Extending the Theory of Planned Behavior to Explore the Influence of Residents’ Dependence on Public Transport

SONG HU1, PHAM MINH CHAU1, JIANCHENG WENG1, PENGFEI LIN1, AND WEI ZHOU2

1Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Beijing 100124, China
2Ministry of Transport of People’s Republic of China, Beijing 100736, China
Corresponding author: Jiancheng Weng (youthweng@bjut.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 52072011 and Grant U1811463.

ABSTRACT The accurate depiction and understanding of the influence mechanism of residents’ dependence on public transport (RDPT) is an important foundation for increasing the proportion of green trips and alleviating urban traffic congestion. To explore the influence mechanism of RDPT, this paper extends the theory of planned behavior (TPB) by introducing objective factors of attributes, environment, and travel characteristics. The agglomerative nesting clustering algorithm and multiple structural equation modeling (SEM) are then developed to identify the RDPT levels and analyze the influence relationships between the integrated influencing factors and RDPT based on travel survey data. The results indicate that the objective variables have indirect impacts on RDPT by influencing psychological variable attitudes, subjective norms and perceived control and travel intention. The residents’ self-selection (RSS) effects of different clusters are all detected under normal conditions, and the environment on RDPT is still significant after controlling for this effect. The findings reflect that the influence mechanisms of the three SEMs for different clusters are distinct, unlike those of the baseline model, and distinctive observation variables have dissimilar explanatory abilities for the travel intention of different residents. Therefore, some beneficial policy implications are proposed for residents, especially those who retain low and relatively low public transport dependence levels, to increase public transport usage while reduce the car usage based on the significant findings.

INDEX TERMS Public transport, dependence, theory of planned behavior, agglomerative nesting algorithm, structural equation modeling.

ABBREVIATIONS
PT Public Transport.
RDPT Residents’ Dependence on Public Transport.
TPB Theory of Planned Behavior.
SEM Structural Equation Modeling.
RSS Residents’ Self-selection.
SP Stated Preference.
RP Revealed Preference.
KMO Kaiser-Meyer-Olkin.
AGNES Agglomerative Nesting.
POI Points of Interest.
TD Travel Days.
TN Travel Number.
M.I. Modification Indices.

I. INTRODUCTION

With the rapid development of the social economy, the way urban residents live and travel are undergoing great changes in China. Currently, the most frequently used mode of transportation for all types of travel is private transportation [1], [2], which has led to various urban problems, including traffic congestion and air and noise pollution. As an intensive, wide-coverage and high-capacity transport mode, public transport (PT) is regarded as an alternative to private car travel [3], which can effectively alleviate various urban transport problems and promote the development of sustainable transport systems. Therefore, traffic management departments have focused on the construction and optimization of urban PT systems to restrict car use and guide transit travel behavior in recent years, with the continuous promotion of the National Transit Metropolis and PT priority in China.
However, transit ridership has not increased significantly in China, showing only minor growth from 16.3% in 2004 to 19.2% in 2017 [4]. Furthermore, due to the diversity of urban traffic modes and the spatial imbalance of infrastructure construction, residents’ mode selection intention and travel behavior are complex, based on heterogeneity characteristics and are difficult to predict. Thus, travelers with different individual attributes have different PT dependence levels, which are influenced by internal and external factors [5]. To better understand travel choice decision-making and improve PT service quality, it is essential to explore the causal relationship between the multidimensional influencing factors and residents’ dependence on PT (RDPT).

There have been many studies on behavioral theory applications and travel behavior modeling in different scenarios. In terms of the theoretical analysis of the travel behavior influence mechanism, the theory of planned behavior (TPB) is one of the most widely used theories. Based on the modified technology acceptance model and the TPB, Chen found that perceived pleasure and subjective norms have the strongest power to influence loyalty by both users and nonusers [6]. Considering the TPB paradigm, researchers developed a four-step analysis to investigate the psychosocial determinants of PT usage behavior [7]. Simsekoglu and Klöckner collected data through an online survey and utilized the structural equation model and the TPB to examine the role of normative and environmental beliefs, perceived attributes, and innovativeness and demographic factors related to e-bike usage [8]. To study the expected usage of future metro services when they become available, Shaaban and Maher used the TPB to predict residents’ intentions of using the forthcoming metro [9]. Borhan et al. proposed three new constructs, and found that attitude, subjective norms, and perceived behavior control in the theoretical framework can directly and positively influence the behavior intention of Libyan car users for intercity travel [10]. These studies provide a new perspective to expand the TPB for analyzing residents’ travel dependence behavior. In addition, few researchers also proposed new constructs for the TPB to strengthen theoretical explanations. Conner and Abraham explored the impacts of past behavior, personality traits on intentions and behavior, as well as the relationship between health protection, exercise, predictions of intentions and self-reported behavior [11]. According to [12], four predictors including situational factors, trust, novelty seeking and external influence are added to the TPB to understand travelers’ willingness to use the train in Petaling Jaya. To understand the commuting travel mode choice among office workers further, Lo et al. extended the contents of attitude, descriptive norm, and perceived control in TPB [13]. The results indicate that these three indicators were consistently associated with intentions. Analogously, Jing et al. proposed an extended TPB to delve into the psychological factors caused by the effects of adults’ cognition and behavioral habits, and explored the factors’ relationship paradigm [14].

Research on the determinants of residents’ travel behavior has also resulted in some findings under different situational conditions. On one hand, the normalized condition was taken as the research background to explore the related studies. Sonmez and Graefe investigated the influence of past international travel experience, types of risk associated with travel, and the overall degree of safety, using information integration theory and protection motivation theory [15]. According to [16], an integrated path analysis-discrete choice model was proposed to examine the influence of commuting mode choice. The results indicated that parking availability and the built environment in residences and workplaces both have significant effects on car use for commuting. Rahimi, Azimi, and Jin presented a comprehensive analysis of people’s attitudes, including preference, perception, reasons and motivations toward shared mobility options and autonomous vehicles, focusing on underlying patterns and determinants [17].

On the other hand, previous studies also show that the adverse weather conditions are observably associated with residents’ travel behavior [18], [19]. Zanni and Ryley collected data from more than 2,000 residents and adopted two improved logit models to capture the impact of adverse weather, including heavy snow and volcanic ash, on long-distance travel behavior [20]. Wu, Liao, and Rose used a survey of residents and subway ridership data to examine the interrelation between weather and travel behavior in Beijing [21]. They found that extreme weather events affect recreational travel, reduce travel demand, and change travel modes.

A network relationship diagram shown in Figure 1, based on the above analytical literature, depicts how the relationship of the keywords to the influence of residents’ travel behavior is identified and understood. Through the analysis of keywords’ network relationships, it is clear that PT is usually regarded as an alternative to cars, and residents’ selection of transport mode is influenced by many factors, such as weather, land use, travel distance, and residential self-selection (RSS) effect, as well as the psychological elements of attitude, perception, satisfaction and loyalty.

Most previous studies regarding the residents’ travel dependence focus on the analysis of residents’ car dependence, while few achievements related to RDPT and its potential for changing determinants have been found. Examples include automobile dependence, expressed through comparative levels of car ownership and usage, and transit service and usage, which varies widely and systematically across a large sample of international cities [22]. In efforts to explore residents’ car dependence in a PT-dominated city, a survey of 401 car owners was implemented. The collected data were utilized to analyze why people owned cars and how dependent car owners were on their vehicles [23]. Ness investigated urban structure matters by capturing the relations between residential location, car dependence and travel behavior [24]. According to the literature [25], household travel survey
data were obtained to analyze the connection between residents’ dependence on cars and physical inactivity. The results showed that there is a large variation in physical activity within auto-dependents. Wang et al. used descriptive statistics and sequencing logistic regressions to analyze the car usage behavior of urban residents [26]. The results showed that the purpose of automobile usage, built environment, and various socioeconomic characteristics have important effects on the intensity and dependence of automobile usage.

