Transportation Research Part A 61 (2014) 249–258

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

Why do passengers choose a specific car of a metro train during the morning peak hours?

Hyunmi Kim [a], Sohee Kwon [a], Seung Kook Wu [b], Keemin Sohn [a,*]

[a] Department of Urban Engineering, Chung-Ang University, 221, Heukseok-dong, Dongjak-gu, Seoul 156-756, Republic of Korea
[b] The Korea Transport Institute, 315 Goyangdaero, Ilsanseo-gu, Goyang-si, Gyeonggi-do, Republic of Korea

ARTICLE INFO

Article history:
Received 17 July 2012
Received in revised form 29 January 2014
Accepted 11 February 2014

Keywords:
Metro crowding
Loading diversity factor
Railway capacity
Train car choice
Nested logit model
Latent variable

ABSTRACT

Crowding on metro trains is an important measure of passenger satisfaction and also provides a criterion for determining service frequency and the number of cars necessary for a train set. Particularly in metropolitan areas during morning peak hours, many studies have revealed a considerable difference in the crowding of specific cars on a single train. To accommodate the impact of this phenomenon in calculating metro capacity, a loading diversity factor has been adopted in many transportation studies. However, the underlying causes behind the uneven nature of carriage loading have rarely been examined in a systematic manner. In particular, there has been no trial to explain the nature of choice within a framework for individual passengers. Under the assumption that the uneven selection might stem from each passenger's intrinsic preference for a specific car, the present study established a nested logit model to investigate the potential factors affecting the choice of a specific car on a train. Passengers were interviewed as they boarded from the platforms of line 7 of the Seoul Metro during the morning peak hours. Results show that the motivation to minimize the walking distance at destination stations turned out to be the most decisive in determining a passenger's choice for a specific car of a train.

1. Introduction

Crowding within a train is a determinant of both the service level for passengers and the supply level for operators. Thus, overcrowding in the morning peak hours has been recognized as a serious problem for metro systems in urban areas (Hirsch and Thompson, 2011; Currie, 2010; Hale and Charles, 2009; Qi et al., 2008). It is apparent that the crowding levels differ across individual cars constituting a single train set. According to the transit capacity and quality of service manual (TCQSM) (TCRP report 100, 2003), the passenger load imbalance between cars on individual trains ranges from +61% to −33% with respect to the average passenger load per car in the Vancouver SkyTrain, and fluctuates even more (from +156 to −89%) in Toronto's Yonge Street subway. The survey conducted for the present study showed that the loading difference varied from +118% to −90% for the Seoul Metro line 7 during the morning peak hours. These phenomena commonly imply that the practical capacity of rail transit could be overestimated without properly taking into account the loading differences between the individual cars of a train. Vuchic (1981) suggested the concept of a loading diversity factor to adjust the imbalance when calculating rail transit capacity, which has been adopted in subsequent studies and in a series of transit capacity manuals (TCRP report 100, 2003; TCRP report 13, 1996; Pudney and Wardrop, 2008; Jong et al., 2011). According to the survey

* Corresponding author. Address: 221, Heukseok-dong, Dongjak-gu, Seoul 156-756, Republic of Korea. Tel.: +82 163058192.
E-mail addresses: aadt@hanmail.net (H. Kim), emsjxnal@naver.com (S. Kwon), wsk115@koti.re.kr (S.K. Wu), kmsohn@cau.ac.kr (K. Sohn).

http://dx.doi.org/10.1016/j.tra.2014.02.015

conducted by Pudney and Wardrop (2008), typical values for the factor are 0.8 for heavy rail, 0.75 for light rail, and 0.6 for electric commuter trains. The latter figure connotes that no less than 40% of capacity reduction is expected for metro services due to the uneven passenger load across cars. The uneven utilization of cars not only decreases metro capacity, but it also imparts great disadvantages in the maintenance of rolling stock.

Although the imbalance itself has often been used to exactly derive rail transit capacities, little attention has been paid to determining the underlying causes behind uneven passenger distribution across the individual cars of a train. Identifying the latent causes is expected to help foment policies to distribute passengers more evenly across the individual cars of a train. For example, there could be measures taken to accomplish a more even distribution by controlling passengers' behaviors and attitudes. Furthermore, it might be possible to streamline the dispatching operation of a train. At least for a projected metro line, the infrastructure and train design could be optimized.

