Research Article

An Application of Hierarchical Structure Model for Trip Mode Choice Forecasting in China

Zhuping Zhou, 1 Wei Wang, 2 and Qizhou Hu 1

1 Department of Traffic Engineering, Nanjing University of Science and Technology, Nanjing 210094, China
2 School of Transportation, Southeast University, Nanjing 210096, China

Correspondence should be addressed to Zhuping Zhou; jzp19850926@163.com

Received 22 July 2014; Accepted 16 October 2014

Copyright © 2015 Zhuping Zhou 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.

Trip mode split is the result of interrelated and mutually independent factors, such as city scale, urban form, economic level, trip distance, and travel time. In order to analyze the formation of traffic structure, it is necessary to make a comprehensive study on the mechanism of these factors and obtain the basic causal relationship of them. Based on this, by using the hierarchical structure model in system engineering, this paper firstly clarifies the logical relationship of different factors. Then, the existing trip survey data of several cities is used to establish the mathematical relationship of various factors of the structure model. Finally, the mode choice forecasting method is proposed based on the structure model of influencing factors. The case study result of six cities shows small bias, indicating that the proposed method is of great practical value. Policy makers can use the results to discover the trip structure feature and grasp the direction of transportation development policy.

1. Introduction

A growing issue for many large cities in China is the increasingly severe traffic congestion. As the problem appears with the advent of private car age, policy makers want to solve traffic congestion problems from the aspect of trip mode. Mode split differs largely in various cities because of the different city characteristics (e.g., urban form, city scale, and population). From this point, finding out the effects of city characteristics on mode choice is important for policy makers to draw up warranted policies and measures.

1.1. Basic Forecasting Models. Trip mode split refers to the number of traveling people shared by various trip modes within or between traffic zones. Trip mode split forecast is an important part of traffic demand forecast. The commonly used forecasting models are diversion curve model, probability model, regressive model, and so on. In addition, discrete choice model has already entered practical stage. Usually, discrete choice used for transportation system analysis is based on random utility theory [1, 2]. Recently, a different approach to choice analysis, based on artificial neural network model, has been proposed by several researchers [3, 4].

1.2. Factors for Mode Choice. Recently, researches in China mainly focus on the trip modes which meet the requirement of ecotransportation and sustainable development, while studies in other countries vary widely. Bergström and Magnusson emphatically discussed the influence mechanism of external factors such as natural conditions, architectural environment, and urban form. Comparatively speaking, the study of internal factors is more prevalent [5]. Festa et al., Shannon et al., and Schmöcker et al. studied the trip mode choice by analyzing selection characteristics of travelers with different personal attributes (e.g., gender, age, and profession) [6–8]. Sunitiyoso and Matsumoto applied an agent-based approach to modeling a social dilemma of travel mode choice considering psychological and sociological aspects [9]. Scheiner reported findings based on longitudinal analysis of the German nationwide travel survey during the period of 1976–2002, focusing on travel mode choice that was subdivided by distance categories and also taking car availability and city size into account [10]. A series of disaggregate mode choice models based on data from 550 workers in Chennai City are developed to capture the effects of vehicle types, vehicle ownership levels, socioeconomic characteristics, and
subjective factors on mode choice [11]. In addition, Verplanken et al. studied the effects of habit discontinuity and self-activation on mode choice [12].

Furthermore, the discussion of public transit share rate is an important issue both in China and abroad. For instance, Thøgersen examined people's attitude towards public transit to gain a way of switching the car mode to bus [13]. Martens discussed the use of bike-and-ride in three countries (the Netherlands, Germany, and the UK) with widely different bicycle cultures and infrastructures [14]. Hensher and Rose used state-of-the-art stated choice designs to parameterize modal choice models for commuting and noncommuting travel futures in the presence of new public transit infrastructure [15].

According to the descriptions above, most of the studies focus merely on a certain point rather than comprehensively considering the integrated transportation system. However, this paper analyzes the mode choice systemically based on the influence of various factors in Section 2.1.1. The inherent change law of traffic structure will be ignored if we only analyze the regression equation between mode split and the related factors. Hence, the structure model should be applied to clarify the logical relationship of various factors.

