Identifying the Group Differences in the Impact of Haze on Residents’ Low-Carbon Travel: Evidence From a Large-Scale Survey in China

Bin Zhang, Beijing Institute of Technology, China
Zhen Xu, Beijing Institute of Technology, China
Liran Sun, Beijing Institute of Technology, China

ABSTRACT
This paper matches the large-scale survey data and the corresponding historical weather data to explore how air pollution impacts on low-carbon travel choices. The K-means algorithm is employed to cluster the personal characteristics of residents into five groups according to their travel behavior. The authors take ordered Logit models to identify the group differences in the impact of haze on the five types of low-carbon travel choices, combining with the theory of responsibility attribution and protection motivation theory. The results show that haze has a significant impact on the two groups, namely young office workers and students. The other three groups will not consider the influence of haze when choosing travel vehicles, travel distance, and travel time. The quantity of personally owned automobiles also has a significant impact on the group differences in low carbon travel choices. It is indicated that low carbon travel policies should be considered in the group differences in the future, and efforts should be made from supply and demand sides to guide residents to choose low-carbon travel.

KEYWORDS
Group Recognition, Haze, Large-Scale Survey Data, Low Carbon, Travel Behavior

INTRODUCTION
Many countries encounter increasing pressure from air pollution to reduce the level of pollutants in the air systematically and appear less haze days. For instance, Mexico City complete a move to soot-free buses in 2018 to ensure cleaner vehicle standards are implemented, besides, it plans to ban private diesel cars by 2025. The main components of haze are sulfur dioxide, nitrogen oxides, PM$_{2.5}$ and PM$_{10}$. Among them, PM$_{2.5}$ will have an acute negative impact on human cardiovascular diseases (Mohammed et al., 2016). A report from the World Health Organization (WHO) website on May 2, 2018 shows that 9 out of 10 people breathe air containing high levels of pollutants, causing some 7 million deaths every year with air pollution. Motor vehicles have great importance in closed space and comfort. Haze weather can significantly reduce cycling, and encourage people to switch to other modes of travel, especially motorized mode of travel (Zhao et al., 2018). The China Motor Vehicle Environmental Management Annual Report (2017) points out that motor vehicle exhaust emissions
are the main source of air pollution, with the highest contribution even reaching 50%. Therefore, motor vehicles cause a large amount of carbon dioxide emissions, aggravating air pollution problems. Residents have a significant willingness to use low-carbon travel vehicles, however, group differences make residents have various intentions when choosing vehicles, such as the intention to use public travel vehicles is 57.9%, and the intention to use electric vehicles is 62% (Shi et al., 2017).

From 1981 to 2010, the annual average number of polluted days in central and eastern China shows an increasing trend. In 2013, the National Environmental Analysis of the People's Republic of China released by the Asian Development Bank and Tsinghua University mentioned that less than 1% of China's 500 large cities meet the WHO air quality standards. Air pollution has become an important issue related to national development and livelihood in China (Wang et al., 2021). The 2020 national motor vehicle and driver data released by the Transportation Administration of the Ministry of Public Security of China shows that the number of vehicles in 13 cities has exceeded 3 million, of which Beijing (6.032 million), Chengdu (5.457 million) and Chongqing (5.044 million) ranked first, second and third respectively. Facing the increasing traffic flow, China has proposed the vehicle restrictions policy. However, the impact of this policy on residents' low-carbon travel decisions is limited (Zhang et al., 2020). On September 6, 2021, the third round of PM$_{2.5}$ source analysis report released by Beijing showed that the contribution rate of mobile sources was 46%, mainly diesel vehicles and gasoline vehicles. Thus, this study uses the Fourth Comprehensive Traffic Survey in Beijing (2010), and the weather data of Beijing in the corresponding period.

At present, the extant studies focusing on the impact of haze on travel choices are less. Due to absent a classification perspective, the present understanding of low-carbon travel choices for different groups is limited. This paper aims to answer what is the impact of haze on residents' low-carbon travel under group differences. In this paper, low-carbon travel refers to travel modes with low carbon dioxide emissions. Our research takes an approach of cluster analysis to increase our understanding by using K-means algorithm to divide all samples into different groups with representative characteristics. Results gained from the research contribute to form the bridge between group differences and the impact of haze on travel choices. The choice of residents' travel is an important theoretical basis for the forecast of traffic supply and demand. Results also help government officials implement policies based on group differences. Our empirical analysis is considerably different from recent studies in two aspects. First, this research uses official data from large scale survey. Therefore, we can more effectively assess group differences in travel choices caused by haze. In so doing, we help to reduce some errors in the research. Second, we use both K-means algorithm and ordered Logit models in this paper.

**LITERATURE REVIEW**

This section reviews the existing literature investigating the impact of haze on residents’ low-carbon travel and the group differences and potential determinants of residents’ travel. The relationship between environment and traffic travel has attracted the attention of scholars. Previous studies have mostly focused on the impact of traffic travel emissions on air pollution. After the frequent occurrence of haze, increasing scholars have begun to explore whether haze affects traffic travel.

The research records of air pollution can be traced back to 1900. At that time, the main research was on the impact of industry on air pollution and the main hazards of air pollution. The impact on traffic was not mentioned until the full development of industry in the 1960s. After the vehicles becoming more mature, air pollution research expands to the field of transportation, and a large number of results have emerged. Haze is the result of the combined effect of natural conditions and human activities, becoming one of the most common air pollution forms. Since the Industrial Revolution, human activities such as transportation have caused a rapid increase in emission intensity and frequent haze weather. Although the PM$_{2.5}$ concept was put forward in 1997, it came into people's sight after
several years. The results by the ordinal logistic regression models show that PM$_{2.5}$ has significant and negative effects on residents’ life satisfaction (Shi et al., 2021).

Xiao et al. (2020) summarized relevant research on residents’ low-carbon willingness. A cross-sectional survey of the Shanghai Children Hospital and Jiading communities indicates that parents’ educational level and average annual household income influence residents’ attitudes towards air quality (Wang et al., 2015). Besides, another web survey of 410 respondents in Singapore finds that attention to media, interpersonal discussion, knowledge, and risk perception are positively associated with intention to take preventive measures on haze (Lin et al., 2017). Based on a cross-sectional design, the majority of 1404 respondents in Ghana are aware of the haze and its adverse effects on health (Odonkor & Mahami, 2020).