It can be inferred that the unevenness of passenger loading can be attributed to the train length, the platform layout, the service frequency, the entrance/exit location, and so forth. However, there is no clear indication of the relative intensity of passenger preference for a specific train car. Furthermore, to the best of our knowledge, individual characteristics and psychometric propensities of metro passengers have never been investigated in association with the choice of a specific car of a train. The present study provides new insight into the potential motives of passengers to choose a specific car of a train within a hierarchical choice structure of individual passengers. The direct objective of the present study was to identify possible reasons for the unevenness of passenger loads across individual cars of a train, which can be utilized for recommending measures to balance the passenger load more evenly.

Section 2 describes how variables affecting train car choice were selected, followed by the data collection methodology in Section 3. A conceptual framework and details of the choice structure is described based on the nested logit model in Section 4. The model estimation results are discussed in Section 5. Section 6 draws conclusions and suggests several measures to promote the even distribution of metro passenger loading across the cars of a train. Further research necessary for advancing the results of the present study is also proposed in Section 6.

2. Determining potential variables

Potential variables that were expected to affect the passenger preference for a specific car of a train were determined based on our pilot survey and a review of the previous literature dealing with conventional mode or route-choice analyses (Sohn and Yun, 2009; Lee et al., 2012; Ben-Elia et al., 2008; Papinski et al., 2009). The chosen variables can be categorized into four groups: individual-specific characteristics, trip-related variables, physical environment, and attitudinal or behavioral propensities.

2.1. Individual-specific characteristics

Conventionally, the individual-specific variables in Table 1 have been regarded as key factors affecting a traveler's decision-making. The background of adopting some unique variables in this category needs to be addressed. Passengers with physical handicaps were expected to find a less crowded car and to try to reduce their walking distance. Obesity might have passengers walking shorter distances or finding a less crowded car.

2.2. Trip-related variables

Variables related to a respondent's current trip encompassed trip purpose, trip frequency, travel times, prior travel experience, number of transfers, awareness of station layout, and whether to transfer (see Table 2). It should be noted that the variables were defined for each respondent with respect to his/her current trip. Two variables regarding transfers were included to test if a transfer passenger had a different preference on choosing a car from non-transfer passengers. While the transfer dummy variable indicated whether a passenger transferred at stations of line 7, the number of transfers included transfers between other lines within a subject's itinerary. There was an expectation that passengers in a party would take on different attitudes when choosing a car of metro trains. Passengers laden with luggage were expected to find a less-crowded car and to try to reduce their walking distance. The variable did not cover handbags under the assumption that women would feel little burden from them. Women in high heels could also be more sensitive to the walking distance. The passenger choice for a car was assumed to vary according to the characteristics of their trip.

2.3. Physical environment

The physical environment around a platform contains entrances/exits, transfer gates, elevators, and escalators. A respondent's accessibility to each facility was quantified as the walking distance between the facility and his/her selected car. The distances were defined only for each respondent's origin and destination stations. These variables were hypothesized to affect the preference of passengers on a specific car of a train. As another physical environment, the crowding within individual cars was expected to be a key factor in the choice of a specific car. Unfortunately, there was no direct way to measure it.

Table 1
Individual-specific variables and the corresponding data description.

Variable	Total		Maximize walking distance at destination		Maximize walking distance at origin		Maximize comfort		Choose a car unintentionally	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Age	37.050	15.818	36.204	15.050	30.077	12.367	36.000	10.583	40.000	17.317
Gender (male: 1; female: 0)	0.494	0.501	0.435	0.497	0.500	0.510	0.400	0.548	0.593	0.493
Marital status (married: 1; unmarried: 0)	0.497	0.501	0.503	0.501	0.654	0.485	0.000	0.000	0.475	0.501
Number of children	0.453	0.516	0.435	0.518	0.269	0.452	0.600	0.548	0.517	0.519
Income (more than 30 million won: 1; less than 30 million won: 0)	0.485	0.501	0.518	0.501	0.462	0.508	1.000	0.000	0.415	0.495
Occupation 1 (student: 1; otherwise: 0)	0.338	0.474	0.372	0.610	0.500	0.510	0.000	0.000	0.314	0.466
Occupation 2(white color: 1; otherwise: 0)	0.374	0.484	0.414	0.494	0.346	0.485	0.600	0.548	0.305	0.462
Education (college-graduated: 1; otherwise: 0)	0.426	0.495	0.440	0.498	0.385	0.496	0.800	0.447	0.398	0.492
Car ownership	0.379	0.559	0.346	0.509	0.385	0.637	0.800	0.447	0.415	0.618
Dwelling type (apartment: 1; otherwise:0)	0.565	0.497	2.487	0.631	2.692	0.618	2.600	0.894	2.525	0.713
Obesity (obesity: 1; otherwise: 0)	0.368	0.483	0.387	0.488	0.231	0.430	1.000	0.000	0.339	0.475
Physical handicap (inconvenience: 1, otherwise: 0)	0.035	0.185	0.042	0.201	0.000	0.000	0.200	0.447	0.025	0.158
Sample size	340		181		43		36		80	