1.3. Hierarchical Structure Model. Hierarchical structure model is one of the system structure models, which was first introduced by Prof. John N. Warfield (the United States, 1973), is used to describe the relationship among various components as well as that between the system and the environment. The modeling approach firstly identifies the relationship of various elements (causal relations, sequence relations, affiliation, subjection relation, etc.) and then builds the system structure model, that is, abstracting a complex problem using the structure relational model which is widely used to analyze social, economic, environmental, management systems and to provide a scientific basis for system planning [16]. In the field of transportation, the system structure model showed its advantages in examining the influencing factors. For example, Fujii and Kitamura used a system structure model to forecast the latent demand effects of the opening of new freeways. The model was applied to determine the effects of commute duration and scheduling factors on after work discretionary activities and their trips [17]. Yan and Wang analyzed the problems of urban transport through urban transport system modeling and then found the key factors affecting urban transport [18]. Recently, researchers have focused on the structure models with latent variables for its high efficiency [19]. So the structure model with latent variables will be used in this paper.

1.4. Research Objectives. The transportation system involves many different elements, of which both the correlation and the structure are not entirely clear. Thus for further analysis, the aims of this paper are to (a) clarify the relationship of various elements through indirect relationship, (b) then establish the structural model, and (c) finally fit the model step by step using the survey data and propose a method based on influencing factors for trip mode split.

The rest of the paper is organized as follows. Section 2 presents the modeling methodology, in which the factors on mode split are described and the multilevel hierarchy model is derived. Section 3 estimates the regression model about the logic relation among different factors by using the travel survey data. Section 4 applies the result to several cities in China. Section 5 concludes the paper and outlines possible future works.

2. Methodology

The aims of this study are to (1) examine the interactions between trip mode choice and the influencing factors; (2) propose a sketch method for mode split based on the interactions; and (3) determine urban factors that influence mode split and address these factors in policies related to resident trip.

2.1. Research Design. According to the analysis above, we should firstly build the structure model, then develop the mode split model by data fitting, and finally analyze the variable effects.

2.1.1. Factors Associated with Mode Choice. From the system view, the traffic mode structure is influenced by various factors which can be classified into internal influence factors and external influence factors. The internal factors include human factors, vehicle and facility factors (e.g., rate of private vehicle, public transit facility, and road network layout and hierarchy), and trip factors (e.g., travel distance, purpose, and time), while external factors include urban characteristics (e.g., city scale, urban form, land-use pattern, and economic level), environmental factors (e.g., natural condition, ecology, energy consumption, and land resources), and policy factors (e.g., social-economic policies and transport policies concerning management, technology, etc.).

Natural conditions can affect the trip mode choice from the following points: (1) natural barriers such as the gulf, rivers, lakes, and mountains will block transport routes or change the road network form. (2) It is difficult for bikes to adapt to the hilly area because of the large slope. (3) Regions with bad climatic conditions, such as extremely cold and high altitude plateau cities, are not suitable for bicycle trip. In addition, the city scale depends mainly on local natural conditions, as suitable natural condition is an important reason for city formation [20]. Likewise, local geographical conditions are the major constraints on the road network layout, so the natural environment can also affect the transport facilities configuration.

Traffic demand and its spatial distance distribution are determined by city scale, because trip distance elongates with the increase of urban land use, which inevitably leads to the reduction of walk and the growth of bus. The inverse relation between the city scale and the per capita trip times shows that the per capita trip times will decrease with the expansion of city scale. According to the travel survey in the cities with different population scale, the share rate of walk is higher in the medium and small cities, but relatively lower in large cities.
Urban form refers to the form of urban space pattern or the external shape, which mainly includes the following patterns: single central type (such as Beijing and Tianjin city), multicenter group type (such as Wuhan city), multicenter zonal type (such as Chongqing), zonal axial type (such as Dalian and Lanzhou), and star-shaped and other types. The increase of average trip distance in the ribbon cities is bound to induce the switch from nonmotorized transport to motorized transport mode [21]. In condition of the same city scale, residents in the zonal axis-type city prefer to choose walk and bus for their trip than that in the cluster centralized city. However, things are opposite for the trend of bicycle trip mode share rate. In the multicenter group cities, trips should be supported by the mass transit system such as rail transit, which is also suitable for the layout of zonal cities along transport corridors.

Urban economy, the social development level, and the urban modernization degree all influence urban traffic demand and supply, and different transport policies would encourage or restrict a certain trip mode. The impact of urban economy includes the following: (a) the rate of cars, motorcycles, and other private vehicles increases along with the development of economy; (b) the scale of city transportation development is directly or indirectly affected by the economy, and the traffic demand increases due to the development of economy, and (c) the investment on urban transportation infrastructure rises as urban economy level increases, which will in turn improve the bus service and induce the bus trip.

