

Research Article

Modelling the Air Ticket Purchase Behavior Incorporating Latent Class Model

Jing Lu ¹, Zheqing Wang ¹, Yingjing Gu ² and Wenhui Yang ³

¹College of Civil Aviation, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China

²Business School, Nanjing Xiaozhuang University, Nanjing 211171, China

³IT Department, Beijing Capital International Airport, Beijing 100621, China

Correspondence should be addressed to Yingjing Gu; guyingjing@126.com

Received 8 August 2019; Revised 14 November 2019; Accepted 17 December 2019; Published 27 February 2020

Academic Editor: Neale R. Smith

Copyright © 2020 Jing Lu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

As passengers are proved to be preference heterogeneous in air travel, this paper tries to model the air ticket purchase behavior incorporating market segmentation. In the research, a latent-NL model, established on the latent class structure and the nested logit model, integrates the personal features as well as the purchase preferences into the forecast of segment-specific purchase probability. In order to calibrate the model, a stated preference survey is designed with the choice profiles using real service information, and the survey is conducted in four cities in China for data collection. The results show that the proposed model provides an effective approach for predicting the air travel demand in particular for air ticket pricing, and the estimation results outperforms the traditional-nested logit model with higher goodness-of-fit. Besides, the model is then adopted to test the efficiency of different pricing strategies, showing its advantages in improving the flight revenues.

1. Introduction

The air ticket price is acknowledged as the key determinate of airline profits in terms of affecting air travel demand [1], it has been constantly optimized to satisfy passengers' needs since the market liberalization of civil aviation [2, 3]. For the airlines, the pricing strategies are traditionally proposed on the assumption that passengers are homogeneous or have the same needs to all the flights as well as other alternative travel modes [4], which may due to the purpose of convenient data collection and easy estimation [5].

However, reported by recent literatures, passengers show the distinct classified preference associated with their travel purpose [6], experiences or even demographic characteristics such as age, income, and occupation [7]. So, it is reasonable to believe that the air transportation market is actually composed of groups of passengers whose air ticket purchase preferences are in similarity [8]. In the sense, the pricing strategies should be proposed on the group-based air ticket purchase behavior.

A lot of studies have focused on the air ticket purchase behavior, the most widely adopted method is the discrete choice model [9]. According to the results, the multinomial logit model (MNL) and the nested logit model (NL) as well as the mixed logit model have the advantages in predicting the probability of purchasing an air ticket with specific price and other characteristics (such as departure and arrival airports, flight time, and on-time performance) [10, 11]. However, the discrete choice models are constructed on the assumption that passengers' preferences are homogeneous, thus the modelling results may overestimate some travelers' price endurance or underestimate others' price tolerance, which may cause the loss in flight demand as well as the profit [7]. Therefore, the group identification method, also recognized as the market segmentation method, should be linked to the discrete choice model for predicting the segment-specific air ticket purchase behavior.

The concept of market segmentation was arisen by Smith in the middle of 19th century, which helps the merchants to diversify their products as well as sales strategies [12]. For the

market segmentation in civil aviation, the conventional way is to divide the population statistically regarding passengers' personal features or travel purpose [13]. Although such method is easily conducted, but the simple classification may not be enough to reveal the complex heterogeneity in the air ticket pricing. For instance, passengers who have higher income may not choose the expansive air tickets in business trip for the limited budget [14], so setting higher air ticket prices for business travelers may be a wrong strategy in some situations. Hence, it is necessary to construct a specialized market segmentation approach for the air ticket pricing by incorporating both of the personal features and the purchase preferences. As the purchase preferences can be well modelled by the discrete choice model, we should find an appropriate market segmentation method, which can be combined with the discrete choice models in our research.

Given the existed market segmentation approaches, the clustering model as well as the latent class model are the two well-known methods. The clustering method uses the statistical correlation to empirically describe the market segmentation, so it is hard to be linked to the econometrical models. However, the latent class model can do the model-based classification which gives out the probability that a passenger belongs to a specific market segment [15], and it is estimated based on the probabilistic distribution of the survey data. In this sense, the latent class model should be a suitable method to be linked to other probabilistic models such as the discrete choice model [16–19].

Therefore, our paper aims to combine the latent class model and the discrete choice model to identify the market segments and to predict the segment-specific purchase behavior for air ticket pricing. Although the latent-discrete choice models have been applied in other fields of research [20, 21], there are also two main challenges in our study: the first is how to construct the suitable latent-discrete choice model with appropriate segment membership function and utility function for air ticket pricing; the second is to find out how to use the estimated results in improving the air ticket pricing strategies.

To solve the abovementioned challenges, our paper is organized as follows. Section 2 reviews previous studies about the latent class model and discrete choice models; Section 3 establishes the latent-discrete choice model for interpreting the segment-based air ticket purchase; Section 4 introduces our data collection and the case in our study; Section 5 shows the estimation results of the proposed model, and Section 6 gives out the sensitivity analysis; finally, Section 7 tries to summarize the limitations and our future research studies.

2. Literature Review

To establish the latent-discrete choice model, we should firstly make clear the mechanism of the latent class model as well as the discrete choice model. The latent class model was originally proposed by Kamakura and Russell [22] to distinguish the households with different brand preferences. In the model, the classification is identified on two probabilities: segmental probability and the conditional probability.

To be specific, the segmental probability represents the chance that an individual belongs to one defined group [23], and it can be estimated by maximizing the loglikelihood of the joint probability of the observable variables [16]. Then, the joint probability is calculated using the posterior probability of an individual who has one specific feature on the condition that he or she belongs to one segment. Moreover, the segment number of the latent class model is determined randomly according to the requirements of the research as well as the efficiency evaluation criteria [24].

Then, for the discrete choice model, it has been widely applied in the research of air travel behavior since Skinner [25] used the MNL model to analyze the airport choice behavior within a multi-airport region. For evolving about 40 years, a lot of researchers have used the MNL, NL, Probit, and mixed logit models to analyze passengers' airport choice, air route choice, airline choice, etc. in different contexts [26]. As the air ticket purchase behavior is closely associated with the passengers' flight choice, here we will refer to the literatures about flight-choosing behavior which use the discrete choice structures.

In the existed literatures about flight choice, most studies only incorporate the service attributes of different flights; however, the car, train, and even the bus may have influences on the choice preferences [27]. Therefore, the prices and other service characteristics of the substitutable travel modes should also be taken into account in the proposed air ticket purchase model. Because there exists strong correlation between different alternatives within one travel mode [28, 29], the nested logit model with a hierarchical structure is more suitable to analyze the air ticket purchase behavior in our research.

Based on the latent class as well as the discrete choice model, some literatures have established the joint structure integrating the two models to predict passengers' heterogeneous travel behavior [30–32]. For the research studies related to the civil aviation, Wen and Lai [33] applied the latent class model into the MNL structure to predict passengers' air carrier choice. According to their results, the proposed latent-MNL model has advantages over the multinomial logit model for better fitness. Meanwhile, their another study constructed latent class nested logit model in the travel mode choice and proved that the model could relax the IIA (Independent from Irrelevant Alternative) property of logit models [24]. Although the latent nested logit has been applied to analyze the travel behavior [34], this paper will construct a specialized latent-NL model for predicting the segment-specific air ticket purchase for the airlines' air ticket pricing.