Haze pollution also affects the green behavior of residents (Zhang et al., 2019). When searching for keywords, the results mainly include research on traffic sign recognition technology and traffic safety in haze weather. There are few research literatures on the impact of haze on traffic travel choices. Campbell et al. (2016) proposes that using bicycles is affected by air quality, and shared bicycles and electric bicycles are affected to different degrees. The public generally believes that when haze weather occurs, private car owners will choose to use the car to travel for better protection and shorter travel time. However, low-income cyclists are less likely to switch to other travel modes on polluted days (Rind et al., 2015). Due to the tolerance to air pollution, they may still ride bicycles in heavy pollution. In comparison, physically vulnerable individuals may be more sensitive to polluted weather, especially the groups of children under 10 years old and the elderly (von Lindern et al., 2016; He et al., 2016). The functional characteristics of bicycles, buses, subways, cars and other travel vehicles are different, and the degree of air pollution is different. When the travel mode is bicycle, the morning commute has a significant weakening effect on air pollution (Apparicio et al., 2016).

Existing research shows that when studying the impact of haze weather on residents’ travel choices, the data level is relatively single, usually from various statistical yearbooks or simple statistical survey. It is difficult to conduct in-depth research from the level of residents’ travel choices. In addition, the existing research literature mostly uses relatively simple methods. It is difficult to analyze the compound influence and degree of multiple factors.

**RESEARCH METHODOLOGY**

**Data Source and Processing**

The large scale survey initiated by the Beijing Municipal Commission of Transportation, involves 46,900 participating families and 284,633 travel records. The scope of the survey covers all urban areas in Beijing, including the six main urban areas of Dongcheng, Xicheng, Chaoyang, Haidian, Fengtai and Shijingshan, as well as peripheral urban areas such as Shunyi and Tongzhou. 33,120 households are surveyed within the six districts of the city, and 13,780 households are surveyed in the outer suburbs, covering almost all sub-districts, counties and district offices in 16 districts and counties of Beijing. The content of the traffic census questionnaire is composed of four modules. The first module is the basic household information of the residents, which surveys the respondent’s detailed home address, basic information of family members, house ownership information and vehicle ownership information. The second module is a personal travel diary, which records the detailed travel history of residents over 6 years old from 3 am of the day to 3 am of the next day within 24 hours, including the travel destination and activity time of each traveler. The third module is the specific travel records of each traveler, including the traveler’s departure time and location, the mode of transportation used on the way, the time to reach each location, the length of stay, and the purpose of a single trip. The fourth module is the detailed information of the resident’s family, covering the gender, age, occupation of family members, industry, education, family income, driver’s license and other information. Historical weather data comes from the Beijing Historical Weather Network, which is recorded and uploaded by the grassroots monitoring points every three hours.
The traffic census covers a wide, detailed and multi-dimensional data, with the possibility of omissions and data contradictions. After data collection, the family and individual members are numbered, and basic information are combined. The vacancies and meaningless items in the data are deleted. After preprocessing, the traffic census data and historical weather data are integrated to lay the foundation for the next step.

**Cluster Variable Description**

The contradiction between the environment and the economy is intensifying. Many scholars are concerned about low-carbon travel, emerging endless researches on the influence of the characteristics of residents on transportation choices. The study finds that residents’ income level (Golob & Hensher, 1998), education level (Golob & Hensher, 1998), gender (Golob & Hensher, 1998), the number of motor vehicles (Susilo et al., 2012), the number of bicycles (Bohler et al., 2006) are important factors that affect travel choices. Based on previous research and consideration of the characteristics of real data, this paper selects the number of car ownership, the annual household income, the age of the residents, the education level of the residents, the occupational status of the residents, the travel time of the residents, and the gender of the residents as clustering variables. We use the residents’ own characteristics to perform cluster analysis, and extract the salient features of each group. The variables involved and the specific conditions of the comparison are shown in Table 1.

### Table 1. Characteristic variables

| Variables              | Code  | Remark                                                                 |
|------------------------|-------|-------------------------------------------------------------------------|
| number of cars owned   | Ncars | Number of cars owned by households                                     |
| family income          | Income| According to the overall family income level, it is divided into 7 income levels from low to high. |
| resident age           | Age   | Specific age of residents                                               |
| education level        | Education | According to the education level of the residents, the education level is arranged from low to high, and a total of 9 levels of education are divided. |
| profession situation   | Job   | According to the original questionnaire, set 14 types of occupation codes. |
| travel time            | Time  | Encoded as a time variable according to the specific conditions of residents’ travel time. |
| gender situation       | Gender| Code the gender of the residents, 1 is male and 2 is female.            |

**The Optimal Number of Clusters**

By clustering and dividing residents with similar crowd patterns, crowd characteristics can be extracted and in-depth exploration of the travel options of different groups of residents in the face of haze. The K-means algorithm is proposed by James MacQueen. It is suitable for clustering processing of large-scale data. With the K value becoming larger, the data objects will be divided into more groups. The cluster centers will increase, and the sum of squared errors (SSE) will gradually decrease. This paper uses the K-means algorithm to divide 88,398 sample data objects into K groups, calculates the distance from each sample data object to K centroids, classifies the sample data into the group with the smallest distance, and calculates the cluster centroids in a loop. When the sum of the squares of the distances from all the data objects to the centroid in the big group is the smallest, the
Clustering result is output. However, when using the traditional K-means algorithm, it is necessary to manually set the number of clusters K based on past experience as the initial starting condition of the algorithm. The data in this paper is rich, and when K is set subjectively based on experience, the actual distribution characteristics of the sample will not be fully considered, resulting in insufficient accuracy and scientific rationality of the clustering results. Therefore, the paper uses the elbow method to determine the optimal K value, input the sample data set and the approximate range of the K value for clustering, and calculate the SSE value corresponding to each K value. When the K value is smaller than the actual number of clusters, with the K value becoming larger, the SSE changes more drastically. When the K value is close to the true number of clusters, with the K value becoming larger, the SSE changes more alleviatively. When drawing an elbow diagram with tools, the point corresponding to the elbow with the largest turning point is the optimal K value point (Hartigan & Wong, 1979). The analysis shows that the optimal number of clusters is 5, and the elbow diagram is shown in Figure 1.