The configuration of transportation facilities can affect the mode choice intensively in that sidewalk continuity, sidewalk width, presence of cycling, and walk paths will have great impact on the nonmotorized mode choice [22].

Generally speaking, travelers have a very strong perception of time when choosing trip mode. In short-distance travel, especially in urban internal travel, they will first consider the impact of travel time. Different traffic modes are suitable for trips of different distance. Therefore, the accumulation process of travelers in space is bound to promote different traffic modes showing different range of trip distance in space. The increased rate of private vehicle ownership will also directly lead to the augment of car mode share rate. The roads, as carriers of the passenger transportation, also have a direct impact on the mode split.

In summary, the interactive relationships of various factors are considered in Figure 1.

2.1.2. Structural Model Building. The corresponding computation of the relationship of the influencing factors (in Figure 1) would be performed as follows. We firstly set up the adjacency matrix according to the directed graph, from which the accessibility matrix could be gained through relevant calculations, then decompose the accessibility matrix, and finally establish multilevel hierarchy of the influencing factors for trip mode.

2.1.3. Calculation of Causal Relationship Matrix


Then set up the adjacency matrix, which is denoted by $R$, and describe the direct relationship between every two factors in the system. The elements of $R$ are defined as

$$R_{ij} = \begin{cases} 1 & \text{accessible from } V_i \text{ to } V_j \\ 0 & \text{no route between } V_i \text{ and } V_j \end{cases}$$

(1)

The results are as follows:

$$R = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ \end{bmatrix}.$$  

(2) Accessibility Matrix Calculation. The accessibility matrix describes the factors’ accessible degree of the system through a certain length of path and it can be obtained by certain computing with the sum of adjacency matrix and unit matrix. Simplifying assumptions are

$$R_1 = (R + I)^1, R_2 = (R + I)^2, \ldots, R_n = (R + I)^n.$$  

(3)

Equation (3) could be calculated to $(R + I)^p = (R + I)^{p+1}$, following the rules of Boolean algebra: $0 + 0 = 0, 0 + 1 = 1, 1 + 1 = 1, 0 \times 0 = 0, 0 \times 1 = 0, \text{ and } 1 \times 1 = 1$. Take matrix $R$ to operation in accordance with such rules to get

$$R^2 = R^3 = M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \end{bmatrix}.$$  

(4)

(3) Accessibility Matrix Decomposition. Accessibility matrix decomposition consists of two steps, namely, demarcating the region and decomposing interlevel. Decomposition algorithm can be described as follows: firstly constituting an accessible set $X(V_i)$ with the factors relevant to $V_i$, and then
According to Table 1, only element 1 satisfies this requirement, so we confirm $L_1 = \{V_1\}$ as the first level and then remove element 1 from the table. Continue calculating; the second level is $L_2 = \{V_2, V_3, V_5, V_6\}$. Similarly, calculate the third and the fourth level. Decomposition results of the interlevel between two and three and between three and four are omitted here. Upon this, rearrange the row vector of the matrix according to the order of $\{1, 2, 3, 5, 6, 4, 7, 9, 10, 8, 11\}$, and then rearrange the column vector likewise to obtain the shifted accessibility matrix

$$
M' = \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 
\end{bmatrix}.
$$

2.1.4. Multilevel Hierarchy Model Building. Minus identity matrix from the shifted accessibility matrix $M'$, we can see from the definition of accessibility that (1) the second level $V_2, V_3, V_5, V_6$ only relates to the first level $V_1$, namely, $V_2 \rightarrow V_1, V_3 \rightarrow V_1, V_5 \rightarrow V_1, V_6 \rightarrow V_1$; (2) removing the column and row whose $V_1$ is in to establish a relationship between the second and the third level, the third level includes $V_3, V_7, V_9, V_{10}$ whose mutual influence relationships are $V_4 \rightarrow V_2, V_4 \rightarrow V_3, V_7 \rightarrow V_5, V_7 \rightarrow V_6, V_7 \rightarrow V_5, V_9 \rightarrow V_2, V_9 \rightarrow V_5, V_{10} \rightarrow V_2, V_{10} \rightarrow V_5$, and $V_{10} \rightarrow V_5$; (3) removing rows and columns whose $V_2, V_3, V_5, V_6$ are in, $V_8, V_{11}$ exist in the fourth grade and the relationships are $V_8 \rightarrow V_2, V_8 \rightarrow V_9, V_8 \rightarrow V_4, V_8 \rightarrow V_{10}$, and $V_{11} \rightarrow V_{10}$. Based on the analysis above, we can establish a four-level hierarchical model as shown in Figure 2.
2.2. Data Source. The practicability and significance of model study would be limited without a large amount of detailed data which can acquire a more precise mathematical model of interrelationship between all levels. The primary data used in the current analysis is drawn from transportation planning reports, which are done by the School of Transportation, Southeast University. These data include urban travel survey data and statistical information of cities in different sizes and patterns, such as Shenyang, Shenzhen, and Suzhou. Part of the statistical results is shown in Table 3. As our research is a macroscopic study, the external factors, particularly urban factors, are the most important ones we considered. The detailed process of data extracting will be carried out in Section 2.3.2.