Overall, although TPB theory is relatively mature in the field of travel behavior, limited efforts have been made to expand TPB theoretical paradigm from a multi-dimensional perspective. Few studies related to travel dependence and influence relationship analysis about travel dependence from the perspective of PT systems have been conducted. In addition, a large number of previous studies on the influence of travel behavior have been carried out, but only from a single perspective of subjective psychology or the objective environment; a comprehensive analysis that combines both factors is lacking. In this context, the RSS effect, which remains a significant contributory factor for certain populations when sustainable modes are considered [27], should be of greater concern because the travel behavior and especially the transit preferences of Chinese residents are dissimilar to those of Europeans and Americans [28].

Compared with previous research, the present study has been conducted to examine objective and psychological factors, including individual attributes, travel environment, travel characteristics, attitude, subjective norms, and perceived behavior control. Although some of these variables have been examined in previous studies separately, they have not been inspected comprehensively. In addition, the TPB is expanded to more comprehensively analyze the causal relationship between the influencing factors and travel behavior performance, and a conceptual framework is also constructed to evaluate these relationships quantitatively. Moreover, this study contributes to a comprehensive understanding of what determines residents’ PT usage behavior and how residents make their travel choices, which supports transportation planning and design, operation management and policy formulation for decision-makers.

The remainder of this paper is organized as follows. The next section proposes a theoretical framework of extended TPB. Section 3 describes the survey design scheme, individual discrimination model, and causal model of RDPT. Section 4 presents the relevant results from the model estimation. Section 5 discusses the implications of the findings, proposes policy implications, provides a conclusion of the main results, and presents future related research ideas.

II. THE THEORETICAL ANALYSIS OF TPB

The TPB is a well-known theoretical approach to the behavior of cognitive determinants and has been successfully used to evaluate the relationship between cognitive and behavioral determinants [29], [30]. Many applications of the TPB in the field of traffic behavior also demonstrate the applicability and validity of the theory in explaining perceived decision behavior [7], [31]. The core idea of TPB is that behavioral intention directly determines actual behavior, and is jointly affected by attitudes, subjective norms and perceived behavior
control [32]. The theoretical framework shows that attitudes, subjective norms and perceived behavior control have direct influence on travel intention, but indirect influence on travel behavior. However, the TPB focuses on behavioral intention analysis from the subjective psychological dimension and lacks an explanation of the objective influence variables on actual behavior.

To enhance the interpretability and theoretical coverage of TPB, many previous studies have added additional constructs to their theoretical frameworks, such as individual attributes, travel characteristics, built environment and situational factors [12], [28], [33]. First, the individual attributes demonstrate the population diversity and basic personal information. Residents can be divided into groups according to their attributes, and different residents consider different self-factors when making travel decisions. Thus, individual attributes are considered influencing factors of travel behavior [34], [35], such as the education level [16], [36]. Second, in terms of travel characteristics, there is an effect of influence and being affected in exploring travel behavior. On the one hand, the travel characteristics of distance, mode and time are treated as travel behaviors that are influenced by external variables [28], [37]. In addition, some research has considered travel distance, time and purpose as factors that influence travel choice behavior [35], [38]. As a result, travel characteristics are introduced as influencing factors of RDPT in this study. Last, considering the diversity of residential communities regarding travel behavior, studies usually focus on the residential built environment. Residential environmental factors, including residential density, mixed land use, distance to transit and street network connectivity, have been proven to have a significant positive influence on PT travel choice behavior [36], [39]. Therefore, the environment is also chosen to enhance the effectiveness of the prediction of RDPT.

In terms of the relationship between these variables, the hypothesis is that the objective variables, including individual attributes, travel characteristics and environment, have an indirect influence on residents’ travel behavior and a direct effect on their psychology [10], [35], while the psychological variables in TPB have a direct effect. Additionally, the RSS indicates that individual attributes or attitude preferences can affect residents’ choice of residence with certain built environment characteristics, and the influence of the built environment on travel behavior could be underestimated if the RSS effect was ignored [28], [40]. Therefore, we take the RSS effect into account and try to control it if the individual attributes have a significant influence on the environment and PT attitude observably influences RDPT. Thus, this paper develops an innovative and extended model framework based on the TPB, shown in Figure 2, that focuses on identifying the influence mechanism of the RDPT.

**III. METHODOLOGY**

Based on travel survey data, this study uses an agglomerative nesting (AGNES) clustering algorithm to identify RDPT, and then multiple structural equation modeling (SEMs) are adopted to explore the influence mechanism of RDPT. The specific research process is shown in Figure 3.

**A. SURVEY DESCRIPTION AND DATA COLLECTION**

In this study, Beijing, which consists of 16 administrative regions, was chosen as the research case. Beijing’s PT system is expanding rapidly as a result of the city’s recent economic development and population growth. By the end of 2020, Beijing had formed a relatively excellent PT infrastructure system. Beijing implemented 24 metro lines and 727 kilometers of operation mileage, and there were more than 1,100 bus lines in the urban area of Beijing and approximately 1,000 kilometers of bus lanes. The 750-meter coverage rate of rail transit stations in the central urban area of Beijing reached 90% in 2020, and the 500-meter coverage rate of bus stations will reach 99.5% in 2022. Beijing was severely disrupted by the COVID-19 epidemic in 2020; the annual total number of bus and subway passenger trips was...
In the context of the expanded TPB paradigm, a stated preference (SP) and revealed preference (RP) survey questionnaire, including individual attributes, travel environment, travel characteristics, PT attitude, subjective norms, perceived behavior control and travel intention, was designed and implemented online in July 2020. The IP address restriction was carried out to collect survey information only of Beijing residents. Through this survey, information related to the trips and the psychological determinants of travel, such as origin, destination, purpose, travel mode, length of the trip, and mode preference, was collected. Data on socioeconomic characteristics and travel environments, such as age, gender, level of education, income, vehicle ownership, mixed land use and distance to transit, were also collected during the survey. The items related to TPB factors were measured using a five-point Likert scale, ranging from 1 = very low to 5 = very high. Higher scores show a higher level of interest in a particular measure. While other items of objective factors are measured according to the actual situation of the respondents.

A total of 408 valid household samples were collected. Several sample control approaches, including prior knowledge information verification and sample proportion control, were adopted to eliminate invalid questionnaires. A final dataset of 307 questionnaires was used in the study, representing an overall response rate of 75.2%. Table 1 shows the statistics of the sample data, and a number of respondents are younger people with higher education. This may be because Beijing is an attractive metropolis in China and attracts many young and highly educated people to work in Beijing every year. By 2020, the number of people from other cities accounts for 38.5% of the permanent resident population in Beijing.

The data recruited from this web-based survey were input into the Social Sciences Software (SPSS) version 23 to test the reliability and validity of the questionnaire data. Statistical analysis was conducted by calculating the Cronbach's alpha value to determine the internal consistency coefficient of the data measuring tool [41], and the Kaiser-Meyer-Olkin (KMO) and Bartlett sphericity tests were used for validity testing [34]. The Cronbach’s alpha values and KMO values of the latent variables for this survey are all above 0.764 and 0.791, respectively, which exceeds the cutoff value of 0.7 recommended by [42]. Hence, all constructs of this questionnaire are deemed to be reliable and valid.