When the latent-NL structure is confirmed, the next important work is to choose the variables of the model. For the latent class part, passengers' demographic characteristics and travel experiences are commonly selected to identify the market segmentation [35]. In the nested logit part, the air route, flight time, air fare, and flight frequency as well as the ground access time are the main factors influencing air ticket purchase [36, 37]. Furthermore, service attributes of other travel modes such as the ticket price, the departure time, and in-vehicle time would also be incorporated [38].

Besides, the data collection is another task of establishing the latent-NL model, as reported by existing literatures, and the stated preference survey (SP) has the advantages of collecting experimental data when the real travel data is hard to obtain, but the revealed preference survey (RP) could explicitly reveal the choice behavior in reality [39]. Considering the difficulty in collecting the previous real travel data in reality, we decided to use the SP survey designed with profiles using real travel alternatives and information in our research.

Based on the abovementioned literature review, Section 3 will introduce the proposed latent-NL model and the estimation method.

3. Modelling Approach

3.1. Model Structure. The structure of the latent-NL model is as shown in Figure 1. In the figure, the whole demand market would be classified into several segments based on the latent class structure, and each segment would have its own air ticket purchase model in terms of the nested logit model. With the model, we could obtain the probability of a passenger who belongs to one segment and purchases a flight with the specific air ticket price.

The aforementioned probability would be calculated on two functions: one is the segment membership function of the latent class structure and the other is the utility function of the NL model. It needs to point out that the parameters of the two functions are jointly estimated with the mechanism in Figure 2. We can see that the joint estimation is the key to construct the latent-NL model, thus to determine the market segmentation and the segment-based air ticket purchase model simultaneously.

3.2. Model Construction. In the model construction, the probability of individual i who belongs to segment s purchases alternative j ($P_{ij|s}$) which is written as follows:

$$P_{ij|s}(x_{ij} = 1 | y_{is} = 1) = p(j | y_{is} = 1) \cdot p(y_{is} = 1). \quad (1)$$

In the model, the variables are defined as follows:

$P_{ij|s}$: probability of individual i who belongs to segment s and purchases alternative j

x_{ij} : dummy variable, equals to 1 when individual i chooses alternative j and 0 otherwise

y_{is} : dummy variable, equals to 1 when individual i belongs to segment s and 0 otherwise

So, $P(j | y_{is} = 1)$ represents the probability that individual i purchases the ticket of flight j , on the condition that individual i belongs to segment s , which can be calculated by the utility function of the nested logit model; $P(y_{is} = 1)$ gives out the probability that i belongs to segment s , which can be calculated by segment membership function of the latent class structure.

First, for the probability $P(y_{is} = 1)$, it is constructed on the segment membership function, and we define the related variables as follows:

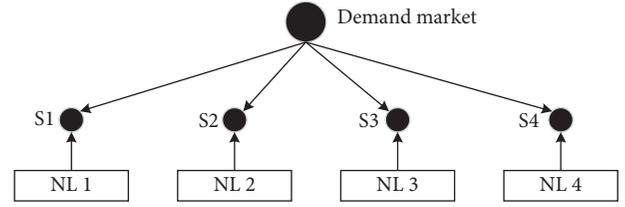


FIGURE 1: The structure of the latent-NL model.

u_{is} : the benefits of individual i belonging to segment s

m_i : variables of describing the personal features of individual i

γ_s : the parameters of m_i for segment s

δ_s : the error variable of the benefit u_{is}

The segment membership function is formulated as equation (2), which establishes the relationship between u_{is} and the personal feature variables m_i . Meanwhile, the probability of individual i belonging to segment s is as equation (3):

$$u_{is} = \gamma_s m_i + \delta_s, \quad (2)$$

$$P(y_{is} = 1) = \frac{\exp(\gamma_s m_i)}{\sum_{s=1}^S \exp(\gamma_s m_i)}. \quad (3)$$

Second, $P(j | y_{is} = 1)$ is predicted in the nested logit frame. The structure of the hierarchical air ticket purchase behavior is described in Figure 3. In the figure, there are four nests and two levels, and within level f , each nest represents one travel mode. All the nests are set under the same scenario, and the alternatives within a nest are attributed to several related service attributes. Within the first nest for the aviation mode, three alternatives are set in level h to represent the peak-hour flight (6:00–9:30; 17:00–19:30) and the normal daytime flight (9:30–17:00) as well as the evening flight (19:30–24:00). It needs to point out that j in equation (1) donates each alternative in level h in the figure.

Besides the “Air” nest, in the nest “Train,” we set two alternatives which are the “normal train” and “high-speed train.” The normal train is the traditional train of which the speed is lower than 300 km per hour, and the speed of high-speed train is over 300 km per hour. As the prices of the normal train and the high-speed train are fixed in China, we do not set multiple alternatives under the “normal train” and the “high-speed train.” Meanwhile, we set “private car” and “rented car” as the choice alternatives of the nest “Car” considering that the travel fees of the two kinds of cars are different. For the nest “Coach,” we set the “daytime coach” and “overnight coach” by taking the ticket price differences into account.

According to the hierarchical structure, substitutable relationship exists between alternatives within every nest. Hence, the error variances in the utility functions of these alternatives are not independently and identically distributed. Therefore, to capture this purchasing process, we develop a two-level nested logit model on the substitutable

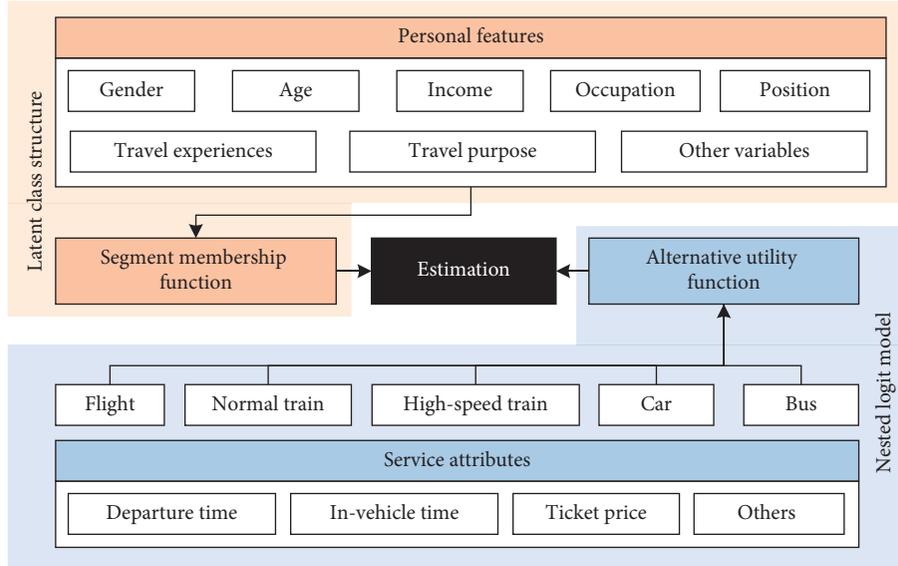


FIGURE 2: The structure of the latent-NL model.