![Elbow diagram](image)

After determining the number of clusters K=5, using the K-means clustering method to divide all sample data objects into 5 groups. In the five groups, the data objects contained in each group are more evenly distributed, with 25,811, 8,203, 10,451, 21,511, and 22,422 sample data respectively. There is no magnitude difference in the data objects of each group, and the effective rate of K-means operation is 100%. The results are shown in Table 2.
According to the results of K-means clustering, the specific clustering calculation distribution results and the clustering characteristic data of each group during the clustering are sorted out. Using Tableau software to analyze the number of cars, the income level of the residents, the age, the education of the residents, occupation of residents, and residents’ travel time. In order to distinguish each group clearly, the sunrise-sunset divergent color palette is used to assign colors from blue to red to the first to fifth groups. The five lines in each figure represent the distribution trend of the five groups under each feature, as shown in Fig. 2.

### Table 2. The number of samples in the final cluster center

| Cluster Type | Numbers of Group |
|--------------|------------------|
| clustering   |                  |
| 1            | 25811.000        |
| 2            | 8203.000         |
| 3            | 10451.000        |
| 4            | 21511.000        |
| 5            | 22422.000        |
| effective    | 88398.000        |
| missing      | 0.000            |

---

**Clustering Result**

According to the results of K-means clustering, the specific clustering calculation distribution results and the clustering characteristic data of each group during the clustering are sorted out. Using Tableau software to analyze the number of cars, the income level of the residents, the age, the education of the residents, occupation of residents, and residents’ travel time. In order to distinguish each group clearly, the sunrise-sunset divergent color palette is used to assign colors from blue to red to the first to fifth groups. The five lines in each figure represent the distribution trend of the five groups under each feature, as shown in Fig. 2.

**Figure 2. Display of various clustering features**

It can be seen from Figure 2 (a) that in each group, the number of people without a car is much larger than the number of people who own cars, and the characteristics of the number of people without
a car in the first and fourth groups are particularly significant. No matter how many cars you own, the number of people in the first group occupies the highest position. Within the range of having one to three vehicles, the number of the first and fifth groups far exceeds that of other groups, and their characteristics are particularly prominent compared with the case of no vehicles.

It can be seen from Figure 2 (b) that the income level of the first group is the highest among the five groups, followed by the fourth and fifth groups, and the second and third groups have the lowest income levels. In other situations which except the lowest income level, the number of people in the first group ranks in the front row. Compared with other groups of people, with the increase in income levels, the trend of the number of people in the first group has not significantly decreased. The increase in income levels of the fourth and fifth groups results in slower changes in the number of people, becoming second only to the first group.

Figure 2 (c) shows the age distribution of the five groups of people. It can be clearly judged from the figure that, from low to high, the age stages of each group are in the second, first, fifth, fourth and third group in order.

Figure 2 (d) is the distribution of education levels of the five groups. It can be seen from the figure that the distribution characteristics of the five groups of people are more significant. The proportion of high education level in the first group is significantly higher than the proportion of low education level. Compared with other groups, the education level of the first group is significantly higher. The fifth group has a larger proportion of high education level, and a larger proportion of low education level, but the proportion of high education level is higher than that of other groups, and its education level is second only to the first group. The level of education in the third group is more evenly distributed, with the level of education being medium to high. The education level of the second group is relatively low.

Figure 2 (e) shows the occupational distribution of various types of residents. In terms of occupations with stable jobs and higher salaries such as corporate employees, civil servants and public institutions, the proportions in the first, fourth and fifth groups are significantly higher. In the third group, the proportion of servicers is relatively high. The characteristics of other groups are not particularly significant.

Figure 2 (f) shows the distribution of travel time for each group. In general, each group reaches the peak of travel around 7 am and 5 pm, but the duration of the peak is different. The peak duration of the third and fourth groups is longer, indicating that the urgency of travel is not as good as that of other groups, and the travel time is relatively free. The travel peaks of the first, second and fifth groups are particularly prominent, indicating that travel is regular and urgent.

After clustering and visualizing the data, it is concluded that each group has the following characteristics, as shown in Table 3. According to the clustering characteristics of various group in the number of cars, income level, age, education level, occupation, and travel time, the sample data objects are roughly divided into young office workers, students, retirees, freelancers and middle-aged office workers.
Model Theory Basis

The behavioral research of residents’ low-carbon travel involves multiple disciplines and even cross-disciplines. Many scholars have conducted research on residents’ low-carbon travel from economics, physics, management, psychology, urban optimization and other disciplines. From the perspective of psychology, this paper expounds the data results in combination with economic data and principles, mainly involving responsibility attribution theory and protective motivation theory.

The responsibility attribution theory comes from Hyde’s attribution theory. Hyde believes that any problem can be attributed to specific reasons, including the actor’s own internal reasons and external environmental factors. In addition, Hyde believes that attribution is the basis of the perpetrator’s responsibility, and the perpetrator is responsible for the results caused by self. Later, Hyde’s naive attribution theory based on personal qualities is widely used in the research of environment and people, deriving the responsibility attribution theory in environmental psychology. Responsibility attribution theory is mainly applied to environmental incidents. By inspiring individuals to think about the causes of environmental problems, it prompts individuals to realize that they are also the attribution of environmental problems, arouse their own environmental awareness, and take active actions to protect the environment. Sloot et al. (2018) finds that pro-environmental behaviors originate from the individual’s belief that the consequences of the environment are caused by himself, and he needs to work hard to maintain the environment. The research conclusions of Mir et al. (2016) show that pro-environmental beliefs and personal awareness of environmental regulations can lead to pro-environmental behaviors. Testers believe that cars will cause environmental degradation. Individuals as car users should be responsible for it and spontaneously reduce the use of cars.

The protective motivation theory was put forward by Rogers et al. in 1975. It emphasizes that when people evaluate information sources, they will make corresponding response patterns through cognition to determine whether they need to protect individuals. Research by Li et al. (2012) shows that people have a very strong sense of self-protection. When people face a situation that may have an adverse effect on themselves, they will react quickly, giving priority to protecting themselves and avoiding the impact of adverse situations. When Meng et al. (2016) studies cyclists, they find that in the face of weather conditions that are not conducive to riding, cyclists will choose to protect themselves and change to other means of travel vehicles based on specific conditions.