### B. THE PT DEPENDENCE IDENTIFICATION MODEL

To accurately explore the influence of RDPT, it is essential to first divide the residents into different groups according to their PT dependence levels. The identification results provide an input variable which belongs to the dependent variable of the causal analysis model for the whole sample, and the subgroup sample dataset for analyzing the PT dependence influencing mechanism of residents in each cluster in the following parts. The AGNES clustering algorithm, a hierarchical clustering algorithm, is a widely used and well-researched unsupervised clustering method [43]. It has been successfully applied in many research fields, such as network user behavior [44], [45], communication signals [46], and medicine [47]. A hierarchical clustering algorithm combines the two most similar data points by calculating the similarity of the two points in the data set and creates a hierarchical nested clustering tree. The dataset is then divided into different clusters of layers. The calculation flow of the algorithm is shown in Figure 4. The algorithm can find the hierarchical relationship between clusters and determine the optimal number of clusters.

There are three methods used to calculate the distance between two composite data points by the hierarchical clustering algorithm: single linkage, complete linkage and average linkage. Among them, the calculation method of average linkage is used to calculate the distance between each data point in the two datasets and all other data points, and the mean value of all distances is taken as the distance between the data points in two datasets. Although the calculated load of this method is large, the result is more reasonable than that of the other two methods. Considering the small scale of the sample size in this paper, the average linkage method

| Item                              | Value | Item                              | Value |
|-----------------------------------|-------|-----------------------------------|-------|
| Gender                            |       | Monthly income                    |       |
| Male                              | 50.2  | Less than 1500¥                   | 12.5  |
| Female                            | 49.8  | 1500-3000¥                        | 4.9   |
| Age                               |       | 3001-5000¥                        | 13.7  |
| Less than 20 years old            | 2.7   | 5001-8000¥                        | 19.4  |
| 21–30 years old                  | 41.9  | 8001-15000¥                       | 31.6  |
| 31–40 years old                  | 32.6  | 15001-20000¥                      | 11.4  |
| 41–50 years old                  | 15.6  | More than 20000¥                  | 6.5   |
| 51 years old and above            | 7.2   | Occupation                        |       |
| Educational levels                |       |                                   |       |
| Senior high school and below      | 8.1   | Student                           | 19.0  |
| Junior college                    | 12.7  | Unemployed and retired            | 1.5   |
| Undergraduate                     | 36.4  | Civil servant                      | 29.3  |
| Graduate and above                | 42.8  | Staff                             | 38.8  |
| Taking the kids to school         |       | Temporary                         | 3.0   |
| Yes                               | 11.4  | Self-employed                     | 3.1   |
| No                                | 88.6  | Freelance                         | 1.9   |
|                                  |       | Other                             | 3.4   |
is expressed as Eq. (1) and used to measure data similarity.

\[ d(u, v) = \sum_{ij} \frac{d(u_i, v_j)}{|u| * |v|} \]  

where \( u \) and \( v \) are different data point sets and \( i, j = 1, 2, 3 \ldots, |u| \) and \(|v|\) indicate the number of data points in the two datasets.

The contour coefficient is a classical cluster performance evaluation index based on sample distance, which can effectively measure the degree of dissimilarity within and between clusters. Therefore, the contour coefficient is introduced to measure the quality of model classification, and it can be defined as follows:

\[ S(i) = \frac{b(i) - a(i)}{\max \{a(i), b(i)\}} \]  

where \( a(i) \) represents the average distance between sample \( i \) and other samples in the cluster, \( b(i) \) indicates the average distance between sample \( i \) and other cluster samples.

C. PT DEPENDENCE INFLUENCING VARIABLE SELECTION AND MODELING

1) MODEL VARIABLE SELECTION

Effective selection of model variables is the basis of successful model construction. The extended TPB theory in section 2 is adopted to extract accurately the model variables to analyze the influence mechanism of the RDPT. Specifically, 3 objective factors and 4 subjective factors containing a total of 23 measurement variables were selected, and the statistical results of the variables are shown in Table 2. The intensity of mixed land (S1), which assesses the diversity of land use, is calculated by classifying the points of interest (POIs) in Beijing into six categories: residence, employment, commerce, scenic spot, transportation, and education. The indicator S1 is explained in expression (3).

\[ E_j = \sum_i \frac{N_{ij}}{A_j} \]  

where \( N_{ij} \) represents the number of POI in category \( i \) in region \( j \), and \( A_j \) indicates the area of the region \( j \).

2) STRUCTURAL EQUATION MODELING FOR PT DEPENDENCE INFLUENCE

The discrete choice models can effectively describe the direct relationship between the PT dependence and its influencing variables, but lack the ability to express the structural relationships between the influencing factors, as well as the influencing factors and the PT dependence. SEM is a multivariate statistical method to analyze the relationship between variables based on the covariance matrix of variables [48]. SEM can be used to analyze complex multivariate data, and explores and tests the causal relationship between the influencing variables and actual behavior. Additionally, SEM can explore the relationship between multiple dependent variables and break through the “black-box” expression paradigm of behavior influence compared with the traditional multivariate statistical method, so the approach is more suitable for the influence analysis of multiple factors synchronously [49]. Therefore, SEM is adopted in this paper to investigate the direct and indirect effects of internal and external factors on RDPT.

To specify both the travel choice behavior and the latent variables, SEM contains two kinds of equations: a) the measurement equation that links the latent variables to the indicator variables and b) the structural equation that relates the latent variables to the explanatory variables [35]. In terms of the measurement equation, the combinatorial calculation model can be written as:

\[ x_{12+1} = \Lambda_x \xi + \delta_{12+1} \]  

\[ y_{12+1} = \Lambda_y \eta + \epsilon_{12+1} \]

where \( x \) is the vector matrix formed by 12 observed values of the exogenous latent variables, \( \xi \) means the vector matrix composed of 3 exogenous latent variables, \( \Lambda_x \) indicates the factor load matrix of \( x \) to \( \xi \), \( \delta \) represents the measurement error vector matrix of 12 exogenous variables, \( y \) is a vector matrix composed of 12 observed values of endogenous latent variables, \( \eta \) indicates the vector matrix composed of 5 endogenous latent variables, \( \Lambda_y \) refers to the factor load matrix of \( y \) to \( \eta \), and \( \epsilon \) demonstrates the measurement error vector matrix of 12 endogenous variables.

The expression of the structural equation is as follows:

\[ \eta_{5+1} = B_{5+5} \eta_{5+1} + \Gamma_{5+3} \xi_{3+1} + \xi_{5+1} \]

where \( B \) represents the structural coefficient matrix of the endogenous latent variables \( \eta \), \( \Gamma \) indicates the structural coefficient matrix of the exogenous latent variables \( \xi \), and \( \xi \) is the error vector of the endogenous latent variables.

IV. RESULTS

A. TRAVELER GROUP CLASSIFICATION

Taking residents who live in Beijing with an advanced PT system as the main research objective in this paper, the three indicators of the ratio of PT travel days (TD), the ratio of PT travel number (TN) and the ratio of PT roundtrip were
TABLE 2. Descriptive statistics of the sample.