Table 2
Trip-related variables and the corresponding data description.

Variable	Total		Maximize walking distance at destination		Maximize walking distance at origin		Maximize comfort		Choose a car unintentionally	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
Trip purpose (discretionary: 1; compulsory: 0)	0.253	0.435	0.168	0.832	0.231	0.430	0.000	0.000	0.593	0.493
Trip frequency	4.532	2.417	4.890	4.890	4.808	2.117	6.800	4.087	3.797	2.590
Travel times	38.985	19.454	39.770	39.770	34.731	12.951	43.000	12.042	38.483	21.407
Prior travel experience	0.850	0.358	0.916	0.278	0.808	0.402	1.000	0.000	0.746	0.437
Number of transfers	0.685	0.602	0.675	0.675	0.692	0.471	1.000	0.000	0.492	0.502
Awareness of station layout	0.850	0.365	0.942	0.234	0.962	0.196	0.800	0.447	0.678	0.487
Whether to transfer (transferred passenger: 1; otherwise: 0)	0.617	0.487	0.675	0.675	0.692	0.471	1.000	0.000	0.492	0.502
Passenger laden with luggage (with luggage: 1; otherwise: 0)	0.324	0.469	0.335	0.473	0.269	0.452	0.000	0.000	0.331	0.472
Uncomfortable shoes (high heels) (uncomfortable shoes: 1; otherwise: 0)	0.3	0.459	0.304	0.461	0.346	0.485	0.400	0.548	0.280	0.451
Number of accompanying persons	0.088	0.323	0.068	0.309	0.154	0.368	0.000	0.000	0.110	0.341
Sample size	340		181		43		36		80	

Instead, the axle load was available and employed as a surrogate for crowding, which was measured before loading and unloading passengers when a train stopped at each station.

2.4. Attitudinal or behavioral propensities

Attitudinal or behavioral propensities were widely adopted to enhance the explanatory power of choice models in the field of mode and route-choice studies (Prato et al., 2012; Sohn and Yun, 2009; Johansson et al., 2006; Morikawa et al., 2002; Ben-Akiva et al., 2002, 2006). Among indicators adopted in the literature, the present study selected 20 indicators in advance (see Table 3), each of which was a simple question that was expected to be associated with passenger behaviors and attitudes toward the choice for a specific car of a train. Indicators of the first factor in Table 3 were adopted with the intention of identifying how different a passenger's car choice behavior is depending on whether he/she is a heavy user of portable electronic devices. A commercial agency is currently providing metro passengers in Seoul with real-time information about the location of entrances/exits and escalators/elevators for their current trip, which would encourage passengers to use a portable electronic device in choosing their own specific car. Indicators of the second factor reflected the ability

Table 3
Psychometric propensities and their indicators.

	Variable	Indicators
1	Preference for electronic devices	I am accustomed to searching for something with a smartphone; I am an early adapter for IT products; I usually get information from newspapers rather than the Internet[a]; I am a morning person[a]
2	Mnemonic ability	I seldom forget the road I have traveled on; I can easily find things in the dark; I only use the road that I know[a]
3	Health condition	I can climb two stories of a building without a break; I take medicine for chronic diseases like diabetes or hypertension[a]; I walk on escalators
4	Punctuality	I always arrive at the expected time; I cannot understand those who are late for work frequently; I can remember the birthdays of my friends
5	Antisociality	I feel comfortable when I solely shop
6	Impatience	I am upset if someone overtakes me while driving; I cannot wait in a long queue for dinner or lunch; I feel irritated when I am in a crowded situation
7	Active propensity	I enjoy exercise on a regular basis; I often meet someone on business for dinner