2.3. Analytical Method

2.3.1. Multiple Regression Analysis. The structural model not only clarifies the relationship between various factors but also simplifies the method for system analysis. The calculation results of the first level variable can be directly got if we know the relationship of variable values between the fourth level and the other levels. Furthermore, the fitting process becomes easier because the structure model eliminates interaction of the same level variables as well as neglecting the linear correlation between variables of the same level.

Regression models are used to examine the mathematical interactions between different levels. For example, we can use the survey data to fit a model on the relationship between trip distance and urban forms, city scale, and travel purpose.

2.3.2. Variable Specification. Variables are key elements in the analysis process, so it is essential to select the variable indicators. In order to complete the analysis, it is better to get indicators from the existing data. According to relevant research data, the variable factors are determined as follows.

(1) Urban Form $C$. Urban form is generally simplified as group shape or ribbon for quantizing. To better distinguish the forms of different cities and achieve higher prediction accuracy, we describe the urban form by adopting the calculation method in urban planning, that is, to select the compact degree as quantitative indicator for the urban form shown as

$$C = \frac{A}{\pi R^2},$$

where $A$ is the area of built-up section of the city ($\text{km}^2$). Urban built-up area means the well developed and constructed area, which has public utilities and infrastructure in the urban administration region. $\pi R^2$ is the area of the city’s smallest circumcircle ($\text{km}^2$), and $R$ is radius of the city’s smallest circumcircle ($\text{km}$) [23].

(2) City Scale $S$. City scale can be stated by different types of indicators, of which the most commonly used are the urban population scale and geographical scale. In the research field related to passenger traffic, the most commonly used is the population scale. Two indicators available in China are urban population and urban nonagricultural population. However, in the view of integrity and comprehensiveness, the urban population is still taken as indicator although the nonagricultural population is the main user of the urban infrastructure [21].

(3) Trip Purpose $I$. The travel purposes include work, school, business, shopping, entertainment, relative-visit, and going home (with return). The existing research mainly focused on the difference between elastic and nonelastic travel. In this paper, the variable of trip purpose will be expressed by the structure of trip purpose, and the indicator for the trip purpose is defined as

$$I = \frac{I_{\text{elastic}}}{I_{\text{nonelastic}}},$$

where $I_{\text{elastic}}$ and $I_{\text{nonelastic}}$ are proportions of elastic travel and nonelastic travel (work, school, and backhaul which means backing home from working or schooling place), respectively (%).
(4) Policy Factors $P$. The city’s policy factor is nonquantitative, but it must be quantified as a parameter. The quantification thinking is to start from one aspect of the qualitative variable. Cities in this paper are divided into 5 groups with reference data for quantification, respectively, according to the public transit priority development policy, as shown in Table 2. On the basis of this classification, we can adopt values in the corresponding range. Besides, cities can be classified according to other aspects such as the availability of bus lane and transit network density if policy-related expatiation cannot be found in the status information.

(5) Facility Configuration $R$. Transport infrastructure includes road network, the number of arteries, bus number in the transit depots, and public traffic facilities [21]. Road traffic, traveling on the urban road, is greatly influenced by road network layout and hierarchical structure [23]. This paper selects road network density as the analysis indicator.

(6) Private Car Ownership Rate $O$. The private car ownership rate imposes on the traffic structure mainly through the private mechanization level. Since motor vehicles ownership is easier to be gained than that of private cars in current archive, we select the motor vehicles ownership per capita (i.e., the number of motor vehicles owned by each urban population) as the analysis indicator.

(7) Travel Time $T$. We select the average travel time in traffic survey data as $T$.