| Factors          | Construct               | Indicators                                      | Mean   | Standard deviation |
|------------------|-------------------------|-------------------------------------------------|--------|--------------------|
| Attributes       | A1                      | Car availability                                | 3.30   | 1.10               |
|                  | A2                      | Bicycle availability                           | 3.65   | 1.06               |
|                  | A3                      | Taking children to school                       | 0.11   | 0.32               |
|                  | A4                      | Age                                             | 2.66   | 0.92               |
|                  | A5                      | Occupation                                      | 3.26   | 0.86               |
|                  | A6                      | Education level                                 | 2.71   | 1.55               |
|                  | A7                      | Income                                          | 4.13   | 1.68               |
|                  | C1                      | Trip purpose                                    | 1.33   | 0.81               |
| Characteristics  | C2                      | Trip distance                                   | 18.91  | 11.57              |
|                  | S1                      | Intensity of mixed land use                      | 0.43   | 0.48               |
| Environment      | S2                      | Distance to transit                              | 25.51  | 12.06              |
|                  | S3                      | Housing price                                    | 6.64   | 2.29               |
|                  | AT1                     | Safety                                          | 3.74   | 1.05               |
| PT attitude      | AT2                     | Convenience                                     | 3.60   | 1.04               |
|                  | AT3                     | Overall satisfaction                             | 3.66   | 0.86               |
| Subjective norms | N1                      | Support degree of PT use from relatives and friends | 3.90 | 0.95 |
| Perceived control| P1                      | Familiarity degree of PT networks                | 3.23   | 1.03               |
|                  | P2                      | Degree of convenience and freedom travel by PT  | 3.70   | 0.98               |
|                  | I1                      | PT travel preference                             | 3.69   | 0.95               |
|                  | I2                      | Car travel preference                            | 3.37   | 1.03               |
|                  | I3                      | Cycling preference                               | 3.44   | 1.10               |
|                  | I4                      | Walking preference                               | 3.30   | 1.10               |

considered representative indicators regarding RDPT [5]. Figure 5 shows the results of the statistical analysis of these three measurement indicators. The results indicate that there are significant heterogeneity characteristics between travel intensity and the travel mode selection of the respondents. This also reflects how different urban residents have distinct PT dependence, which is the core issue we attempt to analyze and investigate in terms of its influence mechanism.

The identification of the RDPT levels and travel group classification was conducted using the AGNES clustering algorithm. The values of the ratio of PT to TD, the ratio of PT to TN and the ratio of PT to roundtrips were input into the initial model. The model was then conducted multiple times, and the model’s parameters were adjusted accordingly. The model’s parameters were then determined, namely, the distance threshold $t = 1$ and the decision tree depth threshold depth $= 2$. Figure 6 shows the hierarchical clustering results. When different cluster numbers were set, the corresponding contour coefficients can be calculated. The results of contour coefficients are shown in Figure 7. The results show that the contour coefficient can achieve the largest value when the cutting height is set at 60, and the optimal cluster number $k$ is 4. The classified clustering samples have significant heterogeneity, which is suitable for analyzing the influencing mechanism of RDPT. Thus, the respondents are divided into four clusters, which are ranked in descending order according to the RDPT. The proportions of respondents in different clusters were 14.1%, 30.4%, 23.2% and 32.3%, respectively. The clustering results provide the dependent variable values of PT dependence influencing model for the whole sample.
TABLE 3. The fit indices value of the baseline model.

| Fit indices | IFI | TLI | CFI | PNFI | PCFI | RMSEA | CMIN/DF |
|-------------|-----|-----|-----|------|------|--------|---------|
| Theoretical range | >0.9 | >0.9 | >0.9 | >0.5 | >0.5 | <0.08 | <3 |
| Value | 0.98 | 0.97 | 0.98 | 0.65 | 0.75 | 0.02 | 1.14 |

1) SEM FOR THE WHOLE SAMPLE
The values of the 23 observed variables presented in Table 2 were input into the SEM to estimate the influence effects of the RDPT. An assessment of the structural coefficient of the overall model was conducted after verifying the measurement model to provide a basis for testing the proposed hypotheses and the extended theory paradigm. To modify the model, various approaches, including adding a covariation relationship between the influence path and the error variable, deleting the path with insignificant influence and adjusting the path weight coefficient, are employed until the modification indices (M.I.) do not prompt modification of the model. The fit indices of the baseline model shown in Table 3 meet the cutoff criterion, which means the model has a good fitting effect on the data including the classification results achieved by the AGNES clustering algorithm. Thus, the rationality and reliability of our research design in Figure 3 can be verified laterally. The direct and indirect influence effects of factors on the PT dependence of the whole sample are presented in Figure 8, and the results of the unstandardized and standardized model estimates are shown in Table 4. The modified model is regarded as a baseline model relative to the following 3 models for the different clusters of respondents.

B. THE INFLUENCE ANALYSIS OF RDPT
To further analyze the PT-dependence influence mechanism of different PT passenger groups, this section will explore the influence relationship between the influencing factors and RDPT for the entire sample and residents with different PT dependence levels.

FIGURE 5. The statistical analysis results of the measurement indicators.

FIGURE 6. The hierarchical clustering results.

FIGURE 7. The results of contour coefficients.

and the subgroup sample dataset of the model for different PT dependence levels.
to live in communities with different environments and have some particular travel behavior characteristics. Therefore, the RSS effect is detected in these results. The environment still has a significant influence on RDPT after controlling the RSS effect. Second, psychological variables, as the mediating variables of the objective variables, have a direct influence on RDPT, and there are significant influencing relationships between the influencing variables. Thus, the validity of the extended TPB theoretical framework and the rationality of hypothesized relationships in section 2 is verified.

The structural equation shown in Figure 8 and Table 4 assesses the indirect effects of the objective variables on the RDPT. Among them, the indirect total effect of attributes on RDPT is \(-0.068 (-0.279 \times 0.815 \times 0.225 + -0.420 \times 0.095 \times 0.225 + -0.181 \times 0.195 \times 0.225)\) via ‘PT attitude + travel intention’, ‘subjective norm + travel intention’ and ‘perceived control + travel intention’. Likewise, the indirect total effects of environment and characteristics on the dependence on PT are \(0.168 (0.617 \times 0.815 \times 0.225 + 0.855 \times 0.095 \times 0.225 + 0.844 \times 0.195 \times 0.225)\) and \(0.004 (0.207 \times 0.095 \times 0.225)\), respectively. The results indicate that only the attributes are negatively associated with RDPT, while the environment and travel characteristics separately have the strongest and weakest indirect effects on RDPT, respectively. Additionally, the direct effects of PT attitudes, subjective norms and perceived control on PT dependence are estimated; the total effects are 0.183, 0.021 and 0.044, respectively. The results suggest that PT attitude significantly affects RDPT, which confirms the expectation that intentions are consistent with actual behavior.

The measurement equation reveals the relationship between the observed variables and the potential variables, and several significant results are found. Bike availability (A2) and occupation (A5) are negatively associated with attributes but are positively associated with RDPT. The other five attributes have the opposite effect, and income (A7) has the highest explanatory ability. Housing price S3 has a relatively high explanatory ability of the environment, which reflects that housing price S3 contributes to the job-residence separation of residents in Beijing [50], exerting a strong influence on PT choice behavior. In terms of the psychological variables, the observed variables all have a positive and high degree of explanatory ability above 0.52 and possess a positive promotion effect on RDPT. According to the explanatory coefficients of the observed variables, residents focus more on convenience (AT2) and overall satisfaction (AT3) when making travel decisions. The degree of convenience and freedom in traveling by PT (P2) has a higher effect on residents’ perceived control (0.87). Then, cycling preference (I3) and walking preference (I4) have a positive impact on PT dependence (0.2 and 0.36, respectively). This is because 78%
### TABLE 4. Unstandardized and standardized model estimates.