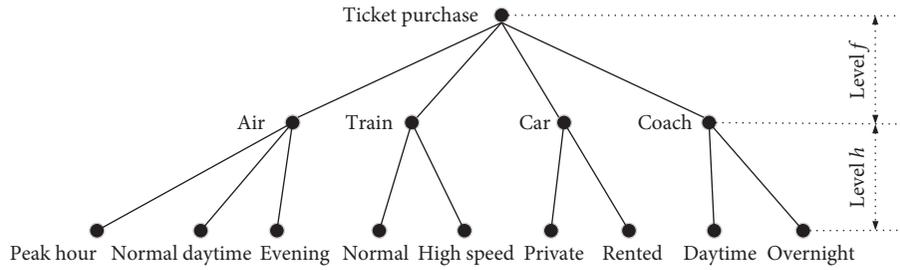


FIGURE 3: Hierarchical structure of air ticket purchase.

relationship between the alternatives. The related variables are demonstrated as follows:

h, f : levels within the nested logit hierarchical structure

$U_{j|s}$: utility of j th alternative cognized by passengers in segment s

J : the volume of alternatives

$U_{q|s}$: utility of q th nest cognized by passengers in segment s

Q : the volume of the nests in the hierarchical structure

$V_{j|s}$: observable portion of the utility function $U_{j|s}$

$V_{q|s}$: observable utility of q th nest cognized by passengers in segment s

z_j : observable variables in the utility function $U_{j|s}$

z_q : observable variables in the utility function $U_{q|s}$

β_{j_s} : parameter of observable variable z_j in the utility function $U_{j|s}$

β_{q_s} : parameter of observable variable z_q in the utility function $U_{q|s}$

α_s : constant variable of the utility function $U_{j|s}$

$\varepsilon_{j|s}$: error variance of $U_{j|s}$, assumed to be i.i.d

$\theta_{f|s}$: inverse of scale parameter, and $0 < \theta_{f|s} < 1$

Γ_{qf} : the maximum utility of alternative j within q in level f

$h_{j|s}$: one individual who belongs to segment s choose alternative j in h level

$f_{q|s}$: one individual who belongs to segment s choose alternative q in f level

On the basis of the nested logit model, the utility function of alternative j for segment s is expressed as equation (4), the utility function of nest q for segment s is as equation (5), and the probability $P(j | y_{is} = 1)$ is as equation (6). And we can see that $P(j | y_{is} = 1)$ is formulated on two related probability: one is the probability that passenger i who belongs to segment s choosing nest q in level h in Figure 3 and the other is that passenger i chooses one alternative j on condition that i chooses the nest q :

$$U_{j|s} = V_{j|s} + \varepsilon_{j|s} = \alpha_s + \beta_{j_s} z_j + \varepsilon_{j|s}, \quad (4)$$

$$U_{q|s} = V_{q|s} + \theta_{f|s} * \Gamma_{qf} = \beta_{q_s} z_q + \theta_{f|s} * \Gamma_{qf}, \quad (5)$$

$$\begin{aligned}
P(j | y_{is} = 1) &= P(h_{j|s} | f_{q|s}) \cdot P(f_{q|s}) \\
&= \frac{\exp(V_{j|s}/\theta_{f|s})}{\sum_{j=1}^J \exp(V_{j|s}/\theta_{f|s})} \cdot \frac{\exp(V_{q|s})}{\sum_{q=1}^Q \exp(V_{q|s})}. \quad (6)
\end{aligned}$$

On the basis of the nested logit model and latent class model, we can rewrite equation (1) based on equations (3) and (6) to develop the equation of the probability that individual i belonging to segment s purchases the ticket of j , which is written as follows:

$$\begin{aligned}
P_{ij|s}(x_{ij} = 1 | y_{is} = 1) &= \frac{\exp(V_{j|s}/\theta_{f|s})}{\sum_{j=1}^J \exp(V_{j|s}/\theta_{f|s})} \\
&\quad \cdot \frac{\exp(V_{q|s})}{\sum_{q=1}^Q \exp(V_{q|s})} \cdot \frac{\exp(\gamma_s m_i)}{\sum_{s=1}^S \exp(\gamma_s m_i)}. \quad (7)
\end{aligned}$$

3.3. Model Estimation. To estimate the parameters (β_{js} , β_{qs} , α_s , γ_s , $\theta_{f|s}$, $\theta_{j|s}$, and γ_s) of the proposed latent-NL model, we do the estimation on maximizing the loglikelihood value of $P_{ij|s}(x_{ij} = 1 | y_{is} = 1)$ using the expectation maximization (EM) algorithm.

It needs to point out that the volume of segment s is fixed before the estimation, which means that we should find out the most suitable segment number with several times of estimation work. Regularly, the segment number and the segment membership need to be determined by testing the model fitness measured by AIC (Akaike Information Criterion) and BIC (Bayesian information criterion), which are written as equations (8) and (9):

$$AIC = -2LL + 2K, \quad (8)$$

$$BIC = -2LL + (\ln(N))K. \quad (9)$$

In the equations, LL is the maximum value of the loglikelihood in the estimation, K is the number of parameters in the proposed model, and N is the sample size. Except for model fitness, these two indicators also consider the number of parameters to prevent overfitting of the model [40]. As BIC places more penalties on the number of parameters than AIC, BIC is recognized as a better criterion to measure the fitness of the latent class model than AIC and loglikelihood.

4. Survey Design

As the statistical trip data may not contain passengers' demographic characteristics which are important for market segmentation, here we should use the survey to do the data collection. Generally, there are two main types of survey design: the revealed preference (RP) survey and the stated preference (SP) survey. The RP survey does the interviews based on the respondents' travel experiences but may not be

able to lead them to memorize the detailed choice situations in the past, so it is hard to use the RP data to predict passengers' potential choice decisions in other situations.

In the contrary, the SP survey does not interview the past trips, but it provides experiments for collecting choice decisions in different scenarios. The experiments are selected out from the profiles which are generated by combining the travel scenario and the choice alternative as well as the service attribute. Commonly, the profiles are designed to reveal the real situation but they may be not existed in reality, so the SP survey data would lead to overestimation [41–43]. Therefore, to update the SP survey, we choose to create the virtual travel scenario for each profile, but to set the service attributes using the real information of flight, train, car, and coach, thus to make sure that the respondents could give the choice results as close as to reality.

4.1. Survey Design. The survey comprises of two parts. The first part collects the data about personal features of respondents, including income, occupation, age, and air travel experience, such attributes are the variables in the segment membership function in equation (2). The second part of the survey gives out three experiments in which the trip scenarios and choice alternatives are listed. In the experiment, the alternatives are described by several attributes which are the variables of utility function in equation (4).

Table 1 gives an overview of the selected variables related to the personal features, which are the variables of the segment membership function, the values of the variables are set according to the suggestions from the airline company. It needs to explain that the purchase dates for business and leisure trips are the duration between the ticket purchase date and the departure date.

Besides, Table 2 demonstrates the environmental variables of the utility function describing the trip scenario for all nests. In the survey, we choose four origin cities to deliver the questionnaires, Dalian, Shanghai, Nanjing, and Chengdu, and choose 50 cities in China as the trip destinations. The respondents would receive the questionnaires in which the trip origin is the city that he/she lives in. Meanwhile, when the origin and destination are confirmed, the travel distance would be calculated and the travel purpose would be set for the survey.

Then, in Table 3, the variables describing the choice alternatives are listed. In the table, 5 to 8 are the generic variables [44] for all alternatives, and then 9 to 15 are specific variables for different nests. The values of the variables are set using the real information collected from websites and the operating data provided by airports. In this sense, there is no need to set the levels of attributes in the SP survey design, and the real travel information would help the respondents make the choice as close as to the reality.

Totally, 213 profiles generated, 43 profiles for Dalian, 79 profiles for Shanghai, 59 profiles for Nanjing, and the rest for Chengdu; the profiles are partly shown in Figure 4. In the figure, each line represents a choice scenario with different alternatives attributed to specific variables using real service information. Based on the generated profiles, each

TABLE 1: Selected attributes of personal feature.