Regarding the influence of haze on the choice of travel modes, the existing research uses the construction of disaggregated models as the main research method. Disaggregated models are established based on the concepts of maximum utility and random utility. The most commonly used

| Group | Cars | Income Level | Age | Education Level | Profession | Time         | Name                |
|-------|------|--------------|-----|----------------|------------|--------------|---------------------|
| 1     | many | high         | medium to low | high | corporate employees | 7-8:00/17-18:00 | young office worker |
| 2     | little | medium to low | medium to low | low | no | 6-7:00/16-17:00 | student             |
| 3     | little | medium to low | high | more even | service | 8-10:00/12-16:00 | retiree             |
| 4     | medium | medium to high | medium to high | medium to low | no | 7-11:00 | free-lance          |
| 5     | many | medium to high | medium | higher | enterprise/ institution | 7:00/17:00 | middle-aged office worker |
types are multinomial Logit model, multinomial Probit model and nested Logit model. Wang et al. (2017) build a nested Logit model by introducing factors such as travel cost, travel distance, station convenience, and number of passengers. In order to measure the influence of values on people’s choice of transportation mode, Koryagin & Dekina (2014) establish a hierarchical Mixed-Logit model that includes latent variables. Taking into account the wide application of disaggregated models in the field of transportation and the research problems of the paper, we select multinomial Logit model to regress the differences in travel impacts of the five groups.

**Statistical Analysis of Variables**

According to the classification method from low-carbon to high-carbon, the travel modes are divided into low-carbon, medium-low-carbon, medium-high-carbon and high-carbon. From green to red, it means from low-carbon to high-carbon, and the densely marked areas are the main urban areas of Beijing. According to the longitude and latitude of the departure location and travel mode, we can know the low-carbon degree of the travel mode of Beijing residents on haze and non-haze days, as shown in Fig. 3. As can be seen from the figure, whether it is haze or not, carbon dioxide emissions of residents’ travel modes in peripheral urban areas are lower than residents’ travel modes in main urban areas. The peripheral urban areas are far away from the center of Beijing. If high-carbon travel vehicles are used for commuting or fun, the travel cost is relatively high. Moreover, the economic conditions of the peripheral urban areas are lagging behind the center. Residents are willing to take a low-carbon mode of travel in peripheral urban areas with short distances:

1. Traveling on non-haze days.
2. Traveling on haze days.

*Figure 3. Low-carbon travel level under different haze weather conditions in Beijing. Notes: Numbers in the legends stand for average values of carbon dioxide emissions. The meaning of leaving district is the urban area that residents live.*
In addition to the comparison between the main urban areas and the peripheral urban areas, the low-carbon travel modes between haze and non-haze days are also different. Compared with haze days, residents in most areas of Beijing have a lower-carbon travel mode on non-haze days. High-carbon travel methods such as cars are more protective. Affected by protection motivation theory, people are more willing to choose a form of travel that prioritizes self-protection when faced with the stimulus of haze weather, leading to a more high-carbon travel mode for Beijing residents under haze weather. It can be seen from Fig. 3 that when haze weather occurs, the colors change from low-carbon to high-carbon in main urban areas is the most significant. Generally speaking, residents living in the center have more favorable family conditions. They usually own high-carbon vehicles at home, and are more likely to use these vehicles.

Notes: Low-carbon category includes walking and bicycles; medium-low-carbon category includes electric vehicles, subways and buses; medium-high-carbon category includes shuttle buses, school buses and small passenger cars; high-carbon category includes the passenger and cargo vehicles, trains, motorcycles and taxis.

By further describing the low-carbon degree of choice of travel modes for different groups of residents under different weather conditions, it is possible to understand whether various groups of residents have changed their travel modes due to the haze weather, as shown in Fig. 4. In general, the proportion of travel among these five groups of residents under different weather conditions has not changed drastically. A separate analysis of each group shows that under the haze weather conditions, the proportion of freelancers and middle-aged office workers adopting low-carbon and medium-low-carbon travel modes decreases, the proportion of young office workers, students and retirees adopting low-carbon and medium-low-carbon travel modes increases. These conclusions are only derived from direct observation of the data without including other factors that may have an impact. Whether the conclusions are correct or not requires more accurate and powerful empirical research and testing in the regression model below.
MULTINOMIAL ORDINAL LOGIT REGRESSION MODEL

The problem studied in this paper is the impact of haze on the choice of residents’ travel mode. The dependent variable is the low-carbon degree of residents’ travel is an ordered multi-categorical variable. Therefore, the ordered Logitic model in multinomial Logit models is selected for this paper. The formula of the ordered Logitic model is as follows:

\[ y^* = a + \sum_{k=1}^{k} \beta_k x_k + \epsilon \]

In the formula, \( y^* \) is the specific internal reaction of the observed variable, \( \epsilon \) is the error term, \( x_k \) is the model independent variable, and \( \beta_k \) is the correlation coefficient of \( x_k \).

According to the group number \( J \) of \( y^* \), set \((J-1)\) demarcation points to calculate the cumulative probability of a given \( X \). Assuming that \( \epsilon \) is a logit distribution, the cumulative probability of a given \( X \) is as follows:

\[
P(y \leq j \mid x) = \frac{\exp[u_j - \alpha + \sum_{k=1}^{k} \beta_k x_k]}{1 + \exp[u_j - \alpha + \sum_{k=1}^{k} \beta_k x_k]}\]

Among them, \( P (y \leq j \mid x) = P (y \leq 1 \mid x) + P (y \leq 2 \mid x) + P (y \leq 3 \mid x) + P (y \leq 4 \mid x) \), \( \alpha = u_1 - \alpha \). After performing natural logarithmic transformation, since it is a logit model, the formula becomes \( P (y \leq j \mid x) = P (y = j \mid x) - P (y = j-1 \mid x) \).