| Path                                      | Estimate | S.E. | C.R. | P     |
|-------------------------------------------|----------|------|------|-------|
| **Structural model**                      |          |      |      |       |
| Attributes → PT attitude                  | -0.404   | 0.251| -1.611| 0.107 |
| Attributes → Subjective norm              | -0.529   | 0.278| -1.903| 0.057 |
| Attributes → Perceived control           | -0.171   | 0.189| -0.906| 0.365 |
| Environment → PT attitude                | 0.307    | 0.164| 1.871 | 0.061 |
| Environment → Subjective norm            | 0.369    | 0.192| 1.927 | 0.054 |
| Environment → Perceived control          | 0.274    | 0.154| 1.780 | 0.075 |
| Characteristics → Subjective norm        | 0.091    | 0.150| 0.609 | 0.543 |
| PT attitude → Travel intentions          | 0.580    | 0.081| 7.122 | 0.000 |
| Subjective norm → Travel intentions      | 0.077    | 0.116| 0.668 | 0.504 |
| Perceived control → Travel intentions    | 0.213    | 0.132| 1.607 | 0.108 |
| Travel intentions → Dependence on PT     | 0.388    | 0.111| 3.502 | 0.000 |
| **Measurement model**                     |          |      |      |       |
| Attributes → A1                          | 0.350    | 0.157| 2.233 | 0.026 |
| Attributes → A2                          | -0.159   | 0.137| -1.163| 0.245 |
| Attributes → A3                          | 0.189    | 0.056| 3.360 | 0.000 |
| Attributes → A4                          | 0.781    | 0.202| 3.864 | 0.000 |
| Attributes → A5                          | -0.513   | 0.151| -3.399| 0.000 |
| Attributes → A6                          | 1.000    | —     |      | —     |
| Attributes → A7                          | 1.662    | 0.438| 3.798 | 0.000 |
| Environment → S1                         | 0.036    | 0.027| 1.343 | 0.179 |
| Environment → S2                         | 1.000    | —     |      | —     |
| Environment → S3                         | 0.264    | 0.167| 1.585 | 0.113 |
| Characteristics → C1                     | 0.144    | 0.030| 0.473 | 0.636 |
| Characteristics → C2                     | 1.000    | 0.142| —     | —     |
| PT attitude → AT1                        | 0.644    | 0.083| 7.740 | 0.000 |
| PT attitude → AT2                        | 1.000    | 0.819| —     | —     |
| PT attitude → AT3                        | 0.869    | 0.078| 11.155| 0.000 |
| Subjective norm → N1                    | 0.905    | 0.104| 8.681 | 0.000 |
| Subjective norm → N2                    | 1.000    | 0.743| —     | —     |
| Perceived control → P1                   | 1.000    | 0.536| —     | —     |
| Perceived control → P2                   | 1.545    | 0.291| 5.305 | 0.000 |
| Travel intentions → I1                   | 1.000    | 0.635| —     | —     |
| Travel intentions → I2                   | 0.019    | 0.113| 0.167 | 0.868 |
| Travel intentions → I3                   | 0.375    | 0.121| 3.103 | 0.002 |
| Travel intentions → I4                   | 0.643    | 0.124| 5.181 | 0.000 |
of the respondents who prefer riding and walking also prefer PT. Noticeably, the effect of car travel preference (I2) on RDPT was not found in Beijing, which differs from previous findings [51]. This is because a number of the respondents who maintain preferences for cars also prefer other travel modes or do not have the right to buy a car, which neutralizes the influence effect of the car travel preference.

2) SEM FOR RESIDENTS WITH A HIGH PT DEPENDENCE LEVEL (PTDL)

Model 1 (Figure 9), which is adjusted according to the M.I. and goodness of fit indices in Table 5, focuses on the inter-relationships between the influencing variables and travel intention instead of RDPT. This is because travel intention has a significantly positive effect on RDPT, which is constant. After evaluating the influence effects of the objective and psychological variables on RDPT, several significant results are found. According to the results of the structural equation, the whole relationship structure of Model 1 is similar to the baseline mode. Additionally, the attributes and environment have a negative interactive relationship, which is opposite to that in the baseline model. This may be because the majority of residents with high-attribute characteristics (such as education level and income) tend to choose to live in the suburbs belonging to a low built environment due to the high price of housing in Beijing.

Noticeably, the influence direction of attributes toward residents’ travel intention is also different from that in the baseline mode but is consistent with previous findings [16]. This result may be because nearly two-thirds of the respondents with high-attribute characteristics have low accessibility to cars, which increases the probability of traveling by PT. In addition, the influence effect of the PT attitude of the residents with high PTDL decreases saliently due to the low travel choice. Thus, the influence effect of perceived control on travel intention seems more important.

The measurement equation indicates that bicycle availability (A2), taking children to school (A3) and income (A7) have...
negative explanatory effects, and the observed variables of attributes, apart from bicycle availability (A2), have a lower explanatory ability than those in the baseline mode. However, the explanatory ability of intensity with regard to mixed land use (S1) and housing price (S3) in Model 1 is higher. Likewise, the trip purpose (C1) also has a higher explanatory ability (0.92) in Model 1. This may be because the majority of the respondents with high PTDL are commuters who have frequent travel demands [52]. Moreover, the explanatory ability of the observed variables of the psychological factors is similar to that in the baseline model. As expected, the preference for car travel shows an obvious negative effect on travel intention, which is in accord with previous findings [53].

3) SEM FOR THE RESIDENTS WITH RELATIVELY HIGH PTDL
Model 2 (Figure 10) has a similar influence structure between latent variables and the baseline model. The results of the fit indices in Table 6 demonstrate that the fit of Model 2 is acceptable. Considering the effects of the objective and psychological variables on RDPT, several significant relations are found in Model 2. According to the structural equation, the relationship between attributes and environment is different from the negative effect in Model 1. This may be determined by the different social attribute characteristics between the residents in clusters 1 and 2, as only one-third of the latter have a higher level of education and income. In addition, the influence direction of attributes on the subjective norms and perceived control is different from that in the baseline model. The total influence effect of attributes on RDPT is consistent. Moreover, a weaker positive association is found between the environment and psychological variables. Notably, only the PT attitude of the psychological factors has a significant positive effect on travel intention. This reflects that the influence of PT attitude generated by individuals on the PT dependence of residents with relatively high PTDL occupies a dominant position.

Regarding the results of the measurement equation, the observed variables of attributes retain a positive explanatory level where only occupation (A5) has no significant effect on the RDPT, which is different from the result

TABLE 6. The fit indices value of model 2.

| Fit indices | IFI | TLI | CFI | PNFI | PCFI | RMSEA | CMIN/DF |
|-------------|-----|-----|-----|------|------|--------|---------|
| Theoretical range | >0.9 | >0.9 | >0.9 | >0.5 | >0.5 | <0.08 | <3 |
| Value | 0.912 | 0.901 | 0.902 | 0.509 | 0.742 | 0.056 | 1.186 |

FIGURE 10. Model results for the residents in cluster 2, model 2.
TABLE 7. The fit indices value of model 3.

| Fit indices | IFI  | TLI  | CFI  | PNFI | PCFI | RMSEA | CMIN/DF |
|-------------|------|------|------|------|------|--------|---------|
| Theoretical range | >0.9 | >0.9 | >0.9 | >0.5 | >0.5 | <0.08  | <3      |
| Value       | 0.932 | 0.901 | 0.924 | 0.538 | 0.712 | 0.042  | 1.204   |

FIGURE 11. Model results for the residents in clusters 3 and 4, model 3.

in the baseline model but is consistent with the previous finding [53]. The corresponding explanatory ability of the observed variables of environment and characteristics is similar to that in the baseline model, while the travel environment shows a lower influence effect on the psychological factors. In addition, the degree of support for PT use (N1), the degree of travel convenience and the freedom of PT (P2) show a lower explanatory ability of the latent variables than those under the whole sample conditions, while the degree of familiarity with PT networks (P1) has a slightly higher explanatory ability. In addition, the explanatory ability of the observed variables of travel intention is consistent with that in the baseline model, and PT travel preference (I1) has a better explanatory effect. In particular, PT travel preference largely determines the PT dependence of the residents in cluster 2; therefore, improving PT service quality is an effective measure to increase residents' willingness to travel by PT.

4) SEM FOR RESIDENTS WITH LOW AND RELATIVELY LOW PTDL

Although residents with low and relatively low PTDL make few trips by PT in their daily lives, they commonly do not exclude the use of PT. Therefore, these residents are important groups for PT operators and managers to further improve the PT sharing rate. The influence mechanism of PT dependence of these two clusters is next investigated jointly.