(8) Trip Distance $D$. There are considerable difficulties in investigating the trip distance directly in resident travel survey, as it is difficult to obtain distance data of most trip modes except for the car which can inform the exact driving distance from its odometer. Most of the domestic cities estimate the trip distance through the average travel time and travel speed of different trip modes, as illustrated in formula (12). According to the basic data from related urban travel survey reports, we assume the average speed of traffic modes including walk, nonmotorcycles, and public transit and personal transport as 3.6 km/h, 9 km/h, 14.4 km/h, and 27 km/h, respectively. Consider

$$D = \sum_i T_i V_i f_i,$$

where $T_i$ is the average travel time-consuming of mode $i$ (s); $V_i$ is the average travel speed of mode $i$; and $f_i$ is the share rate of mode $i$. According to the above paragraph, the average velocity $V_i$ for walk, bicycle, bus, and car is 1 m/s, 2.5 m/s, 4 m/s, and 7.5 m/s, respectively.

(9) Mode Share Rate $M$. Mode share rate could be obtained from the traffic survey data directly. For the convenience of study, we classify the domestic multifarious urban traffic mode into walk, nonmotorcycles (including motorcycle and bicycle), public transit (including bus and company commute bus), and personal transport (including private car, social vehicle, and taxi). The reason for merging the motorcycle into bicycle is that motorcycle is an alternative to bicycle from the point of trip distance and role. From the aspect of occupying road resources and capacity, taxi mode can be classified into personal transport mode. This paper does not consider urban rail transit as it is not common in the surveyed cities; therefore public transit mainly means conventional public transit.

Statistical data of urban and trip characteristics are shown in Table 3.

### 3. Results and Discussion

Considering that the complicated natural conditions are difficult to quantify, it is better to work up mathematical models from the third level. The modeling process is twofold. Firstly, we provide an accurate prediction model that is calibrated by travel survey data for mode choice. Secondly, we test the model precision through a detailed validation based on external data.

#### 3.1. Regression Model for the Second Level and the Third Level

Firstly, set up the relationship between trip distance $D$ and urban form $C$, city scale $S$, and trip purpose $I$ as well as the relationship between travel time $T$ and city scale $S$ and travel purpose $I$ and then establish a model about facility configuration $R$ and urban form $C$ and policies $P$; lastly discuss the interaction between private car ownership rates $O$ and policy factors $P$.

(1) Data Fitting $D = F_1(C, S, I)$. Consider

$$D = 7.853 \times 10^{-2} \sqrt[3]{S} + 0.157C - 0.208 \exp(I) + 2.269 \left(R^2 = 0.736\right).$$

$R^2$, known as the coefficient of determination, is a comprehensive measurement for the fitting degree of regression model. The value range of $R^2$ is $0 \leq R^2 \leq 1$ and the closer it is to one, the more accurate the model is. The precision coefficient $R^2$ is not very satisfying as it is not very close to one, which is influenced by the complexity of cities, data errors, and the accuracy of fitting. The model, though not perfect as we thought, reflects the statistic data outcome of many cities after all and can be used as an analytical tool in that it is a manifestation of certain data rules.