No.	Variables	Values
1	Age	≤ 22 ; $>22 \& \leq 30$; $>30 \& \leq 40$; $>40 \& \leq 50$; >50
2	Occupation	None; company worker; merchant; civil servant; farmer; others
3	Position	Staff; first-line; middle-level; top-level
4	Income, unit: 10,000RMB per year	≤ 5 ; $>5 \& \leq 10$; $>10 \& \leq 20$; $>20 \& \leq 40$; >40
5	Frequent travel purposes	Business; family; study; leisure; others
6	Long-distance frequency per year	≤ 1 ; $>1 \& \leq 3$; $>3 \& \leq 5$; $>5 \& \leq 8$; >8
7	Purchase date for business	≤ 3 ; $>3 \& \leq 7$; $>7 \& \leq 14$; $>14 \& \leq 30$; >30
8	Purchase date for leisure	≤ 7 ; $>7 \& \leq 14$; $>15 \& \leq 30$; $>30 \& \leq 60$; >60

TABLE 2: Selected environmental variables for all nests.

No.	Variables	Values
1	Origin	Dalian; Shanghai; Nanjing; Chengdu
2	Destination	Beijing, Ha'erbin, Hefei, Urumqi, etc. (50 cities)
3	Travel purpose	Business; family; study; leisure; others
4	Travel distance, unit: km	≤ 800 ; $>800 \& \leq 1200$; $>1200 \& \leq 1600$; >1600

TABLE 3: Selected variables for the alternatives.

No.	Variables
5	Departure time
6	Travel time, unit: hour
7	Delay time, unit: hour
8	Access time, unit: hour
9	Air fare, unit: 100RMB
10	Aircraft type
11	Normal train fare, unit: 100RMB
12	High-speed train fare, unit: 100RMB
13	Rented car fare, unit: 100RMB
14	Private car fare, unit: 100RMB
15	Coach fare, unit: 100RMB

questionnaire will be set with three experiments which are randomly selected from the profile set.

4.2. Data Collection and Descriptive Analysis. Then, the questionnaires were distributed to respondents face to face in four cities where four airports are located in; they are Dalian, Shanghai, Nanjing and Chengdu, from October 25th to 31th in 2019. As the survey duration does not cover the holiday as well as the starting or ending of school season, we would confirm that this period is valid for conducting the survey without the special service characteristics and abnormal choice decisions. The scale of the four airports are different, Shanghai Pudong airport is one of the largest airports in China. Dalian, Nanjing, and Chengdu airports are the regional hub airports in different areas. Meanwhile, the four cities are well reachable by other travel modes (train, car, and coach), making the selected cities the suitable cases for validating our proposed model.

In the end, 153, 122, 169, and 136 valid questionnaires were collected in Shanghai Pudong, Dalian, Nanjing, and Chengdu airport, respectively, collecting 1740 sets of choice data, and the response rate was 87.3%. A preliminary descriptive analysis is conducted. It shows that 26.1% of the respondents are young travelers (≤ 30), 48.3% of them are

middle-age person ($>30 \& \leq 55$), and others are old people (>55). Besides, 36.2% of the respondents only travel by air for once a year, 12.1% of them have over 8 times of air travel per year, and others' air travel frequency drops in the range between 1 to 8 times per year.

On the basis of the survey data and the proposed latent-NL model, we could conduct the estimation work. Section 5 introduces the estimation results.

5. Estimation Results

Before estimating the latent-NL model, we firstly try to estimate the parameters of the variables in the nested logit structure and set the goodness-of-fit of the NL model as the benchmark in our research.

5.1. Results of Nested Logit Model. Based on the nested logit structure, the parameters are estimated using the maximum likelihood estimation method. The alternative "Coach 2" is chosen as the referent alternative, and the associated constant is zero. Results are shown in Table 4. The goodness-of-fit is satisfactory, with the pseudo R^2 equal to 0.35. The inclusive value parameter is 0.54 for "Air" nest, 0.43 for "Train" nest, 0.48 for "Car" nest, and 0.61 for "Coach" nest. All the inclusive value parameters fall within 0 and 1, which supports the nested logit structure of the air ticket purchasing model.

In the table, the variables written in italic type are the constants of the utility functions. By comparing the constants of different alternatives, we could sort the travel modes in a preference-descending sequence for long-distance trips: air, high-speed train, normal train, car, and coach. Then, for the variables related to personal features, the coefficients of age, position, and long-distance travel frequency are statistically significant, showing that the three variables may have strong influences on the market segmentation. Meanwhile, the negative sign of long-distance travel frequency states that passengers with frequent travel

No.	O	D	Dis	Pur	Flight1D	Flhgt1T	Flight1F	Flight2D	Flhgt2T	Flight2F	Flight3D	Flhgt3T	Flight3F	TrainND	TrainNT	TrainNF	TrainHD	TrainHT	TrainHF	CarPT	CarPF	CarRT	CarRF
1	4	1	4	3	9	2.8	11.5	14	2.8	7.8	20	2.8	19.3	8.6	27.8	3.99	10.7	10	7.78	22	24.7	23	26.7
2	4	1	4	5	7	2.8	12	13	2.8	17.2	22.5	2.8	13.5	11.7	22.3	4.34	15	7.9	7.78	22	24.7	23	26.7
3	4	1	4	1	18	2.8	14.8	15	2.8	17.6	21	2.8	15.4	19.1	16	4.37	7	7.8	7.78	22	24.7	23	26.7
4	4	2	4	3	6.1	4	5.6	15	5.5	13.46	21.5	4.5	10	7.7	50	5.76	15	11.1	12.8	36.4	31	37.4	33
5	4	2	4	2	9.5	5.5	8.9	15.2	4	21.7	21.5	4.5	10	7.7	50	5.76	12.4	11.7	12.8	36.4	31	37.4	33
6	4	3	3	2	8.5	2	10.1	16.9	2	10.1	21	2	10.07	8.5	20.9	3.63	8.4	11.4	4.99	17.5	16	18.5	18
7	4	4	4	3	6.3	3.8	8.6	13.5	3.8	15.8	21.8	3.6	13.7	20.2	35.1	4.91	12	32.3	7.85	33	29	34	31
8	4	4	4	1	8.1	3.8	15.8	13.7	5.1	10.1	19.9	3.6	15.8	20.2	35.1	4.91	12	32.3	7.85	33	29	34	31
9	4	4	4	4	17	3.6	18.5	15.3	3.8	13.7	22.1	3.6	15.8	20.2	35.1	4.91	12	32.3	7.85	33	29	34	31
10	4	5	2	5	6.8	14.2	4.16	14.5	1.35	5.24	21	1.5	4.9	12.4	8.5	15.8	9.6	4	4.37	7.5	7.5	8.5	9.5
11	4	6	2	3	6.6	1.6	13.8	13.1	1.5	14.15	22.1	1.5	4.7	0	0	0	14.1	6.8	4.87	10.5	8.7	11.5	10.7
12	4	6	2	3	17.2	1.6	15.2	13.9	1.5	12.9	21.3	1.5	6.87	0	0	0	15.1	6.6	4.87	10.5	8.7	11.5	10.7
13	4	6	2	2	17.5	1.5	12.9	15.5	1.6	15.1	23.6	1.7	4.2	0	0	0	8.9	5.6	4.87	10.5	8.7	11.5	10.7
14	4	6	2	1	9.1	6.7	9.62	10.8	1.5	11.1	23.1	1.6	4.7	0	0	0	9.7	6.3	4.87	10.5	8.7	11.5	10.7
15	4	7	3	2	8	2.3	6.5	15.1	2.3	8.9	20.2	2.3	6	9.5	27.5	3.53	7.1	8.7	7.24	18	15.7	19	17.7
16	4	7	3	5	9.1	2.3	10.8	14.2	2.3	7.6	21.6	2.3	8.03	22.6	22.5	3.77	8.2	9.1	7.24	18	15.7	19	17.7
17	4	8	4	4	7.3	2.4	7.64	12.7	2.5	12.9	19.9	2.4	10.1	9.1	27.7	3.81	8	9.9	5.41	18.5	15.2	19.5	17.2
18	4	8	4	5	17.3	2.4	10.1	14.5	2.4	17.2	21.4	2.4	8.8	15.2	25	4.8	14.2	9.7	5.66	18.5	15.2	19.5	17.2
19	4	8	4	1	8.5	2.4	9	11.1	2.4	11.2	22.5	2.4	13.1	17.9	34.7	4.08	6.7	9.5	5.66	18.5	15.2	19.5	17.2
20	4	8	4	2	9.2	2.4	10.1	12	2.4	10.1	21.1	2.4	9.2	12.5	29	4.14	13.6	9.7	5.74	18.5	15.2	19.5	17.2
21	4	8	4	3	9.5	2.4	10.9	16.8	2.5	10.1	22.5	2.4	11.2	15.2	25	4.8	9.4	9.9	5.74	18.5	15.2	19.5	17.2
22	4	10	1	3	6.1	1.5	7.85	13.7	1.5	4.4	21.7	1.5	3.7	8.1	13.8	1.94	10.1	4.4	2.63	9.2	7	11.2	9
23	4	11	3	1	7.1	2	9.4	11.3	2	8.9	20.7	2	8.4	13.9	25.5	3.28	11.5	7.1	5.02	14.5	11.87	15.5	13.87
24	4	11	3	5	9.3	2	9	14.7	2	7.8	21.8	2	7.6	11.6	16.1	3.37	15	5.9	5.02	14.5	11.87	15.5	13.87