The probability model formula of the final ordered Logitic model is shown in (3), where \( x_i \) is the independent variable, \( \alpha_i \) is the constant term, and \( \beta_{ij} \) is the regression correlation coefficient:

\[
P\left( Y \leq \frac{j}{x} \right) = \frac{\exp(a_i + \sum_{j=1}^{j} \beta_{ij} x_i)}{1 + \exp(a + \sum_{j=1}^{j} \beta_{ij} x_i)} (j = 1, 2, 3, 4)\]

In the multinomial ordered Logitic model, \( Y \) represents the four types of low-carbon travel for residents, namely low-carbon, medium-low-carbon, medium-high-carbon, high-carbon. In the previous section, the sample data objects are divided into five groups. However, a resident may have multiple travel records. It is not possible to judge the low-carbon degree of a resident’s traffic travel based on only one of the records. Therefore, it is necessary to match the existing clustering results with the travel records of a person to obtain a new regression sample. The independent variable \( x_i \) includes haze, rainfall, whether working day, travel time, travel distance, whether to have commonly used motor vehicles, whether to have the driver’s license, the number of bicycles and the number of motorcycles. The description of the dependent and independent variables is shown in Table 4.
RESULTS AND DISCUSSION

When using stata15.1 statistical software for regression analysis, stata defaults to the first one as the reference, that is taking low-carbon travel mode as the reference to analyze whether the haze affects the willingness of different residents to choose high-carbon travel mode, as well as making regression test on other factors that affect residents’ choices of low-carbon travel, the results are shown in Table 5. It can be seen from Table 5 that the lowest pseudo $R^2$ of the regression results of the five groups is 0.284. Generally, as long as the pseudo $R^2$ is above 0.2, the model has a good goodness of fit. The pseudo $R^2$ of the five groups are all higher than 0.2, indicating that the fitting of the five groups regression model is good and the model setting has strong scientific and rationality.

| Name                      | Code  | Remark                                                                 |
|----------------------------|-------|------------------------------------------------------------------------|
| low-carbon travel          | Y     | Assign values in the order from low-carbon to high-carbon, 1 means low-carbon, 2 means medium-low-carbon, 3 means medium-high-carbon, 4 means high-carbon. |
| haze                      | Haze  | Expressed by dummy variables, 0 means no haze, 1 means haze.           |
| rainfall                   | Rain  | Encode the amount of rainfall, 1 is rain and fog and no rain, 2 is showers, and 3 is obvious rain and heavy rain. |
| is it a working day         | Week  | Expressed by dummy variables, 1 means Monday to Friday, 0 means Saturday and Sunday. |
| travel time                | Time  | Encode the time, 1 means less than 10 minutes, 2 means 10 minutes to 30 minutes, 3 means 30 minutes to 60 minutes, 4 means more than 60 minutes. |
| travel distance            | Distance | Encode the distance, 1 means the distance is within two kilometers, 2 means the distance is between two kilometers and five kilometers, 3 means the distance is between five kilometers and ten kilometers, and 4 means the distance is above ten kilometers. |
| own common motor vehicles  | Car   | Expressed by dummy variables, 1 means there is at least one commonly used motor vehicle, 0 means there is no commonly used motor vehicle. |
| have the driver’s license  | Driving license | Expressed by dummy variables, 1 means there is the driver’s license, 0 means there is no driver’s license. |
| number of bicycles         | Bike  | Encode the quantity, 0 means no, 1 means one bicycle, 2 means two or more. |
| number of motor-cycles     | Motorcycle | Encode the quantity, 0 means no, 1 means one motorcycle, 2 means two or more. |
From Table 5, we can know that there is a significant negative correlation between haze and the low-carbon travel choices of students. The correlation coefficient is -0.175, which is significant at statistical levels of 5%. The stata software uses low-carbon travel as a reference. This shows that for

| Explanatory Variables | Young Off-Ice Worker | Student | Retiree | Freelance | Middle-Aged Office Worker |
|-----------------------|----------------------|---------|---------|-----------|---------------------------|
| Haze                  | -0.153***            | -0.175**| -0.038  | 0.009     | 0.067                     |
|                       | (-4.00)              | (-2.07) | (-0.35) | (0.17)    | (1.53)                    |
| Distance              | 0.209***             | 0.468***| 0.711***| 0.44***   | 0.274***                  |
|                       | (138.95)             | (77.58) | (86.44) | (123.18)  | (125.39)                  |
| Time                  | -2.884***            | -6.909***| -9.895***| -6.209***| -3.928***                 |
|                       | (-77.03)             | (-55.29)| (-61.60)| (-85.03)  | (-79.32)                  |
| Week                  | -0.053***            | -1.158***| -0.034  | -0.062***| -0.08***                 |
|                       | (-3.28)              | (-4.33) | (-0.90) | (-2.92)   | (-4.36)                   |
| Car                   | 0.763***             | 0.486***| -0.106* | 0.316***  | 0.659***                  |
|                       | (44.57)              | (14.43) | (-1.82) | (13.59)   | (35.25)                   |
| Bike                  | -0.092***            | -0.13***| -0.165***| -0.166***| -0.15***                 |
|                       | (-13.70)             | (-8.98) | (-9.75) | (-18.18)  | (-19.54)                  |
| Motorcycle            | 0.487***             | 0.413***| 1.088***| 0.682***  | 0.527***                  |
|                       | (11.69)              | (5.41)  | (10.75) | (16.27)   | (13.18)                   |
| Driving license        | 0.387***             | 0.089   | 0.202** | 0.411***  | 0.537***                  |
|                       | (23.79)              | (1.37)  | (2.46)  | (17.47)   | (29.70)                   |
| Rain                  | 0.04*               | -0.017  | -0.22***| 0.004     | 0.044*                    |
|                       | (1.82)               | (-0.36) | (-4.25) | (0.15)    | (1.82)                    |
| cut1                  | 0.207               | 0.616   | 0.678   | 0.962     | 0.771                     |
| cut2                  | 3.631               | 4.292   | 6.855   | 4.754     | 3.263                     |
| cut3                  | 8.256               | 9.777   | 10.397  | 12.327    | 10.139                    |
| pseudo $R^2$          | 0.284               | 0.392   | 0.54    | 0.434     | 0.341                     |
| chi-square test value | 53416.56            | 17050.95| 27075.56| 55049.62  | 54091.8                   |
| observed sample size  | 84144               | 21882   | 33069   | 69336     | 70020                     |
| P value               | 0                   | 0       | 0       | 0         | 0                         |

Notes: Numbers in parentheses are standard errors; *** p<.01, ** p<.05, * p<.1.
students, the occurrence of haze weather will make it more difficult to choose high-carbon travel. The individuals belonging to the youngest age group are more likely to use more of less-polluting modes and less of more-polluting modes (Saigal et al., 2021). Students need to go to school regularly, and their regular travel characteristics will not change easily. Schools and family backgrounds are beneficial to increase secondary education students’ environmental consciousness (Ntanos et al., 2018). Students are immersed in the strong environmental education atmosphere, and their sense of responsibility for environmental protection is stronger than the awareness of self-protection. Affected by responsibility attribution theory, when haze weather occurs, students are less likely to choose the high-carbon travel mode.