### Table 2: Policy intensity of prioritizing public transportation.

<table>
<thead>
<tr>
<th>Policy of prioritizing public transportation (policy)</th>
<th>Very strong</th>
<th>Strong</th>
<th>Moderate</th>
<th>Weak</th>
<th>None</th>
</tr>
</thead>
<tbody>
<tr>
<td>Value</td>
<td>0.75–1</td>
<td>0.5–0.75</td>
<td>0.25–0.5</td>
<td>0–0.25</td>
<td>0</td>
</tr>
<tr>
<td>City</td>
<td>Survey year</td>
<td>City scale</td>
<td>Urban form</td>
<td>Facility Configuration</td>
<td>Total Population</td>
</tr>
<tr>
<td>-----------</td>
<td>-------------</td>
<td>------------</td>
<td>------------</td>
<td>------------------------</td>
<td>------------------</td>
</tr>
<tr>
<td>Bengbu</td>
<td>2002</td>
<td>77.16</td>
<td>0.417</td>
<td>7.06</td>
<td>340.3</td>
</tr>
<tr>
<td>Huzhou</td>
<td>2005</td>
<td>107.72</td>
<td>0.438</td>
<td>7.4</td>
<td>271.8</td>
</tr>
<tr>
<td>Qinhuangdao</td>
<td>2007</td>
<td>73.97</td>
<td>0.43</td>
<td>4.47</td>
<td>273.3</td>
</tr>
<tr>
<td>Shangyu</td>
<td>2006</td>
<td>20.49</td>
<td>0.543</td>
<td>5.28</td>
<td>77.28</td>
</tr>
<tr>
<td>Shenzhen</td>
<td>2002</td>
<td>404.2</td>
<td>0.421</td>
<td>7.88</td>
<td>617.5</td>
</tr>
<tr>
<td>Shenyang</td>
<td>2006</td>
<td>492.34</td>
<td>0.597</td>
<td>5.3</td>
<td>698.5</td>
</tr>
<tr>
<td>Shijiazhu</td>
<td>2007</td>
<td>32.54</td>
<td>0.425</td>
<td>4.8</td>
<td>64.163</td>
</tr>
<tr>
<td>Suzhou</td>
<td>2003</td>
<td>127.94</td>
<td>0.560</td>
<td>2.25</td>
<td>576.2</td>
</tr>
<tr>
<td>Taicang</td>
<td>2003</td>
<td>13.69</td>
<td>0.445</td>
<td>6.32</td>
<td>44.89</td>
</tr>
<tr>
<td>Wenzhou</td>
<td>2007</td>
<td>144.16</td>
<td>0.529</td>
<td>6.1</td>
<td>852.2</td>
</tr>
<tr>
<td>Guiyang</td>
<td>2006</td>
<td>83.6</td>
<td>0.459</td>
<td>5.45</td>
<td>371.8</td>
</tr>
<tr>
<td>Wujia</td>
<td>2003</td>
<td>45.8</td>
<td>0.416</td>
<td>4.11</td>
<td>77.3</td>
</tr>
<tr>
<td>Yangzhou</td>
<td>2009</td>
<td>110</td>
<td>0.485</td>
<td>4.89</td>
<td>452</td>
</tr>
<tr>
<td>Yinchuan</td>
<td>2008</td>
<td>75.82</td>
<td>0.407</td>
<td>2.58</td>
<td>137.7</td>
</tr>
<tr>
<td>Changsha</td>
<td>2008</td>
<td>23.5</td>
<td>0.346</td>
<td>4.02</td>
<td>104</td>
</tr>
<tr>
<td>Suzhou</td>
<td>2009</td>
<td>75.84</td>
<td>0.416</td>
<td>4.32</td>
<td>514</td>
</tr>
<tr>
<td>Changde</td>
<td>2010</td>
<td>66</td>
<td>0.446</td>
<td>5.59</td>
<td>607.7</td>
</tr>
<tr>
<td>Huaibei</td>
<td>2005</td>
<td>68.73</td>
<td>0.467</td>
<td>4.34</td>
<td>202</td>
</tr>
<tr>
<td>Kunshan</td>
<td>2006</td>
<td>19.30</td>
<td>0.429</td>
<td>5.1</td>
<td>120</td>
</tr>
</tbody>
</table>
3.2. Regression Model for the Second and the First Level. Set up the relationship between mode split rate $M$ and transport facility configuration $R$, trip distance $D$, travel time $T$, and private car ownership rate $O$.

In order to capture the impact of various factors on mode choice more clearly, we use the partial least-squares regression (PLS) to fit data. PLS consists of the basic functions of multiple linear regression analysis, canonical correlation analysis, and principal component analysis, and by using PLS we can not only obtain more accurate fitting results but also analyze the impact of each factor, that is, the contribution rate of each factor (contribution, VIP, variable importance point).

(1) Walk-Mode Share Rate Fitting $M_\alpha = F_4(R, D, T, O)$. Consider

$$M_\alpha = -0.476R - 7.265D + 0.274T - 13.942O + 55.757 \quad (R^2 = 0.893),$$

where $M_\alpha$ is the walk share rate (%).

Trip distance has the largest impact on the walk share rate, for walk occupying the largest proportion among the mode choice of short trips and decreasing rapidly as the distance increases. Hence, it is necessary to select the walk mode from microcosmic view. The VIP value of per capita ownership has also been over 1, because statistics in this paper are about the ownership of motor vehicle which includes motorcycle and the other two. Unlike walk mode, bicycle, bus and car mode relate to the factor of ownership. As the walk mode split rate can be regarded as total trips minus the rate of this three modes, the walk mode split rate also relates to the factor of ownership.

(2) Bicycle-Mode Share Rate Fitting $M_\beta = F_4(R, D, T, O)$. Consider

$$M_\beta = 3.190R + 14.518D - 1.458T - 3.458O + 19.195 \quad (R^2 = 0.967),$$

where $M_\beta$ is the bicycle share rate (%).