FIGURE 4: The sample of generated profiles.

TABLE 4: Results of the nested logit model.

Variable	Coefficient
Peak-hour flight	3.162 (0.000***)
Normal day-time flight	3.245 (0.000***)
Evening flight	3.076 (0.065**)
Normal train	1.890 (0.158*)
High-speed train	2.764 (0.076**)
Private car	0.561 (0.111*)
Rented car	0.874 (0.107*)
Daytime coach	0.145 (0.000***)
Age	-0.231 (0.000***)
Occupation	0.086 (0.354)
Position	0.149 (0.027**)
Income	0.017 (0.216)
Frequent travel purpose	0.159*** (0.000)
Long-distance frequency	-0.356 (0.000***)
Purchase date for business	-0.002 (0.218)
Purchase date for leisure	-0.016 (0.219)
Origin	0.001 (3.654)
Destination	-0.001 (1.769)
Travel purpose	0.032 (0.000***)
Travel distance	-0.005 (0.119*)
Departure time	-0.184 (0.379)
Travel time	-1.017 (0.062*)
Delay time	-0.514 (0.000***)
Access time	-0.043 (0.650)
Air fare	-1.015 (0.000***)
Normal train fare	-0.547 (0.095*)
High-speed train fare	-0.972 (0.000***)
Rented car fare	-0.867 (0.821)
Private car fare	-0.708 (0.104*)
Coach fare	-0.559 (2.732)
Loglikelihood function:	-2596.37
Pseudo R^2 (ρ^2):	0.35

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level.

experiences may have higher requirements on the service quality of the travel alternatives.

For the coefficients of the environmental variables, the travel purpose is the significant one in the estimation; its positive sign indicates that passengers would obtain gains from the trip. Meanwhile, the travel distance is significant

but seems that it does not have strong impacts on the alternative choice, which may be due to the development of the high-speed travel mode. In addition, it can be easily found that passengers put more concerns on the money costs than the travel time according to the coefficient value and the significant status. Besides, comparing the coefficient of air fare and those of other alternatives, the same rise in ticket price would lead to larger decrease in the air travel utility, showing that passengers have much more critical concern on the air fare increases.

Another significant coefficient is for the delay time, its value shows higher negative impacts of the delay time on the ticket purchase than the access time and even the fare of some travel modes. Moreover, the insignificance of the coefficients for rented car fare as well as coach fare may be due to the reason that the two travel modes are not popular for long-distance trip, especially when the travel distance is longer than 800 km.

5.2. Results of Latent-NL Model. To determine the model specification with the best fit, we tested a variety of models by varying the segment number. Table 5 summarizes the goodness-of-fit of two-, three-, and four-segmented model, based on which we could determine the model specification by comparing the AIC and BIC value, and the values are calculated using the maximum loglikelihood value (LL) based on equations (8) and (9). It can be seen that the lowest BIC value appears when the segment number equals to three; hence, the three-segment latent-NL model is the best choice of our research statistically.

The estimation results of the three-segment latent-NL model are shown in Table 6, and the first segment (S1) is set as the base of segment two and three (S2 and S3) in the estimation. In the following paragraphs, we will introduce in great detail of the segment-based characteristics of market, and further compare the results across different segments with the purpose of testing segment-specific attitude in air ticket purchase.

In the table, the first column demonstrates the label of the variables, the second column shows the coefficients of the NL model in Table 4, and the rest columns list the coefficients

TABLE 5: Estimation efficiency with different segment numbers.

	Parameters	LL	AIC	BIC
Logit model	24	-2596.37	5260.74	5364.34
Two-segment	34	-2436.12	4940.24	5127.24
Three-segment	44	-2327.81	4743.62	4985.62
Four-segment	54	-2293.98	4695.96	4992.96

estimated for different segments (S1 to S3) within the latent-NL model. From the results, the respondents in our research are classified into three groups which comprise 41.04%, 37.89%, and 21.07% of the population, respectively. The pseudo R^2 for each segment is listed in the last line of the table, indicating that the latent-NL model can explain the air ticket purchase behavior more precisely than the NL model.

The coefficients of latent variables are listed in the fourth and fifth column; S1 does not get such coefficients because it is the base in the estimation. Then, taking all the coefficients into joint consideration, we could title each segment according to the segment-specific features as follows.

5.2.1. S1: Low-Fare Seekers. The distinctive characteristic of S1 is that the memberships have highest negative attitudes to the travel fare but lowest negative attitudes to the travel time; hence, the passengers within the group are defined as the low-fare seekers.

As per the personal features, 79.6% of the group members lie in the age thresholds of ≤ 22 and > 55 , 36.8% of them do not have any job or retired, 87.6% of the employed members are first-line workers, and the average salary may not exceed 100 thousand RMB per year. Besides, the long-distance travel frequency is less than 3 times per year and the frequent travel purpose is concentrated in leisure and family. In addition, 92.5% of the members purchase the tickets for business and leisure trips at the dates about 30 and 60 days, respectively, in advance of the departure, indicating that they may be the customers enjoying the early bird prices.