Besides, we can also know that there is a significant negative correlation between haze and the low-carbon travel choices of young office workers. The correlation coefficient is -0.153, which is significant at statistical levels of 1%. Compared with students, when haze occurring, the negative impact on young office workers’ choice of high-carbon travel is more significant. The engagement of employees in green behaviors is driven by autonomous motivation (Budzanowska-Drzewiecka & Tutko, 2021). Young office workers are at the stage of just entering society. Their education level and income level are the highest among the five groups. Efficient environmental education taken by students throughout their four study years in the university created an awareness and changes in their attitudes (Boca & Saracli, 2019), generally believing that their choice of high-carbon travel will further deteriorate the environment. The occurrence of haze will also stimulate the sense of responsibility of young office workers, prompting them to choose low-carbon travel mode.

It can also be seen from the table that haze has no significant impact on the choice of travel modes for retirees, freelancers and middle-aged office workers. Travel distance, travel time, number of bicycles and number of motorcycles have a significant impact on the travel modes of the three groups. When the haze occurs, young office workers and students are unwilling to choose high-carbon travel methods that are prone to congestion due to the needs of school and commuting. However, retirees and freelancers do not have travel time requirements, their travel time is longer and uneasily affected. The influence of whether is weekend or not to the travel choices of retirees is not significant, indicating that the elderly does not depend on week to make travel choices. Compared with young office workers and students, retirees, freelancers and middle-aged office workers are older and have formed a complete self-regulation system, which is uneasy to change travel mode due to the non-rapidly harmful weather.

There is a significant positive correlation between travel distance and the low-carbon travel choices of the five groups, indicating that the increase in travel distance will prompt travelers to make high-carbon travel options. The longer the travel distance, the more people pay attention to the comfort of travel. In addition, the convenience of cars is high, avoiding transfers on the way, which helps people reduce their exposure to the haze environment. There is a significant negative correlation between the number of bicycles and the low-carbon travel options of the five groups, indicating that the increase in the number of bicycles will make people become more willing to adopt low-carbon travel. The number of motorcycles and whether they have the driver’s license have a significant positive impact on the low-carbon travel choices of the five groups, indicating that with the increase in the number of motorcycles and when travelers have the driver’s license, people will be more willing to adopt high-carbon travel. Whether there are commonly used motor vehicles has different influence on low-carbon travel choices of different groups. In addition, weather characteristics have a significant impact on the choice of travel mode.

In order to ensure the validity and reliability of the conclusions, this paper adopts the multinomial ordered Probit regression model to analyze the existing sample data and conduct robustness test. The results are shown in Table 6. Comparing Table 5 with Table 6, it can be seen that the sign and significance level of the key variable of haze have not changed much. Only its coefficient has changed slightly. Other variables still have a significant impact on residents’ choice of low-carbon travel mode. The overall result has strong stability.
Table 6. Ordered Probit regression results of the impact of haze on residents’ low-carbon travel choices

| Explanatory Variables | Young Off-Ice Worker | Student | Retiree | Freelance | Middle-Aged Office Worker |
|-----------------------|----------------------|---------|---------|-----------|--------------------------|
| Name                  | Young Off-Ice Worker | Student | Retiree | Freelance | Middle-Aged Office Worker |
| Haze                  | -0.093***            | -0.175** | -0.025  | 0.009     | 0.44*                    |
|                       | (-4.26)              | (-2.31) | (-0.46) | (0.32)    | (1.76)                   |
| Distance              | 0.088***             | 0.468*** | 0.229*** | 0.141***  | 0.099***                 |
|                       | (163.94)             | (94.90) | (115.62)| (158.58)  | (149.37)                 |
| Time                  | -0.847***            | -6.909*** | -2.06*** | -1.181*** | -0.917***                |
|                       | (-77.44)             | (-48.91)| (-61.71)| (-73.87)  | (-71.82)                 |
| Week                  | -0.041***            | -1.158*** | -0.045  | -0.052*** | -0.046***                |
|                       | (-4.45)              | (-4.21) | (-2.41) | (-4.56)   | (-4.43)                  |
| Car                   | 0.21***              | 0.486   | -0.163* | 0.127***  | 0.161***                 |
|                       | (8.55)               | (5.52)  | (-1.53) | (3.60)    | (5.97)                   |
| Bike                  | -0.06***             | -0.13*** | -0.086*** | -0.098*** | -0.093***                |
|                       | (-15.74)             | (-8.28) | (-10.19)| (-20.14)  | (-21.40)                 |
| Motorcycle            | 0.355***             | 0.413*** | 0.719*** | 0.5***    | 0.457***                 |
|                       | (15.94)              | (5.82)  | (14.88) | (22.73)   | (21.22)                  |
| Driving license        | 0.228***             | 0.089**  | 0.234*** | 0.309***  | 0.381***                 |
|                       | (24.79)              | (2.25)  | (6.04)  | (24.53)   | (36.25)                  |
| Rain                  | 0.03**               | -0.017  | -0.113*** | 0.004     | 0.019                    |
|                       | (2.45)               | (-0.59) | (-4.40) | (0.25)    | (1.43)                   |
| cut1                  | 0.018                | 0.039   | 0.035   | 0.022     | 0.021                    |
| cut2                  | 0.019                | 0.042   | 0.045   | 0.024     | 0.022                    |
| cut3                  | 0.024                | 0.055   | 0.053   | 0.035     | 0.03                     |
| pseudo $R^2$          | 0.253                | 0.331   | 0.46    | 0.369     | 0.297                    |
| chi-square test value | 47555.23             | 14383.35| 23039.24| 46858.38  | 47118.85                 |
| observed sample size  | 84144                | 21882   | 33069   | 69336     | 70020                    |
| P value               | 0                    | 0       | 0       | 0         | 0                        |

Notes: Numbers in parentheses are z-statistics; *** p<.01, ** p<.05, * p<.1.
CONCLUSION

Regarding the relationship between transportation and the environment, past studies mainly focused on how emissions from traffic travel would affect air pollution such as haze. With the frequent occurrence of haze weather, scholars have begun to pay attention to the adverse effects of haze on traffic travel. However, the conclusions of various scholars are inconsistent and even contradictory. This research uses data from large scale survey to explore the group differences in the haze impact on low-carbon travel choices. The main conclusions of this paper are as follows.