The goodness of fit shown in Table 7 meets the cutoff criterion; thus, the results of Model 3 are also valid. Model 3 (Figure 11) indicates that the relationship between endogenous and exogenous latent variables is consistent with the above models. As expected, the RSS effect is also found among the residents with low and relatively low PTDL. The relation between attributes and the environment coincides with that of model 1. This may be because the residents in clusters 3 and 4 have similar socioeconomic attributes to residents in cluster 1.
Regarding the results of the structural equation, several compelling results are found. The attributes of the residents are negatively associated with the psychological variables while having a lower influence effect than that in the baseline model. In addition, the environment also shows a weaker positive effect on the psychological variables, while travel characteristics have a more significant positive effect on the subjective norm (0.59). In addition, PT attitude and perceived control show a similar influence effect on travel intentions, indicating that the PT dependence of residents with low and relatively low PTDL is mainly affected by their PT attitudes. Surprisingly, the subjective norm has a negative effect on travel intentions, and the degree of support for PT use from relatives and friends (N1) has a higher explanatory ability toward the subjective norm than the degree of influence (N2). According to the travel survey, although the relatives and friends of 78.7% of the respondents in cluster 3 demonstrate a high level of support toward PT, almost two-thirds of them have a relatively short trip distance, within 15 kilometers, which is suitable for cycling [54], [55]. Namely, although their relatives and friends are supportive of PT, most of them adopt active modes such as cycling for their relatively short trips.

Regarding the results of the measurement equation, it is noticeable that only the variables of taking children to school (A3), age (A4) and education level (A6) have a positive relationship with attributes, which shows heterogeneity with the results in other models. This result may be related to the attitude preference and the different socioeconomic attributes of the respondents in clusters 3 and 4. In addition, the trip distance (C2) has a higher explanatory ability than the trip purpose (C1), and it shows a total negative effect on the RDPT. This is mainly related to the inconvenience of taking PT for 46.2% of respondents in clusters 3 and 4 and the relatively short daily trip distance. In addition, the degree of support for PT use from relatives and friends (N1) and the degree of familiarity with PT networks (P1) show a more significant explanatory ability. This may be because of the complex PT network around respondents’ residences and their young attributes. Moreover, car travel preference (I2) is positively correlated with travel intention. This may be explained by the fact that only 13.6% of the respondents who have a positive reference for car travel in clusters 3 and 4 have a high availability of a vehicle, so they have to choose the PT.

**V. DISCUSSION, POLICY IMPLICATIONS**

In terms of the overall structural influence relations of the residents with different PT dependence levels, the findings demonstrate that residents’ travel dependence behavior is directly affected by psychological factors. These factors are significantly affected by external and objective conditions which have an indirect mediating effect on RDPT. Thus, the innovative and extended model framework proposed in this paper has been verified. In addition, the RSS effect was found in different clusters in our research context, which is consistent with previous findings [28]. The heterogeneity of RSS effects for different clusters is mainly related to diverse socioeconomic attributes and PT attitudes.

Specifically, the model results indicate that the multivariate determinants of RDPT show similar or different influence effects in clusters. Unlike other models, the variable of attributes in Model 1 has a positive direct effect on psychological variables and a total positive effect on travel intention. This conclusion is consistent with previous findings [7]. The environment has a significant effect on travel intention in the baseline model, while the effect in Models 1, 2 and 3 is relatively low. The reason may be that the influence effect of the environment is weakened through group classification and controlling the RSS effect. In addition, the travel characteristics of the residents in clusters 3 and 4 have a more significant positive effect on the subjective norm and result in a total negative effect on RDPT. This finding is mainly related to the inconvenience of taking PT and the relatively short trip distance. Moreover, it is an interesting finding that attitudes toward PT have positive direct effects on travel intention in different group models, and own the highest effects on RDPT in Models 1 and 2 and the baseline model at the psychological level, which is consistent with many previous findings [8], [28]. However, perceived PT control has the most influence on travel intention only in Model 1. This finding reflects the fact that PT is the optimal travel choice for residents with high PTDL, which results in a weaker effect of PT attitudes and subjective norms on travel intention. It is worth noting that PT travel preference has the highest variable explanatory ability of travel intention, which indicates that most residents have consistent recognition of green and sustainable PT travel mode in the current social environment.

The results showed that the influence mechanism of PT dependence was significantly different among different clusters. Thus, some efficient measures and policies are proposed to guide or influence travel choices for residents in different clusters according to the influence effects of the observed variables of the RDPT. Once a strong habit is developed through policy incentives, its effect is expected to result in more frequent usage of PT in the future [30], which is beneficial for RDPT. The findings show that car availability retains a general negative influence on PT dependence. Policies restraining the use of private cars, such as reducing parking bays in congested areas, imposing high toll and parking fees [52], [56], promoting the ‘park and ride’ (P+R) model, implementing the odd-even car using the scheme and car traveling restrictions [57], should be implemented. However, it is worth noting that low car availability may incentivize residents with low and relatively low PTDL to use cars to travel in Beijing. Thus, the implemented intensity of private car restriction policies that cover approximately 46.4% of residents should be analyzed further. In addition, expanding the bike-sharing network [58] and providing diversified payment methods for bike rentals are conducive to boosting bicycle availability to improve the RDPT for these residents, which accounts for 44.5% of the entire sample. Furthermore, the
changeable variables of the intensity of mixed land use and distance to transit have a significant positive effect on RDPT for different residents. Improving the diversity of land use to build a multigroup city [59] and optimizing the layout of PT stations could result in 38.2% of the total sample improving their PTDL.

Considering the significant variables of PT attitudes for the residents in clusters 2, 3 and 4, other policies could also improve residents’ subjective cognition and willingness to use PT. Convenience and overall satisfaction are of greater concern to residents when they travel by PT. In terms of convenience, the government should develop some applicable policies and measures, such as providing real-time information on PT services [60], increasing the frequency of high-frequency bus lines, adding bus stops at places with high travel demands and encouraging multimode integration (P+R and MASS). Additionally, these policies are conducive to improving residents’ overall satisfaction with PT. In addition, adjusting PT service schedules (frequencies, operating times, number of routes) and providing congestion levels inside PT vehicles [61], improving the carriage environments and the comfort of the seats [62], and opening exclusive bus lanes in areas or periods with heavy traffic [63] are also key factors further promoting residents’ positive attitudes and affective satisfaction. Wherein, the provision of ramps on buses, popularizing low floor buses, and increasing the number of reserved seats for senior citizen could be adopted to improve PT carriage environments [64]. Moreover, the significant impacts of subjective norms and perceived control on the perception of travel intention and RDPT indicate the importance of creating an overall pro-public transit social atmosphere and providing a highly user-friendly PT service system. In such situations, it is expected that residents would be more satisfied with PT and be more willing to voluntarily travel by PT. Meanwhile, employing the above measures to increase PT ridership can also reduce the private car share on the city roads [57].

VI. CONCLUSION

This study aims to investigate the influence mechanism of RDPT to better understand how the influencing variables affect residents’ PT selection behavior, based on which policies and strategies could be developed in an attempt to manage their behavior. By taking a more comprehensive perspective than previous studies, this study has four objectives: (1) to analyze residents’ transport dependence behavior from the perspective of PT instead of private cars, (2) to extend the TPB framework and include both objective and psychological variable measures of the influence of RDPT, (3) to consider the RSS effect, and (4) to develop multigroup SEMs for different clusters of respondents to specifically explore the influencing mechanism of RDPT.

A better understanding of the interaction relations between objective and psychological variables and the RDPT can be used to obtain optimal operation strategies and thus achieve better PT service quality. The research results indicate that the proposed theoretical concept, hypotheses of variable relationships, and models are effective in measuring the influence mechanism of RDPT. The findings show that the indirect effect of objective variables and the direct effect of objective variables on RDPT and the RSS effect were detected in different travel clusters. Moreover, some management and operational policies are proposed from the perspective of objective conditions and social psychology based on different influence effects on RDPT. It is hoped that the implications would work effectively in the current background at least on a certain group of individuals.