Besides the feature of low-fare seeking, members in S1 also have other segment-specific characteristics in the ticket purchase. For example, the constants of the utility functions state that they prefer using air and train in long-distance journey but do not show the favorite mode; moreover, the travel purpose seems to have weaker impact on the purchase decision than that it has in other segments, demonstrating that passengers in this group would choose the cheapest alternative regardless of the travel purpose; furthermore, the group members may not concern a lot about the service quality such as delay and departure time, that is because the low cost and high service quality could not exist simultaneously.

In general, members in S1 are attributed to low-fare seek and early purchase; thus, they might be treated as the target passengers of airlines at the start-up time of air ticket booking horizon.

5.2.2. S2: Experienced Waverers. Compared to S1, members in S2 show lower negative attitude to the travel fare but stronger negative attitude to the travel time, so they inherently prefer air and high-speed train in the choice.

TABLE 6: Estimation results of the latent-NL model.

	NL	S1	S2	S3
<i>Constants</i>				
Peak-hour flight	3.162***	2.134*	3.072***	4.671***
Normal day-time flight	3.245***	2.093	3.018*	2.387***
Evening flight	3.076**	2.110*	3.159*	3.762**
Normal train	1.890*	2.272*	0.995**	0.265**
High-speed train	2.764**	1.958*	3.012**	2.316*
Private car	0.561**	0.006	0.821	0.043**
Rented car	0.874*	0.019*	0.439	0.019*
Daytime coach	0.145	0.189	0.062	0.004
<i>Environmental variables</i>				
Origin	0.001	0.000	0.007	0.012*
Destination	-0.001	0.000	-0.015*	-0.016**
Travel purpose	0.032***	0.012***	0.019***	0.033***
Travel distance	-0.005*	-0.004*	-0.006	-0.002
<i>Generic variables</i>				
Departure time	-0.184	-0.035	-0.232*	-0.711***
Travel time	-1.017*	-0.569**	-0.876*	-1.284**
Delay time	-0.514***	-0.082	-0.414***	-1.617**
Access time	-0.043	-0.024	-0.089	-0.286*
<i>Specific variables</i>				
Air fare	-1.015***	-1.558***	-0.992**	-0.538***
Normal train fare	-0.547*	-0.627***	-0.413**	-0.294*
High-speed train fare	-0.972***	-1.059*	-0.815***	-0.419**
Rented car fare	-0.867	-0.966	-0.812	-0.421
Private car fare	-0.708*	-0.893	-0.782*	-0.305
Coach fare	-0.559	-0.789*	-0.497	-0.207
<i>Personal feature variables</i>				
Age	-0.231***		-0.114**	-0.275*
Occupation	0.086		0.076	0.085**
Position	0.149**		0.186*	0.214**
Income	0.017		0.007	0.013*
Frequent travel purpose	0.159***		0.213***	0.165***
Long distance frequency	-0.356***		-0.218*	-0.492***
Purchase date for business	-0.002		-0.035**	-0.006
Purchase date for leisure	-0.016		-0.089**	-0.027
%	—	41.04	37.89	21.07
Pseudo R^2	0.35	0.37	0.45	0.41

***Significant at the 1% level. **Significant at the 5% level. *Significant at the 10% level.

However, seen from the constants of utility functions, the group members do not show obvious preferences to air or high-speed train, declaring that their purchase decisions depend on, to large extents, the service attributes such as ticket price, departure time, and delay time. In the sense, they may easily switch between air and high-speed train with the variations of service attributes, so members in S2 can be called waverers in the market.

Besides, when we check the social-economical features of the group, 82.4% members drop in the age range between 30

and 40, 86.1% of them are employed, and 45% are the middle-level managers with their salary over 100 thousand per year. Because of the high employment rate, 45.7% of the members commonly travel for business and commonly buy the tickets 7–15 days in advance of the trip. More importantly, their average travel frequency (5–8 times per year) is much higher than that in S1; thus, we can define the group passengers in S2 as experienced waverers.

Because of the cumulated experiences, passengers in S2 may be familiar with the flights as well as the high-speed trains to various destinations, so the coefficient for the destination is significant. In the circumstance, the members would have the ability to quickly compare the service attributes of the two modes in their mind without asking for lots of information. Apart from the prices, other service attributes are also highlighted by the members in S2; for instance, they have strong negative attitude to the flight delay and also care about the departure time.

As summarized above, the experienced waverers have higher preferences to air travel than the ones in S1, but they are also interested in using high-speed train. Meanwhile, they concern a lot about the service quality of the travel modes, so the way of airlines to capture them is to optimize the price strategy and improve the service quality at the same time.

5.2.3. S3: Loyal Flyers. Comparing to S2, members in S3 show highest inherent preferences to the air travel, indicating that they are the loyal passengers of airlines. The formation of loyalty is due to several reasons; first is the high income. To be specific, 42.5% of the group members are leadership managers and 52.7% of which could earn the income over 400 thousand RMB per year, so they could afford the air ticket prices even if the prices are very high just before the flight departure. The second reason is due to the time value. As declared by the coefficient of travel time, it is easy to find that the group members have the strongest negative attitude to the time costs, which is because 62.6% of them need to travel for more than 8 times every year with business purpose; thus, the aviation which can provide quick and comfortable connections becomes the first choice.

Due to the loyalty, members in S3 would not easily shift to other modes in reality. However, they also have higher requirements on the airlines' service quality, for example, they have lowest tolerance of flight delay and prefer early morning and evening flights for the convenience of day-time work. Besides, they are strict on the access time as well as the comfort during the trip as 23% of the group members are with the age over 55. Therefore, members in S3 are the airlines' loyal passengers, and they would not easily switch to other modes for high prices but would leave for weak services.

On the basis of previous analysis, it is clear that airlines' markets are statistically composed of three segments of passengers which could be defined as low-fare seekers, experienced waverers, and loyal flyers. According to the segment-specific ticket purchase behavior, the low-fare seekers pursue cost saving; the experienced waverers who

highlight both of the fare and the service quality may easily switch between the air and the high-speed train; the loyal flyers who habitually prefer the flights would put more emphasizes on the service quality than on the prices. In the sense, the air ticket pricing strategies would be set according to the different price attitudes kept by the segment members. In Section 6, we would do the sensitivity analysis for testing the segment-specific reaction to the change of ticket prices.

6. Discussion

The sensitivity analysis is conducted in three situations: increasing the price of peak-hour flight by 10%, decreasing the price of normal day-time flight by 10%, and decreasing the price of evening flight by 10%. Under each situation, we will calculate the changes in the purchase probability to the changes in the travel fare for each market segment. Table 7 illustrates the calculation results. We need to note that, when changing the travel fare of a specific alternative, the changes in the choice probability of other alternatives should be the same. This phenomenon is due to the independence from irrelevant alternative (IIA) property of discrete choice models.

In the table, columns 3 to 5 demonstrate the segment-based changes of the choice probabilities, and the probability can be also recognized as the market share of different alternatives. In situation 1, the peak-hour flight may totally lose 1.59% market share with 10% increment of its air ticket price. In situations 2 and 3, the normal day-time flight as well as the evening flight would win 10.96% and 11.38% market share, respectively, for cheaper air fare. When we go through the market share variations in each segment, it is easy to find out that the largest win and loss exists in S2, proving that the members in S2 are the experienced waverers. Meanwhile, passengers in S1 also do large contributions in improving the market share of the normal day-time flight and the evening flight, which is because they are low-fare seekers. However, because of the price insensitive, only a few members within S3 change their choices. Therefore, the pricing strategies should be proposed on the segment-specific price sensitivity.