Only indicators that had a factor loading value larger than 0.5 are shown.
[a] A negative factor loading.

to remember, with the expectation that those who have a good memory might have advantages in the use of dispersed station facilities. Indicators of the third factor were included to identify how the health condition of a passenger affects his/her car choice. Indicators of the fourth factor were associated with a passenger's sense of time. The remaining indicators were selected since they were expected to be related to a passenger's sociability, activeness, and patience, which also might affect car choice.

3. Survey and data acquisition

3.1. Overview of survey

The test-bed of the present study was line 7 of the Seoul Metro. Line 7 is 46.9 km long with 41 stations and is operated at an average speed of 32.3 (km/h). The line connects from the north-east to the south-west via the Gangnam area, the second biggest employment center in the Seoul metropolitan area. The line dispatches eight-car trains every 5 min during peak hours. It takes 87 min to run the line in one direction. Out of 41 stations of the line, 11 are transfer stations where passengers can transfer to or from other lines. The morning peak hours of line 7 range from 7 AM to 9 AM, which was when the survey was conducted. The physical layout of each station was examined to identify the location of facilities. Line 7 has four types of station layouts. Each type shares common locations of facilities such as entrances/exits, elevators/escalators, and transfer gates.

We randomly selected respondents, so that the entire sample could be distributed proportional to the number of boarding passengers at each station. The sample size was 340, after eliminating data associated with incomplete or contradictory replies. Because of the lack of surveyors, it took 4 weeks to collect the sample. For weekdays of the period, three surveyors chose their own convenient days and conducted the survey during the morning peak hours. They were instructed to randomly choose a subject on a platform, considering the distribution of waiting passengers across each car's stop position. However, this created a possible bias in the sample, since surveyors had to depend on their own visual inspection to find the distribution of passengers. Another bias was an abnormal concentration of subjects in a particular station, which was due to the convenience of the station location that happened to be closest to our laboratory. The results were, however, certainly free from bias since the study focused only on the car-choice behavior of metro riders, which was unlikely to differ from the trip origin.

3.2. Data for dependent variables

The most important thing in the survey was to collect data for the dependent variables of the model. The dependent variable was set as the respondents' responses to the following two questions. The first question asked passengers, who were waiting for boarding in the platform, whether or not they chose a specific car intentionally. Those who answered "yes" were asked a subsequent question concerning their motivation for the choice. Three main options for the motivation were confirmed by a week-long pilot survey in advance. Although a finer categorization of motivations could have been done, these three options were finally adopted because, in the context of statistical modeling, the parsimony was important and there was actually an ignorable number of respondents who had a distinct motivation that should fall in a category other than one of the three motivations. As a result, 76.6% of the respondents reported choosing a specific car intentionally. Among them, 69.7% stated that their motivation was to minimize the walking distance to exit when they disembarked at a destination station, 16.6% reported that they sought to minimize the distance from the entrance when they boarded at an origin station, and the remaining 13.5% reported that they wanted to pursue comfort while traveling. There were also minor groups who had other motivations in the main survey, but they were negligible and eliminated from the sample.

3.3. Data for independent variables

3.3.1. Individual-specific data

Data descriptions for the independent variables that fell on the individual-specific characteristics are shown in Table 1. Surveying the variables did not depend on a self-answering questionnaire to avoid a potential bias. Rather, the first ten variables in Table 1 were collected directly from the interview, while the remaining two variables were observed by the surveyor. Despite the possibility that this would cause some bias associated with surveyors, there was no robust way to survey the variables without a privacy infringement. Surveyors were instructed to see the walking pace of subjects when they entering the platform and to classify them to have obesity if their pace was much slower than the average pace at which most other people walk. A physical handicap was defined as the disability to walk. A person who appeared to be lame or used a wheelchair or crutches was classified as handicapped.