Admittedly, this study has three limitations that call for further investigation to develop a greater understanding of overall travel mode choice behavior. First, the study focuses on the impact of RDPT in a normal environment but lacks a comparative analysis of the influence effects under special conditions, especially public health events. Future studies should attempt to analyze the influencing mechanism of RDPT under different research conditions to better understand and guide residents’ PT usage behavior. Second, future research on trip-based behavior should consider the relationships between companions and respondents in each trip to detect how the heterogeneities between the companions chosen for different travel purposes influence the respondents. Third, we carried out this study only in the case of Beijing, then explored horizontally the impact of different attributes of residents within the same city level on their PT dependence behavior. The vertical comparison of the influencing factors of RDPT in different implementation cities and countries will be further explored in the future work.

REFERENCES

[1] A. N. H. Ibrahim, M. N. Borhan, N. A. Zakaria, and S. K. Zainal, “Effectiveness of commuter rail service toward passenger’s satisfaction: A case study from Kuala Lumpur, Malaysia,” Int. J. Eng. Technol., vol. 8, nos. 1–2, pp. 50–55, 2019.
[2] L. L. Losada-Rojas, C. Gkartzonikas, V. D. Pyrialakou, and K. Gkritza, “Exploring intercity passengers’ attitudes and loyalty to intercity passenger rail: Evidence from an on-board survey,” Transp. Policy, vol. 73, pp. 71–83, Jan. 2019.
[3] C. Su, J. Lu, and P. Xu, “Analysis of urban transport carbon emissions and low-carbon development mode—A case study of Shanghai,” J. Highway Transp. Res. Develop., vol. 29, no. 3, pp. 142–148, 2012.
[4] L. Yu, B. L. Xie, and E. H. W. Chan, “How does the built environment influence public transit choice in urban villages in China,” Sustainability, vol. 11, no. 1, pp. 1–15, 2019.
[5] S. Hu, J. Weng, W. Zhou, P. Lin, and Z. Liu, “A data-fusion based method for identifying individual passenger travel dependence on public transportation,” in Proc. IEEE Intell. Transp. Syst. Conf. (ITSC), Oct. 2019, pp. 2669–2674.
[6] S.-Y. Chen, “Using the sustainable modified TAM and TPB to analyze the effects of perceived green value on loyalty to a public bike system,” Transp. Res. A, Policy Pract., vol. 88, pp. 58–72, Jun. 2016.
[7] X. Fu and Z. Juan, “Exploring the psychosocial factors associated with public transportation usage and examining the ‘gendered’ difference,” Transp. Res. A, Policy Pract., vol. 103, pp. 70–82, Sep. 2017.
[8] O. Simsekoglu and C. A. Klöckner, “The role of psychological and sociodemographic factors for electric bike use in Norway,” Int. J. Sustain. Transp., vol. 13, no. 5, pp. 315–323, 2018.
[9] K. Shaaban and A. Maher, “Using the theory of planned behavior to predict the use of an upcoming public transportation service in Qatar,” Case Stud. Transp. Policy, vol. 8, no. 2, pp. 484–491, Jan. 2020.
[10] M. N. Borhan, A. N. H. Ibrahim, and M. A. A. Miskeen, “Extending the theory of planned behaviour to predict the intention to take the new high-speed rail for intercity travel in libya: Assessment of the influence of novelty seeking, trust and external influence,” Transp. Res. A. Policy Pract., vol. 130, pp. 373–384, Dec. 2019.

[11] M. Conner and A. Armitage, “Consciousness and the theory of planned behavior: Toward a more complete model of the antecedents of intentions and behavior,” Personality Social Psychol. Bull., vol. 27, no. 11, pp. 1547–1561, Nov. 2001.

[12] A. Madha, A. D. Salman, H. D. Hussain, N. Borhan, and R. A. O. K. Rahmat, “Analysis of travel behaviour in petaling Jaya, Malaysia: An application of the theory of planned behaviour,” Arch. Transpor., vol. 38, pp. 29–38, Jun. 2018.

[13] S. H. Lo, G. J. P. van Breukelen, G.-J. Y. Peters, and G. Kok, “Commuting travel mode choice among office workers: Comparing an extended theory of planned behavior model between regions and organizational sectors,” Travel Behav. Soc., vol. 4, pp. 1–10, May 2016.

[14] P. Jing, J. Wang, L. Chen, and Q. Zha, “Incorporating the extended theory of planned behavior in a school travel mode choice model: A case study of Shaoxing, China,” Transp. Plan. Technol., vol. 41, no. 2, pp. 1–19, 2017.

[15] S. F. Sönmez and A. R. Graefe, “Determining future travel behavior from past travel experience and perceptions of risk and safety,” J. Travel Res., vol. 37, no. 2, pp. 171–177, Nov. 1998.

[16] C. Y. Yin, C. F. Shao, and X. Q. Wang, “Influence of urban built environment on car commuting considering parking availability,” J. Jilin Univ., vol. 50, no. 2, pp. 115–120, 2020.

[17] A. Rahimi, G. Azimi, and X. Jin, “Examining human attitudes toward shared mobility options and autonomous vehicles,” Transp. Res. F. Traffic Psychol. Behav., vol. 72, pp. 133–154, Jul. 2020.

[18] M. Dijst, L. Böcker, and M.-P. Kwan, “Exposure to weather and implications and behavior,” Netw. Spatial Econ., vol. 20, no. 1, pp. 1–17, Mar. 2020.

[19] A. M. Zanni and T. J. Ryley, “The impact of extreme weather conditions on long distance travel behaviour,” Transp. Res. A. Policy Pract., vol. 77, pp. 305–319, Jul. 2015.

[20] J. Wu and H. Liao, “Weather, travel mode choice, and impacts on subway ridership in Beijing,” Transp. Res. A. Policy Pract., vol. 135, pp. 264–279, May 2020.

[21] J. R. Kenworthy and F. B. Laube, “Patterns of automobile dependence in cities: An international overview of key physical and economic dimensions with some implications for urban policy,” Transp. Res. A. Policy Pract., vol. 33, nos. 7–8, pp. 691–723, Sep. 1999.

[22] S. Cullinane and K. Cullinane, “Car dependence in a public transport dominated city: Evidence from Hong Kong,” Transp. Res. D. Transp. Environ., vol. 8, no. 2, pp. 129–138, Mar. 2003.

[23] P. Nars, “Urban structure matters: Residential location, car dependence and travel behaviour,” J. Environ. Plan. Manage., vol. 50, no. 4, pp. 567–573, Sep. 2016.

[24] S. Chakrabarti and E. J. Shin, “Automobile dependence and physical inactivity: Insights from the California household travel survey,” J. Transp. Health, vol. 6, pp. 262–271, Sep. 2017.

[25] F. L. Wang and D. G. Wang, “Characteristics and determinants of car use in Beijing,” Acta Geoge. Sin., vol. 69, no. 6, pp. 771–781, 2014.

[26] J. Humphreys and A. Ahern, “Is travel based residential self-selection a focus on methodologies,” Transp. Res. B. Methodol., vol. 42, no. 3, pp. 204–228, Mar. 2008.

[27] Y. Vilmaz and E. Ari, “The effects of service quality, image, and customer satisfaction on customer complaints and loyalty in high-speed rail service in turkey: A proposal of the structural equation model,” Transportmetrica A. Transp. Sci., vol. 13, no. 1, pp. 67–90, Jan. 2017.