Apart from the segment-specific attitude to the travel fare, the segment membership volume would also be essential for air ticket pricing. In order to do further interpretation, we create an experiment to compare the flight revenues resulted from different pricing strategies. The information of all the alternatives in the experiment is listed in Table 8.

Based on the abovementioned service attributes, we fix the ticket prices of all alternatives except for that of Flight 2 and then use the surveyed personal features and the estimated latent-NL model to calculate the revenues of Flight 2 with different pricing strategies. Table 9 declares the results.

For the strategies in Table 9, each of them contain four classes attributed to different air ticket prices and seat numbers; we call this kind of strategies as class allocations for airline marketing. Comparing the sold seats and revenues in the table, Strategy 2 would be better than Strategy 1 for the higher average revenue per seat. The success of Strategy 2 is due to the reason that it captures the passengers who would

TABLE 7: Demand sensitivity to ticket price in different segments.

	Alternatives	S1 %	S2 %	S3 %	Total %
Situation 1	Peak-hour flight	-0.0	-1.04	-0.53	-1.59
Situation 2	Normal day-time flight	3.85	5.27	1.84	10.96
Situation 3	Evening flight	2.75	5.48	1.15	11.38

TABLE 8: Information about choice alternatives.

Alternatives	Departure time	Travel time	Delay time	Access time
Flight 1 (120 seats)	7:00	1.5 h	0.25 h	1.5 h
Flight 2 (120 seats)	14:00	1.5 h	0.1 h	1.5 h
Flight 3 (120 seats)	20:00	1.5 h	0.25 h	1.5 h
Normal train	18:00	9 h	0	0.5 h
High-speed train	14:00	4.5 h	0	1 h
Private car	0	7 h	0.5 h	0
Rented car	10:00	8 h	0.5 h	0.25 h
Day-time coach	8:30	9 h	0.5 h	0.5 h
Overnight coach	20:00	8 h	0	0.5 h

TABLE 9: Flight 2's revenues with different pricing strategies.

Strategy	30-60 days	15-30 days	7-15 days	0-7 days	Revenue
1	480 (50 seats)	680 (30 seats)	980 (20 seats)	1680 (20 seats)	80900
Sold tickets	50	30	20	10	
2	480 (20 seats)	680 (60 seats)	980 (30 seats)	1680 (10 seats)	86940
Sold tickets	20	53	25	10	

pay higher prices for the same seats. In other words, the class allocation set by Strategy 2 is proposed on the right distribution of passengers with different segment-specific price-sensitivity. Therefore, the market segmentation would have great effects on improving airlines' profits with its usage in air ticket pricing, or more precisely, in the class allocation.

7. Conclusion

Regarding the fact that passengers' air ticket purchase behavior is relevant to the heterogeneous preferences, this paper established an air ticket purchase model incorporating the market segmentation. In the model construction, the discrete choice model which had advantages in predicting purchase preferences was integrated into the latent class structure which could classify the market segments, thus to established a latent-NL model for predicting the segment-specific air ticket purchase behavior.

The proposed model was estimated using the SP survey data, and three groups of passengers were identified, titled with low-fare seekers, and experienced waverers as well as loyal flyers. With the comparison to the common classifications such as the "business-and-leisure" used in practice, our proposed approach could distinguish the passengers according to their purchase predispositions derived from the personal features and the purchase preferences. Because of the precise segmentation, the estimated segment-specific purchase models had been proved with advantages over the nested logit model for better fitness, indicating that the latent-NL has advantages in predicting the air ticket

purchase behavior. Therefore, by constructing the latent-NL model with the careful variable selection as well as the survey design, we found an effective way to predict air ticket purchase behavior especially for the job of air ticket pricing.

Then, on the basis of the sensitivity analysis, we made clear that passengers' feedbacks to the price fluctuations would vary across the segments. Hence, the same percentage change in the air ticket price would lead to the diverse variations in the demand of different flights which were preferred by passengers with differentiated price sensitivities. In this sense, the airlines could use the proposed model to precisely evaluate the effects of varying the air ticket price in improving the flight market share. Furthermore, the estimated latent-NL model was adopted to compare the revenues of one flight under different pricing strategies. The results proved that the market segmentation would support the airlines to predict the demand volume for each price class, thus to capture the high-value passengers. Therefore, our second contribution was that we proved the importance of market segmentation in the air ticket pricing and gave out the benefits of adopting the latent-NL model in raising the flight revenue.

However, there are some limitations existed in our work. First, the SP survey was conducted within only one week, which might lead to the difficulty in covering all kinds of ticket purchase behavior. Although 4 cities were selected to deliver the questionnaires with the purpose of offsetting the limitation led by the short survey duration, we would try to collect much more survey data to verify our results in the future. Second, apart from the demographical features, the

experiences, and the service attributes, some other factors such as the passenger location, the airport scale, and the seasonal factor may also affect the airlines' market segmentation in air ticket pricing, and all these factors should be incorporated in the future work.

Data Availability

The survey data used to support the findings of this study have not been made available because the data will be used to support another ongoing research, and meanwhile the data only can be used when permitted by the respondents as there are some private information incorporated in the data.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This research was sponsored by the National Key Research and Development Project (Grant no. 2018YFB1601200) and National Natural Science Foundation of China (Grant nos. 71701094 and 51608268).