Bicycles, servicing primarily to short-distance travel in metropolis of our country, are greatly influenced by trip distance. In addition, other factors such as travel time cost, the per capita holdings, and facilities configuration whose VIP values are close to one also play a significant role.

(3) Bus-Mode Share Rate Fitting $M_\gamma = F_4(R, D, T, O)$. Consider

$$M_\gamma = -2.694R - 5.682D + 1.126T + 6.247O + 12.540 \quad (R^2 = 0.926),$$

where $M_\gamma$ is the bus share rate (%).

Travel time has the largest effect on bus mode choice, which illuminates the obvious time-consuming characteristics of public transit mode; in other words, whether people choose public transit or not depends mainly on the travel time. So the key to increase bus travel is focusing on the
reducing of bus delay and the improvement of the service level.

(4) Car-Mode Share Rate Fitting $M_ω = F_i(R, D, T, O)$. Consider

$$M_ω = 0.219R - 1.741D + 0.216T + 10.337O + 6.594$$

$$+ 6.594 \left( R^2 = 0.903 \right),$$

(19)

where $M_ω$ is the car share rate (%).

The number of motor vehicles itself determines the personal transport split which is also affected by road conditions, because road traffic facilities would induce private car travel. The trip distance of personal transport has a wide coverage, so distance and time-cost characteristics have little influence on it.

3.3. Model Integration. Concluding the above analysis, we can calculate the share rate of the various modes using formulas (21) and (22) if we know urban form $C$, city scale $S$, trip purpose $I$, and policy factors $P$. Consider

$$D = 7.853 \times 10^{-2} \sqrt{S} + 0.157C - 0.208 \exp(I) + 2.269,$$

$$T = 3.580 \ln(S) - 1.386 \exp(I) + 7.676,$$

$$O = -70.418P^4 + 127.33P^3 - 70.814P^2$$

$$+ 14.661P - 0.1207,$$

$$R = -9.25 \times 10^{-2} \exp \left( \exp \left( \frac{(C - 0.211)}{0.432} \right) \right)$$

$$+ 2.942 \times 10^{-2}P^{-1.974} + 5.987,$$

$$M_α = -0.476R - 7.265D + 0.374T - 13.942O + 55.757,$$


$$M_γ = -2.694R - 5.682D + 1.126T + 6.247O + 12.540,$$

$$M_ω = 0.219R - 1.741D + 0.216T + 10.337O + 6.594.$$  (21)

4. Application to Six Cities in China

As a final extension of the research, a brief forecasting application is conducted to verify the fit goodness of the model developed above and reality system and to test whether the model can reflect the characteristics and changing rule of reality system or not. In addition, the validation process is an important way to analyze the problem-solving ability of our proposed model. Thus in order to calculate the mode split rate and analyze the validity of model, we compare the forecasting results with actual survey data taken from Wuhan, Changchun, Nanning, Hefei, Anqing, and Nantong.

We substitute the original data of city scale $S$, urban form $C$, trip purpose $I$, and traffic policy $P$ into formulations (20) and (21) to obtain the share rate of every mode $M_i$ ($i = α, β, γ, ω$) and then control total amount to 100 using formulation (22) to get the final predictive value $M'_i (%)$. Finally, error rate $Δ_i (%)$ is calculated by comparison of $M'_i$ with actual survey data $N_i (%)$. Consider

$$M'_i = M_i \times \left( \frac{100}{\sum M_i} \right).$$

(22)

Predictive values and validation results are listed in Table 4 which indicates that the model direct calculation results $\sum M_i$ are very close to 100, so this method for mode split is accurate and there is no need to control the total amount. In addition, the forecasting of walk and bicycle shows high precision, while bus and car mode show some errors. The problem in essence may be caused by the forecast method itself and the reasons may be the following: (a) the foundational data is inadequate, that is, the sample size is not large enough to consider cities of all types; (b) the mode split calculation is acquired by several formulations, leading to more probability of producing bias; and (c) there are qualitative variables in our model, which bring difficulties to the forecasting.