When haze weather occurs, young office workers and students are less likely to choose high-carbon travel, and young office workers are more likely to adopt low-carbon travel than students. The common feature of young office workers and students is that they all belong to the age when they have just started to work or are still studying at school. These two groups have received long-term education about environmental protection, and have high acceptance of environmental protection concepts. In the haze days, young office workers and students are significantly affected by the theory of responsibility attribution, realizing that their choice of high-carbon transportation will continue to aggravate the haze, and are more willing to choose low-carbon transportation when traveling.

Under the haze weather, the groups of retirees, freelancers and middle-aged office workers will not be significantly affected in the choice of travel modes. These three types of groups are aged, with more travel restrictions, and most of them have the established normative system. When faced with haze weather which is not directly harmful, retirees, freelancers and middle-aged office workers are unwilling to easily change their original travel choices.

Resident’s travel characteristics have an impact on the choice of travel modes. Travel distance, the quantity of personal owned motorcycles and license availability are all positively correlated to the low-carbon travel options of the five groups. Comparatively, the quantity of personal owned bicycles shows a negative correlation. Whether or not to own a motor vehicle has the least significant impact on the travel choices of retirees.

Policy Suggestions

Based on the starting point of green environmental protection, some governments have proposed many guiding policies in recent years. However, there is no immediate publicity effect. Based on the research conclusions, this paper puts forward the following policy suggestions.

For students, the government should form an immersive propaganda atmosphere of environmental protection education in the whole society, to encourage students make low-carbon travel choices. By long-term immersive environmental protection publicity, students will strengthen their understanding and memory of environmental protection. Moreover, students will clarify their status and responsibilities in the environment by means of introduction to haze consequences. After promoting environmental protection concepts in a long period, the government should cultivate students’ environmental awareness, and guide them to drive family members and friends around to choose low-carbon travel.

For the groups of office workers, the government should consider young office workers and middle-aged office workers respectively. Young office workers always have high acceptance of environmental protection concepts. Therefore, the government should provide travel with low-carbon and quickness. Middle-aged office workers usually have their own private cars for family travel. Faced with this situation, the government ought to promote middle-aged office workers to purchase new energy vehicles. At the same time, a restraint of driving cars on specific days should be taken. The government should also actively set up bicycles parking spots in densely populated areas to encourage office workers to choose bicycles for travel.

For the groups of retirees and freelancers, the government should decrease the costs of new energy vehicles and provide multiple subsidies to enhance the attractiveness of low-carbon travel. When facing these two groups, environmental protection education should be increased, for example new
publicity methods including short video APP and other network software. In addition, the government should provide smoother travel conditions for low-carbon vehicles such as bicycles and public traffic.

**Future Research Directions**

Our literature review and proposed models enable us to highlight some opportunities for future research. First, data from large scale survey is only sourced from Beijing, which shows that the suggestions is regional. Therefore, future research is supposed to expand the scope of regions and extract the characteristics of different regions. One major direction for future research is merging key characteristics of Beijing and other regions. The other area, which needs attention is willingness of the sample to travel when conducting survey. The findings can be essential for academia and practice because the differences between regions influence effectiveness of suggestions.

Second, this study shows that the survey has mainly collected residents’ travel conditions in one day. Although results from this research can enhance our understanding of haze influence in the short term, it can also prevent researchers from conducting long-term follow-up research on the residents. Because the signal theory shows that the impact of events will weaken with the passage of time, this research may not illustrate long-term impact of haze. Therefore, future researchers can extend the research period to offer a more accurate and comprehensive understanding of residents’ travel choices in a long period.

Third, our research shows that the responsibility attribution theory and the protective motivation theory are main theory basis to analyze models. However, psychological factors have not been involved in this survey of Beijing, resulting in a decline in the simulation accuracy of models. Future research needs to add questions and options related to psychological factors in the survey. The findings could be insightful for the relationship between haze and residents’ low-carbon travel.

**ACKNOWLEDGMENT**

This study is supported by National Natural Science Fund of China (Reference No. 72174023, 72141302, 71774014, 91746208, 71521002), Key Research Projects of Philosophy and Social Sciences of China Ministry of Education (Reference No.21JZD027), Beijing Social Science Foundation Project (Reference No.20JCC108, 20JJC023), National Science Fund for Distinguished Young Scholars (Reference No. 71625003), National Social Science Fund of China (Reference No. 17ZDA065), MOE Project of Humanities and Social Sciences for Young Scholars(Reference No. 20YJC630056), Science and Technology Innovation Program for innovative talents in Beijing Institute of Technology (Reference No.2021CX0125), Joint Development Program of Beijing Municipal Commission of Education, and Science and Technology Project of the Ministry of Housing and Urban-Rural Development of the People’s Republic of China (Reference No. 2021-K-106).
REFERENCES

Apparicio, P., Carrier, M., Gelb, J., Seguin, A. M., & Kingham, S. (2016). Cyclists’ exposure to air pollution and road traffic noise in central city neighbourhoods of Montreal. *Journal of Transport Geography, 57*, 63–69. doi:10.1016/j.jtrangeo.2016.09.014

Boca, G. D., & Saracli, S. (2019). Environmental Education and Student’s Perception, for Sustainability. *Sustainability, 11*(6), 1553. Advance online publication. doi:10.3390/su11061553

Bohler, S., Grischkat, S., Haustein, S., & Hunecke, M. (2006). Encouraging environmentally sustainable holiday travel. *Transportation Research Part A, Policy and Practice, 40*(8), 652–670. doi:10.1016/j.tra.2005.12.006

Budzanowska-Drzewiecka, M., & Tutko, M. (2021). The impact of individual motivation on employee voluntary pro-environmental behaviours: The motivation towards the environment of Polish employees. *Management of Environmental Quality, 32*(5), 929–948. doi:10.1108/MEQ-11-2020-0268

Campbell, A. A., Cherry, C. R., Ryerson, M. S., & Yang, X. M. (2016). Factors influencing the choice of shared bicycles and shared electric bikes in Beijing. *Transportation Research Part C, Emerging Technologies, 67*, 399–414. doi:10.1016/j.trc.2016.03.004

Golob, T. F., & Hensher, D. A. (1998). Greenhouse gas emissions and Australian commuters’ attitudes and behavior concerning abatement policies and personal involvement. *Transportation Research Part D, Transport and Environment, 3*(1), 1–18. doi:10.1016/S1361-9209(97)00006-0