[28] M. N. Borhan, A. N. H. Ibrahim, M. A. A. Miskeen, R. A. O. K. Rahmat, and A. M. Alhodaidi, “Predicting car drivers’ intention to use low cost airlines for intercity travel in Libya,” J. Air Transp. Manage., vol. 65, pp. 88–98, Oct. 2017.

[29] A. Rehn, A. Possenniers, and J. Holdsworth, “Hierarchical clustering for single-dimensional data using CUDA,” in Proc. Australas. Comput. Sci. Week Multiconf., Jan. 2018, pp. 1–10.

[30] C. Zha, Y. Dou, M. Guo, and Y. Dong, “A new hybrid clustering algorithm based on simulated annealing,” in Proc. 5th Int. Conf. Intell. Hum.-Mach. Syst. Cybern., Aug. 2013, pp. 94–99.

[31] X. Geng, Y. Zhang, Y. Jiao, and Y. Mei, “A novel hybrid clustering algorithm for topic modeling on Chinese microblogging,” IEEE Trans. Comput. Social. Syst., vol. 6, no. 2, pp. 289–300, Apr. 2019.

[32] Z. Zhao, K. Xu, G. Hui, and L. Hu, “An energy-efficient clustering routing protocol for wireless sensor networks based on AGNES with balanced energy consumption optimization,” Sensors, vol. 18, no. 11, pp. 3938, Nov. 2018.

[33] K. Swapna and P. Babu, “A critical study on cluster analysis methods to extract liver disease patterns in Indian liver patient data,” Int. J. Comput. Intell. Res., vol. 13, no. 10, pp. 2379–2390, 2017.

[34] P. Hu and X. G. Yang, “Multiple-factor perceived features of traffic quality influencing trip decision,” J. Traffic Transp. Eng., vol. 17, no. 2, pp. 117–125, 2017.

[35] H. Ren and H. Folmer, “Determinants of residential satisfaction in urban China: A multi-group structural equation analysis,” Urban Stud., vol. 54, no. 4, pp. 1407–1425, May 2017.

[36] L. Li and X. Wu, “Housing price and intergenerational co-residence in urban China,” J. Housing Econ., vol. 45, pp. 1–27, Sep. 2018.

[37] M. Namgung, B. J. Lee, A. Fujiwara, and J. Zhang, “Car preference model and analysis of international tourism behavior between South Korea and Japan-according to Asian-highway,” J. Korean Soc. Civ. Eng., vol. 27, no. 1D, pp. 27–33, May 2018.

[38] S. Hu, Q. Liang, H. Qian, J. Weng, W. Zhou, and P. Lin, “Frequent-pattern growth algorithm based association rule mining method of public transport travel stability,” Int. J. Sustain. Transp., vol. 15, no. 1, pp. 1–14, Oct. 2020.

[39] R. Arroyo, T. Ruiz, L. Mars, S. Rasaoli, and H. Timmermans, “Influence of values, attitudes towards transport modes and companions on travel behavior,” Transp. Res. F. Traffic Psychol. Behav., vol. 71, pp. 8–22, May 2020.
[54] P. A. Plazier, G. Weitkamp, and A. E. Van den Berg, “Cycling was never so easy!” An analysis of e-bike commuters’ motives, travel behaviour experiences using GPS-tracking interviews,” J. Transp. Geogr., vol. 65, pp. 25–34, Dec. 2017.

[55] W. H. Zhou and Z. X. Xu, “Discussion on effect from the development of electric bicycle on traffic configuration in domestic cities,” Commun. Standardization, vol. 141, no. 5, pp. 116–120, 2005.

[56] Y. Wang, A. Monzon, and F. D. Cimmino, “Assessing the accessibility impact of transport policy by a land-use and transport interaction model—The case of Madrid,” Comput., Environ. Urban Syst., vol. 49, pp. 126–135, Jan. 2015.

[57] B. B. Majumdar, M. Jayakumar, P. K. Sahu, and D. Potoglou, “Identification of key determinants of travel satisfaction for developing policy instrument to improve quality of life: An analysis of commuting in Delhi,” Transp. Policy, vol. 110, pp. 281–292, Sep. 2021.

[58] Y. Zhang, T. Thomas, M. J. G. Brussel, and M. F. A. M. Van Maarseveen, “Expanding bicycle-sharing systems: Lessons learnt from an analysis of usage,” PLoS ONE, vol. 11, no. 12, pp. 1–25, 2016.

[59] Q. Li, “The relationship between resident trip characteristics and spatial evolution of the ribbon group city: A case study in Lanzhou,” in Proc. CICTP, Jul. 2012, pp. 1–12.

[60] F. Zhang, Q. Shen, and K. J. Clifton, “Examination of traveler responses to real-time information about bus arrivals using panel data,” Transp. Res. Rec., J. Transp. Res. Board, vol. 2082, no. 1, pp. 107–115, Jan. 2008.

[61] J. Zhang, “Transport policymaking that accounts for COVID-19 and future public health threats: A PASS approach,” Transp. Policy, vol. 99, pp. 405–418, Dec. 2020.

[62] A. J. Filtness, E. Mitsopoulos-Rubens, and C. M. Rudin-Brown, “Police officer in-vehicle discomfort: Appointments carriage method and vehicle seat features,” Appl. Ergonom., vol. 45, no. 4, pp. 1247–1256, Jul. 2014.

[63] J. Zhao and X. Zhou, “Improving the operational efficiency of buses with dynamic use of exclusive bus lane at isolated intersections,” IEEE Trans. Intell. Transp. Syst., vol. 20, no. 2, pp. 642–653, Feb. 2019.

[64] S. H. K. Soltani, M. Sham, M. Awang, and R. Yaman, “Accessibility for disabled in public transportation terminal,” Proc. Social Behav. Sci., vol. 35, pp. 89–96, Jan. 2012.

SONG HU was born in Xingtai, Hebei, China, in 1992. He received the M.S. degree in transportation engineering from Chongqing Jiaotong University, China, in 2018. He studied at the University of Nevada, Reno for exchange visiting in the USA, in 2017. He is currently pursuing the Ph.D. degree in transportation engineering with Beijing University of Technology (BJUT), Beijing, China. His research interests include traffic data mining and travel behavior modeling.

JIANCHENG WENG received the M.S. and Ph.D. degrees in transportation planning and management from BJUT, in 2007. He studied at the University of Minnesota in systems for exchange visiting in the USA, in 2010. He has authored or coauthored more than 70 academic papers in prestigious international journals, including the International Journal of Sustainable Transportation, Networks and Spatial Economics, and Physica A: Statistical Mechanics and its Applications, and top-level conferences, such as Transportation Research Board Annual Meeting and IEEE Intelligent Transportation Systems Conference. He is currently a Professor with the Faculty of Intelligent Transport, BJUT. His research interests include traffic data mining, travel behavior modeling, and traffic operation analysis.

PHAM MINH CHAU received the M.S. degree in urban management from Ho Chi Minh City University of Architecture, Vietnam, in 2012. He is currently pursuing the Ph.D. degree in transportation planning with Beijing University of Technology (BJUT), Beijing, China. His research interests include traffic data mining and travel behavior modeling.

PENGFEI LIN received the B.S. degree in transportation engineering from Chongqing Jiaotong University, Chongqing, China, in 2015, and the Ph.D. degree in transport engineering, BJUT, in 2021. He studied at Vrije Universiteit Amsterdam, The Netherlands, in 2019, as a Government Student. He currently holds a postdoctoral position in transportation planning at BJUT. His research interests include traffic data modeling and intelligent transportation analysis.

WEI ZHOU received the M.S. and Ph.D. degrees in traffic engineering from Chang’an University, China, in 1991 and 2000, respectively. He is currently a Professor in transportation planning at Chang’an University and BJUT. He is also the Chief Engineer of the Ministry of Transport, China. His research interests include transportation planning and management.

* * *