<table>
<thead>
<tr>
<th>City</th>
<th>$C$</th>
<th>$S$</th>
<th>$I$</th>
<th>$P$</th>
<th>$M_α$</th>
<th>$M'_α$</th>
<th>$N_α$</th>
<th>$Δ_α$</th>
<th>$M_β$</th>
<th>$M'_β$</th>
<th>$N_β$</th>
<th>$Δ_β$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wuhan</td>
<td>418.22</td>
<td>0.37</td>
<td>0.507</td>
<td>0.5</td>
<td>29.876</td>
<td>29.589</td>
<td>30.60</td>
<td>-3.03</td>
<td>44.641</td>
<td>44.211</td>
<td>43.18</td>
<td>2.389</td>
</tr>
<tr>
<td>Changchun</td>
<td>306.824</td>
<td>0.577</td>
<td>0.354</td>
<td>0.45</td>
<td>47.263</td>
<td>46.886</td>
<td>47.38</td>
<td>-1.04</td>
<td>17.499</td>
<td>17.359</td>
<td>18.53</td>
<td>-6.318</td>
</tr>
<tr>
<td>Hefei</td>
<td>107.49</td>
<td>0.568</td>
<td>0.223</td>
<td>0.4</td>
<td>30.011</td>
<td>29.987</td>
<td>31.34</td>
<td>-4.31</td>
<td>42.560</td>
<td>42.525</td>
<td>42.37</td>
<td>0.366</td>
</tr>
<tr>
<td>Nanning</td>
<td>128.3</td>
<td>0.542</td>
<td>0.456</td>
<td>0.35</td>
<td>26.265</td>
<td>26.410</td>
<td>25.26</td>
<td>4.555</td>
<td>54.957</td>
<td>55.261</td>
<td>57.72</td>
<td>-4.258</td>
</tr>
<tr>
<td>Anqing</td>
<td>60.3</td>
<td>0.507</td>
<td>0.316</td>
<td>0.4</td>
<td>30.800</td>
<td>30.923</td>
<td>32.70</td>
<td>-5.432</td>
<td>43.103</td>
<td>43.275</td>
<td>44.88</td>
<td>-3.574</td>
</tr>
<tr>
<td>Nantong</td>
<td>54</td>
<td>0.436</td>
<td>0.485</td>
<td>0.3</td>
<td>30.860</td>
<td>30.699</td>
<td>32.91</td>
<td>-6.716</td>
<td>52.722</td>
<td>52.447</td>
<td>51.48</td>
<td>1.879</td>
</tr>
</tbody>
</table>

Table 4: Model validation results.
5. Conclusions

From the view of data analysis, we use the factor contribution rate in the model to estimate the effects of motor vehicle population, trip distance, road traffic facilities, and travel time on trip mode split. The analysis results can provide theoretical basis for trip mode structure policy. For example, in order to improve the bus share rate, decision makers should increase the service level of public transit to make sure that the bus travel time is in the reasonable range. In addition to travel time, the travel distance also affects the trip mode split greatly. Thus, using advanced technology to collect information on passenger trip distance is very important. As for private cars, we can reduce travel in private cars from the following two aspects. On the one hand, control the private car ownership by limiting the amount of license. On the other hand, considering the impact of trip distance, increase the attractiveness of public transit to share part of the traffic volume.

It is a new attempt to develop the hierarchical structure model for modal split. The methods and models presented in this paper are not the most advanced mode split modeling tools, but they comprehensively consider nearly all aspects of influencing factors in the transportation system and reflect dynamic development process of urban trip mode structure. Moreover, calculation results which are the factual manifestation of traffic mode structure in this city can be of great help for policymakers to grasp the traffic structure in a macro level and plan the further transportation more profoundly.

However, the accuracy of decision coefficients (R square) in some formulations is not as high as we expected, indicating a certain gap between actual situations, which the deeper work can be traced to solve. A way of building up a more parsimonious predictive tool should be on the policy value of trip mode choice. Using this method, decision-makers can analyze the impacts of transport policies on trip mode split in the micro level and make scientific decisions based on urban feature.

In addition, compared with disaggregate models and neural networks, the proposed methods and models ignore the properties of the individual trips and do not take into account the impacts of traveler's physical, psychological, and other characteristics on trip mode split. Therefore, the impacts of micro individual travel characteristics and macro urban characteristics on trip mode split should be analyzed comprehensively, which will be the most important improvements directions of this paper.

Conflict of Interests

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

Acknowledgments

This research is supported by the National Natural Science Foundation of China (51308298 and 51308311), Project of Ministry of Housing and Urban-Rural Development of China (2013-K5-20), Project of Jiangsu province joint innovation fund of industry, education, and research Prospective joint research project (BY2013004-04), and Project of Nanjing University of Science and Technology.

References


