian and Chaoyang districts, which are two rather big zones with many residential land-uses, but the residence density is not very high; therefore, households tend to locate in places with high population density, which is also a kind of reflection of population clustering effect. Conversely, in classes 2, 3, and 4, most households live in other six districts of Beijing, which are rather small areas with very high residence density; therefore, households in these three classes tend to reside in areas with low population density, which reflects that low residence density and comfortable community environment are more important for these people.

TABLE 5: Estimation results of four latent classes.

Variables	Class 1	Class 2	Class 3	Class 4
House renting price	-0.9568 (-18.48)	-0.1823 (-2.40)	-0.1734 (-4.41)	-0.2153 (-8.60)
Commute distance (M)	-0.1953 (-12.63)	-0.1174 (-3.72)	-0.5049 (-10.10)	-0.2662 (-9.83)
Commute distance (S.D.)	0.3115 (14.49)	0.1762 (5.06)	0.2878 (6.65)	-0.1417 (-4.72)
Commute time (M)	-0.03623 (-2.37)	-0.0305 (5.22)	-0.0128 (-2.89)	-0.0582 (4.96)
Commute time (S.D.)	-1.14E - 04 (0)	-0.0139 (-1.02)	7.69E - 05 (0)	-2.10E - 04 (0)
Household monthly income	1.0919 (24.82)	2.0387 (15.55)	1.0537 (13.47)	1.3504 (16.82)
Population density	0.9441 (18.36)	-0.9983 (-4.24)	-0.1624 (-15.63)	-0.0763 (-6.01)
Number of employment opportunities	0.0648 (10.31)	0.0665 (4.74)	0.1113 (3.66)	0.2311 (4.67)
GDP of workplace	0.3617 (14.24)	0.2288 (9.06)	0.3973 (6.60)	0.1594 (3.40)

Here M means "mean value"; S.D. means "standard deviation."

Further comparisons of the estimations among four latent classes reveal the following results.

- (1) For all 4 latent classes, the magnitude of household monthly income is much bigger than other parameters. It indicates that this factor is much more important than other factors for household residence location and workplace choices. Moreover, house renting price also has rather big magnitude, showing that housing price is also a very important factor in location choices.
- (2) For latent class 1, house renting price, population density, and GDP of workplace have much bigger magnitudes, showing that these three factors are more important for households in Haidian and Chaoyang districts to make their residence location and workplace choices.
- (3) For latent class 2, much bigger magnitudes in population density and GDP of workplace again prove their importance. As stated before, the sign of population density is negative, because most households in class 2 reside in Xicheng and Dongcheng districts, which locate at the center city of Beijing, with very high residence density. Therefore, different from class 1, these households tend to live in areas with low population density and comfortable environment.
- (4) For latent class 3, commute distance has rather big magnitude, which means that people consider more about commute distance when they make residence location and workplace choices. Results from the latent class analyses give the reason; that is, most households in this group use walk and bicycle in commute trips, and these two kinds of modes are more sensitive on trip distance.

- (5) For latent class 4, the magnitude of number of employment opportunities is obviously bigger than others. Once again, the reason can be achieved from the latent class analyses. Most households in this group reside and work in Fengtai and Shijingshan districts, which are relatively underdeveloped in economic level. There are less employment opportunities in these two districts than other six zones, and income is also rather low. Therefore, households in this group pay more attention to number of employment opportunities in residence location and workplace decision behaviors.

Generally, all the estimation results are consistent with expectations. The detailed analyses based on latent classes explore many interesting and logical results.

6. Conclusions

This paper addresses an integrated model of joint residence-workplace location choice and commute behavior using latent class and mixed logit methods. The general model framework consists of two single models. We first present a latent class model to extract households' different choice characteristics and cluster households into several groups. Based on the latent class analyses, we further combine the residence location and workplace together as the joint choice alternative and formulate a joint residence-workplace location choice model using mixed logit method. A large amount of data is extracted from the personal trip survey data of Beijing in 2005 for case study. Estimation results of the latent class model show that households are properly clustered into four groups, and every kind of household has different characteristics. The mixed logit models for all four latent

classes are then estimated, respectively, using maximum simulated likelihood method. Estimated parameters show that all the estimations are consistent with expectations. For all latent classes, household monthly income and housing price are much important for residence location and workplace choices. Further comparisons of the estimated parameters among four latent classes prove that there exist much big differences in the location choice behaviors, and the joint residence-workplace location choice model using latent class and mixed logit methods is very effective.

Future researches are directed towards the following aspects. The first is to employ more recent socioeconomic data, census data, and trip survey data and to update the case studies of this research. The second is to further explore the differences among different decision makers, for instance, male, female, and children in the same household and to analyze more detailed choice behaviors. The third is to track the development histories of residential and employment land-uses based on panel data.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

This research has been supported by National Natural Science Foundation of China Project (51208024), Beijing Philosophy and Social Science Project (14CSC014), Beijing Nova Programme (Z151100000315050), Science and Technology Project of Ministry of Housing and Urban-Rural Development of China (2013-K5-6), and the Importation and Development of High-Caliber Talents Project of Beijing Municipal Institutions (CIT&TCD201404071).