Hartigan, J. A., & Wong, M. A. (1979). Algorithm AS 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series A (General), 28*(1), 100–108. doi:10.2307/2346830

He, G. J., Fan, M. Y., & Zhou, M. G. (2016). The effect of air pollution on mortality in China: Evidence from the 2008 Beijing Olympic Games. *Journal of Environmental Economics and Management, 79*, 18–39. doi:10.1016/j.jeem.2016.04.004

Koryagin, M., & Dekina, A. (2014). Optimization of the road capacity and the public transportation frequency which are based on Logit-Model of travel mode choice. *Information Technologies and Mathematical Modelling, 487*, 214–222. doi:10.1007/978-3-319-13671-4_26

Li, Y. J., Kenrick, D. T., Griskevicius, V., & Neuberg, S. L. (2012). Economic decision biases and fundamental motivations: How mating and self-protection alter loss aversion. *Journal of Personality and Social Psychology, 102*(3), 550–561. doi:10.1037/a0025844 PMID:22003837

Lin, T. T. C., Li, L., & Bautista, J. R. (2017). Examining how communication and knowledge relate to Singaporean youths’ perceived risk of haze and intentions to take preventive behaviors. *Health Communication, 32*(6), 749–758. doi:10.1080/10410236.2016.1172288 PMID:27392280

Meng, M., Zhang, J., Wong, Y. D., & Au, P. H. (2016). Effect of weather conditions and weather forecast on cycling travel behavior in Singapore. *International Journal of Sustainable Transportation, 10*(9), 773–780. doi:10.1080/15568318.2016.1149646

Mir, H. M., Behrang, K., Isaii, M. T., & Nejat, P. (2016). The impact of outcome framing and psychological distance of air pollution consequences on transportation mode choice. *Transportation Research Part D, Transport and Environment, 46*, 328–338. doi:10.1016/j.trd.2016.04.012

Mohammed, M. O. A., Song, W. W., Ma, W. L., Li, W. L., Li, Y. F., Afed, U. K., Ibrahim, M. A. E. M., Osman, A. M., Ahmed, A. A., & Ambuchi, J. J. (2016). Potential toxicological and cardiopulmonary effects of PM$_{2.5}$ exposure and related mortality: Findings of recent studies published during 2003-2013. *Biomedical and Environmental Sciences, 29*(1), 66–79. doi:10.3967/bes2016.007 PMID:26822514

Ntanos, S., Kyriakopoulos, G. L., Arabatzis, G., Palios, V., & Chalikias, M. (2018). Environmental behavior of secondary education students: A case study at central Greece. *Sustainability, 10*(5), 1663. Advance online publication. doi:10.3390/su10051663

Odonkor, S. T., & Mahami, T. (2020). Knowledge, attitudes, and perceptions of air pollution in Accra, Ghana: A critical survey. *Journal of Environmental and Public Health, 2020*, 1–10. Advance online publication. doi:10.1155/2020/3657161 PMID:32104187
Rind, E., Shortt, N., Mitchell, R., Richardson, E. A., & Pearce, J. (2015). Are income-related differences in active travel associated with physical environmental characteristics? A multi-level ecological approach. *The International Journal of Behavioral Nutrition and Physical Activity, 12*(1), 73. Advance online publication. doi:10.1186/s12966-015-0217-1 PMID:26044262

Saigal, T., Vaish, A. K., & Rao, N. M. (2021). Is the choice of less-polluting modes of transport for non-work purposes affected by socio-demographic factors? Evidence from India. *Management of Environmental Quality, 32*(3), 488–505. doi:10.1108/MEQ-09-2020-0212

Shi, H. X., Fan, J., & Zhao, D. T. (2017). Predicting household PM$_{2.5}$-reduction behavior in Chinese urban areas: An integrative model of theory of planned behavior and norm activation theory. *Journal of Cleaner Production, 145*, 64–73. doi:10.1016/j.jclepro.2016.12.169

Shi, X. M., Li, X. P., Chen, X. Y., & Zhang, L. P. (2021). Objective air quality index versus subjective perception: Which has a greater impact on life satisfaction? *Environment, Development and Sustainability*. Advance online publication. doi:10.1007/s10668-021-01730-4

Sloot, D., Jans, L., & Steg, L. (2018). Can community energy initiatives motivate sustainable energy behaviours? The role of initiative involvement and personal pro-environmental motivation. *Journal of Environmental Psychology, 57*, 99–106. doi:10.1016/j.jenvp.2018.06.007

Susilo, Y. O., Williams, K., Lindsay, M., & Dair, C. (2012). The influence of individuals’ environmental attitudes and urban design features on their travel patterns in sustainable neighborhoods in the UK. *Transportation Research Part D, Transport and Environment, 17*(3), 190–200. doi:10.1016/j.trd.2011.11.007

Wang, R., Yang, Y. Y., Chen, R. J., Kan, H. D., Wu, J. Y., Wang, K. R., Maddock, J. E., & Lu, Y. A. (2015). Knowledge, attitudes, and practices (KAP) of the relationship between air pollution and children’s respiratory health in Shanghai, China. *International Journal of Environmental Research and Public Health, 12*(2), 1834–1848. doi:10.3390/ijerph120201834 PMID:25664694

Wang, S. J., Zhou, H. Y., Hua, G. H., & Wu, Q. (2021). What is the relationship among environmental pollution, environmental behavior, and public health in China? A study based on CGSS. *Environmental Science and Pollution Research International, 28*(16), 20299–20312. doi:10.1007/s11356-020-11951-1 PMID:33405117

Wang, Y., Yan, X. D., Zhou, Y., & Xue, Q. W. (2017). Influencing mechanism of potential factors on passengers’ long-distance travel mode choices based on structural equation modeling. *Sustainability, 9*(11), 1943. Advance online publication. doi:10.3390/su9111943

Xiao, J., Zhen, Z. L., Tian, L. X., Su, B., Chen, H., & Zhu, A. X. (2020). Green behavior towards low-carbon society: Theory, measurement and action. *Journal of Cleaner Production, 278*, 123765. Advance online publication. doi:10.1016/j.jclepro.2020.123765

Zhao, P. J., Li, S. X., Li, P. L., Liu, J. X., & Long, K. F. (2018). How does air pollution influence cycling behaviour? Evidence from Beijing. *Transportation Research Part D, Transport and Environment, 63*, 826–838. doi:10.1016/j.trd.2018.07.015