References

- [1] G. Gallego and G. van Ryzin, "A multiproduct dynamic pricing problem and its applications to network yield management," *Operations Research*, vol. 45, no. 1, pp. 24–41, 1997.
- [2] A. W. Donovan, "Yield management in the airline industry," *The Journal of Aviation/Aerospace Education & Research*, vol. 14, no. 3, pp. 11–20, 2005.
- [3] K. Kulkarni, A. Gosavi, S. Murray, and K. Grantham, "Semi-markov adaptive critic heuristics with application to airline revenue management," *Journal of Control Theory and Applications*, vol. 9, no. 3, pp. 421–430, 2011.
- [4] E. A. Grantham and R. Kallesen, "Practice papers: the science of revenue management when passengers purchase the lowest available fare," *Journal of Revenue and Pricing Management*, vol. 3, no. 2, pp. 171–177, 2004.
- [5] Y.-C. Chiou and C.-H. Liu, "Advance purchase behaviors of air tickets," *Journal of Air Transport Management*, vol. 57, pp. 62–69, 2016.
- [6] T. A. Arentze and E. J. E. Molin, "Travelers' preferences in multimodal networks: design and results of a comprehensive series of choice experiments," *Transportation Research Part A: Policy and Practice*, vol. 58, pp. 15–28, 2013.
- [7] T.-H. Tsai, "Homogeneous service with heterogeneous products: relationships among airline ticket fares and purchase fences," *Journal of Air Transport Management*, vol. 55, pp. 164–175, 2016.
- [8] I. W.-Y. Chiang, G.-S. Liang, and S. Yahalom, "The fuzzy clustering method: applications in the air transport market in Taiwan," *Journal of Database Marketing & Customer Strategy Management*, vol. 11, no. 2, pp. 149–158, 2003.
- [9] Y.-C. Chiou and C.-H. Liu, "Advance purchase behaviors of air passengers: a continuous logit model," *Transportation Research Part E: Logistics and Transportation Review*, vol. 93, pp. 474–484, 2016.
- [10] J. Ishii, S. Jun, and K. Van Dender, "Air travel choices in multi-airport markets," *Journal of Urban Economics*, vol. 65, no. 2, pp. 216–227, 2009.
- [11] C.-W. Yang, J.-L. Lu, and C.-Y. Hsu, "Modeling joint airport and route choice behavior for international and metropolitan airports," *Journal of Air Transport Management*, vol. 39, pp. 89–95, 2014.
- [12] W. R. Smith, "Product differentiation and market segmentation as alternative marketing strategies," *Journal of Marketing*, vol. 21, no. 1, pp. 3–8, 1956.
- [13] B. P. Y. Loo, "Passengers' airport choice within multi-airport regions (MARs): some insights from a stated preference survey at Hong Kong International Airport," *Journal of Transport Geography*, vol. 16, no. 2, pp. 117–125, 2008.
- [14] M. Brons, E. Pels, P. Nijkamp, and P. Rietveld, "Price elasticities of demand for passenger air travel: a meta-analysis," *Journal of Air Transport Management*, vol. 8, no. 3, pp. 165–175, 2002.
- [15] T. Rietveld, E. Shehu, and I. von Wartburg, "Customer segmentation revisited: the case of the airline industry," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 1, pp. 227–242, 2008.
- [16] W. H. Greene and D. A. Hensher, "A latent class model for discrete choice analysis: contrasts with mixed logit," *Transportation Research Part B: Methodological*, vol. 37, no. 8, pp. 681–698, 2003.
- [17] J. Zhang, M. Kuwano, B. Lee, and A. Fujiwara, "Modeling household discrete choice behavior incorporating heterogeneous group decision-making mechanisms," *Transportation Research Part B: Methodological*, vol. 43, no. 2, pp. 230–250, 2009.
- [18] H. Shao, W. H. K. Lam, and M. L. Tam, "A reliability-based stochastic traffic assignment model for network with multiple user classes under uncertainty in demand," *Networks and Spatial Economics*, vol. 6, no. 3–4, pp. 173–204, 2006.
- [19] T. Tang, Y. Shi, Y. Wang, and G. Yu, "A bus-following model with an on-line bus station," *Nonlinear Dynamics*, vol. 70, no. 1, pp. 209–215, 2012.
- [20] L. Kattan, K. M. N. Habib, S. Nadeem, and T. Islam, "Modeling travelers' responses to incident information provided by variable message signs in Calgary, Canada," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2185, no. 1, pp. 71–80, 2010.
- [21] M. Islam, M. Mohamed, C. D. Higgins, E. Abotalebi, and P. Abotalebi, "How open are Canadian households to electric vehicles? a national latent class choice analysis with willingness-to-pay and metropolitan characterization," *Transportation Research Part D: Transport and Environment*, vol. 58, pp. 208–224, 2018.
- [22] W. A. Kamakura and G. J. Russell, "A probabilistic choice model for market segmentation and elasticity structure," *Journal of Marketing Research*, vol. 26, no. 4, pp. 379–390, 1989.
- [23] L. Wu, J. Zhang, and A. Fujiwara, "Representing tourists' heterogeneous choices of destination and travel party with an integrated latent class and nested logit model," *Tourism Management*, vol. 32, no. 6, pp. 1407–1413, 2011.
- [24] C.-H. Wen, W.-C. Wang, and C. Fu, "Latent class nested logit model for analyzing high-speed rail access mode choice," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 2, pp. 545–554, 2012.
- [25] R. E. Skinner, "Airport choice: an empirical study," *Journal of Transportation Engineering*, vol. 102, no. 4, pp. 871–883, 1976.

- [26] H. Yang and A. Zhang, "Effects of high-speed rail and air transport competition on prices, profits and welfare," *Transportation Research Part B: Methodological*, vol. 46, no. 10, pp. 1322–1333, 2012.
- [27] S. Hess, "Treatment of reference alternatives in stated choice surveys for air travel choice behaviour," *Journal of Air Transport Management*, vol. 14, no. 5, pp. 275–279, 2008.
- [28] Y.-C. Chiou and Y.-H. Chen, "Service quality effects on air passenger intentions: a service chain perspective," *Transportmetrica*, vol. 8, no. 6, pp. 406–426, 2012.
- [29] G. Debrezion, E. Pels, and P. Rietveld, "Modelling the joint access mode and railway station choice," *Transportation Research Part E: Logistics and Transportation Review*, vol. 45, no. 1, pp. 270–283, 2009.
- [30] K. Arunotayanun and J. W. Polak, "Taste heterogeneity and market segmentation in freight shippers' mode choice behaviour," *Transportation Research Part E Logistics & Transportation Review*, vol. 47, no. 2, pp. 138–148, 2011.
- [31] A. Vij and J. L. Walker, "Preference endogeneity in discrete choice models," *Transportation Research Part B: Methodological*, vol. 64, pp. 90–105, 2014.
- [32] F. El Zarwi, A. Vij, and J. L. Walker, "A discrete choice framework for modeling and forecasting the adoption and diffusion of new transportation services," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 207–223, 2017.
- [33] C. H. Wen and S. C. Lai, "Latent class models of international air carrier choice," *Transportation Research Part E Logistics & Transportation Review*, vol. 46, no. 2, pp. 211–221, 2010.
- [34] D. Li, B. Fu, Y. Wang et al., "Percolation transition in dynamical traffic network with evolving critical bottlenecks," *Proceedings of the National Academy of Sciences*, vol. 112, no. 3, pp. 669–672, 2015.
- [35] R. Chen and Z. Lei, "Airport dominance and airline pricing power: an investigation of hub premiums in the Chinese domestic market," *Transportation Research Part A: Policy and Practice*, vol. 103, pp. 509–524, 2017.
- [36] T. M. Lei, "Airfare pricing determinants in hub-to-hub markets," *Journal of Transport Geography*, vol. 14, no. 1, pp. 15–22, 2006.
- [37] Q. Zhang, H. Yang, Q. Wang, and A. Zhang, "Market power and its determinants in the Chinese airline industry," *Transportation Research Part A: Policy and Practice*, vol. 64, pp. 1–13, 2014.
- [38] M. L. Zhang, W. H. K. Lam, and H. P. Lo, "Incorporating passenger perceived service quality in airport ground access mode choice model," *Transportmetrica*, vol. 6, no. 1, pp. 3–17, 2010.
- [39] E. Pels, P. Nijkamp, and P. Rietveld, "Airport and airline choice in a multiple airport region: an empirical analysis for the San Francisco Bay area," *Regional Studies*, vol. 35, no. 1, pp. 1–9, 2001.
- [40] J. L. Walker and J. Li, "Latent lifestyle preferences and household location decisions," *Journal of Geographical Systems*, vol. 9, no. 1, pp. 77–101, 2007.
- [41] S. Hess, T. Adler, and J. W. Polak, "Modelling airport and airline choice behaviour with the use of stated preference survey data," *Transportation Research Part E: Logistics and Transportation Review*, vol. 43, no. 3, pp. 221–233, 2007.
- [42] E. Marcucci and V. Gatta, "Regional airport choice: consumer behaviour and policy implications," *Journal of Transport Geography*, vol. 19, no. 1, pp. 70–84, 2011.
- [43] E. Marcucci and V. Gatta, "Dissecting preference heterogeneity in consumer stated choices," *Transportation Research Part E: Logistics and Transportation Review*, vol. 48, no. 1, pp. 331–339, 2012.
- [44] D. A. Hensher and W. H. Greene, "Specification and estimation of the nested logit model: alternative normalisations," *Transportation Research Part B: Methodological*, vol. 36, no. 1, pp. 1–17, 2002.