References

- [1] S. R. Lerman, "Location, housing, automobile ownership, and mode to work: a joint model," *Transportation Research Record*, no. 610, pp. 6–11, 1976.
- [2] O. Freedman and C. R. Kern, "A model of workplace and residence choice in two-worker households," *Regional Science and Urban Economics*, vol. 27, no. 3, pp. 241–260, 1997.
- [3] D. McFadden, "Modeling the choice of residential location," *Transportation Research Record*, no. 672, pp. 72–77, 1978.
- [4] B. N. Boots and P. S. Kanaroglou, "Incorporating the effects of spatial structure in discrete choice models of migration," *Journal of Regional Science*, vol. 28, no. 4, pp. 495–509, 1988.
- [5] S. A. Gabriel and S. S. Rosenthal, "Household location and race: estimates of a multinomial logit model," *The Review of Economics and Statistics*, vol. 17, no. 2, pp. 240–249, 1989.
- [6] P. Waddell, "Exogenous workplace choice in residential location models: is the assumption valid?" *Geographical Analysis*, vol. 25, no. 1, pp. 65–82, 1993.
- [7] J. E. Abraham and J. D. Hunt, "Specification and estimation of a nested logit model of home, workplace and commuter mode choice by multiple worker households," *Transportation Research Record*, no. 1606, pp. 17–24, 1997.
- [8] M. E. Ben-Akiva and J. L. Bowman, "Integration of an activity-based model system and a residential location model," *Urban Studies*, vol. 35, no. 7, pp. 1131–1153, 1998.
- [9] Y. Deng, S. L. Ross, and S. M. Wachter, "Racial differences in homeownership: the effect of residential location," *Regional Science and Urban Economics*, vol. 33, no. 5, pp. 517–556, 2003.
- [10] L. M. Hunt, B. Boots, and P. S. Kanaroglou, "Spatial choice modelling: new opportunities to incorporate space into substitution patterns," *Progress in Human Geography*, vol. 28, no. 6, pp. 746–766, 2004.
- [11] C. R. Bhat and J. Y. Guo, "A mixed spatially correlated logit model: formulation and application to residential choice modeling," *Transportation Research Part B: Methodological*, vol. 38, no. 2, pp. 147–168, 2004.
- [12] P. Waddell, C. Bhat, N. Eluru, L. Wang, and R. M. Pendyala, "Modeling interdependence in household residence and workplace choices," *Transportation Research Record*, no. 2003, pp. 84–92, 2007.
- [13] P. Jiao and N. Harata, "Residential location choice for different households: methodology and case study," *Journal of the Eastern Asia Society for Transportation Studies*, vol. 7, pp. 1280–1295, 2007.
- [14] A. Vega and A. Reynolds-Feighan, "A methodological framework for the study of residential location and travel-to-work mode choice under central and suburban employment destination patterns," *Transportation Research Part A: Policy and Practice*, vol. 43, no. 4, pp. 401–419, 2009.
- [15] X. Li, C. Shao, T. Qu, L. Yang, and J. Wang, "Application of network generalized extreme value model to estimate residential location and commute mode choice," *Acta Scientiarum Naturalium Universitatis Pekinensis*, vol. 46, no. 6, pp. 926–932, 2010.
- [16] X. Li, C. Zhuge, X. Zhang, J. Gao, and H. Zhang, "Multiobjective optimization model of residential spatial distribution," *Mathematical Problems in Engineering*, vol. 2014, Article ID 167495, 9 pages, 2014.
- [17] Á. Ibeas, R. Cordera, L. Dell'Olio, and P. Coppola, "Modelling the spatial interactions between workplace and residential location," *Transportation Research Part A: Policy and Practice*, vol. 49, pp. 110–122, 2013.
- [18] P. Jiao, T. Sun, J. Guo, and Y. Li, "Joint residence-workplace location choice model based on household decision behavior," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 908387, 8 pages, 2015.
- [19] P. F. Lazarsfeld, "The logical and mathematical foundation of latent structure analysis," in *Studies in Social Psychology in World War II, Volume 4: Measurement and Prediction*, E. A. Schulman, P. F. Lazarsfeld, S. A. Starr, and J. A. Clausen, Eds., pp. 362–412, Princeton University Press, Princeton, NJ, USA, 1950.
- [20] L. A. Goodman, "The analysis of systems of qualitative variables when some of the variables are unobservable. Part I-A modified latent structure approach," *The American Journal of Sociology*, vol. 79, no. 5, pp. 1179–1259, 1974.
- [21] S. J. Haberman, *Analysis of Qualitative Data*, Academic Press, New York, NY, USA, 1979.
- [22] J. K. Vermunt, *Log-Linear Models for Event Histories*, Sage Publications, Thousand Oakes, Calif, USA, 1997.
- [23] S. T. Lanza and B. L. Rhoades, "Latent class analysis: an alternative perspective on subgroup analysis in prevention and treatment," *Prevention Science*, vol. 14, no. 2, pp. 157–168, 2013.

- [24] A. S. Harvey and M. E. Taylor, "Activity settings and travel behaviour: a social contact perspective," *Transportation*, vol. 27, no. 1, pp. 53–73, 2000.
- [25] T. Arentze and H. Timmermans, "Social networks, social interactions, and activity-travel behavior: a framework for microsimulation," *Environment and Planning B*, vol. 35, no. 6, pp. 1012–1027, 2008.
- [26] C. R. Bhat, "Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences," *Transportation Research Part B*, vol. 37, no. 9, pp. 837–855, 2003.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