3.3.2. Trip-related data

Data for variables related to a respondent's current trip was also obtained by interview except for the variable of uncomfortable shoes. To collect data for this variable, surveyors used physical observation to approximate the height of the shoes, and shoes that seem higher than 7 cm fell into the high-heel category. Each surveyor asked respondents their trip purpose, trip frequency, travel times, prior travel experience, number of transfers, awareness of station layout, and transfer stations. Five categories of trip purpose were predetermined through the pilot study, so that they could be used in the subsequent modeling. Actually, the five trip purposes were reclassified into two categories of trip purpose and were taken into account in the modeling: compulsory or discretionary trips. Compulsory trips encompassed commuting, work, and school trips, while discretionary trips were for shopping, leisure, and visiting purposes (Otuzar and Willumsen, 2004). The awareness of station layout was measured with a five-point Likert scale. For every metro line, there were transfer stations where passengers could transfer to or from other lines. Since there is no cross-platform transfer in line 7, every transfer station had one or two gates linked to other lines, which was not quite different from normal exits/entrances. The dummy variable for transfer was set at 1 only for respondents who transferred. Origin and destination stations for a respondent's trip were confined to line 7, which, however, does not mean transferred passengers from other lines were excluded. The confinement was only for brevity in modeling. Four types of metro trips were identified in the survey. The first type encompassed trips that used only line 7. In this case, there was no difficulty in determining their origin and destination stations. The second type corresponded to trips that started from other lines, transferred at a station of line 7, and terminated at another station of line 7. In these cases, the transfer station was regarded as an origin station. The third type included trips that started at a station of line 7 and transferred from another station of line 7 to other lines. In these cases, the transfer station was regarded as a destination station. The last category included trips that started and terminated at stations of other lines. In these cases, both transfer stations were origin and destination stations, respectively. In addition, the trip frequency had to be distinguished from the travel experience. The former was applied for respondents with compulsory trips on a regular basis, while the latter was for discretionary trips.

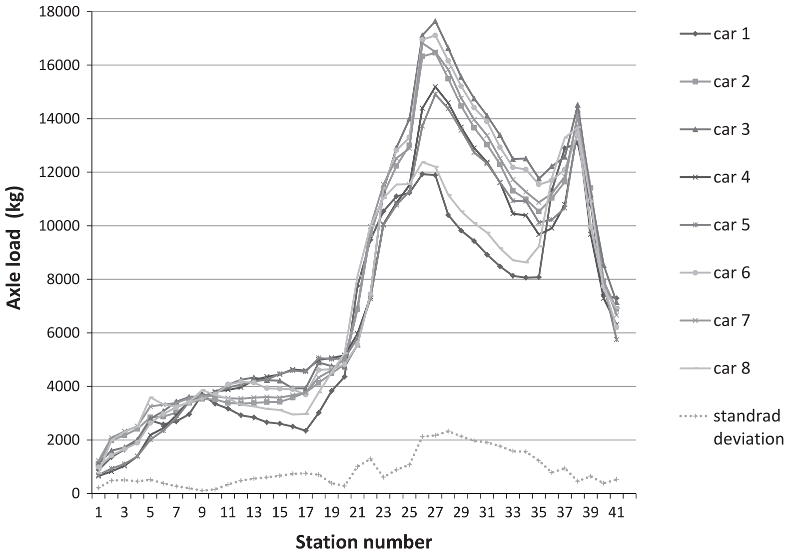
3.3.3. Data for the physical environment

Regarding the collection of variables indicating a respondent's physical environment, all physical aspects of each station had been examined in advance through the pilot survey, so that these, afterwards, could be matched with subjects. To constitute the physical environment variables, surveyors examined distances from a respondent's selected car to each facility in his/her origin and destination stations. Such distances were measured in the laboratory after finishing the field survey. For convenience, rather than using the metric system, the distance was measured by counting the number of cars along a platform between a facility and the stopping position of a respondent's selected car. The distance ranged from 1 to 8 since line 7 of the Seoul Metro, which was a test bed for the present investigation, dispatched eight-car trains in the morning peak hours. The accessibility variable, which was actually used in the modeling, took on the inverse value of the distance variable, so that the variable for the longest distance was set at 1, the variable for the shortest distance at 8, and if the corresponding facility was unavailable or unnecessary for a respondent, the variable was set at zero. For example, a passenger who transferred at a station to other lines tried not to reduce the distance to the exit but, rather, to reduce the distance to the transfer gate. To accommodate this in the modeling, such a respondent was endowed with zero accessibility to the exit.

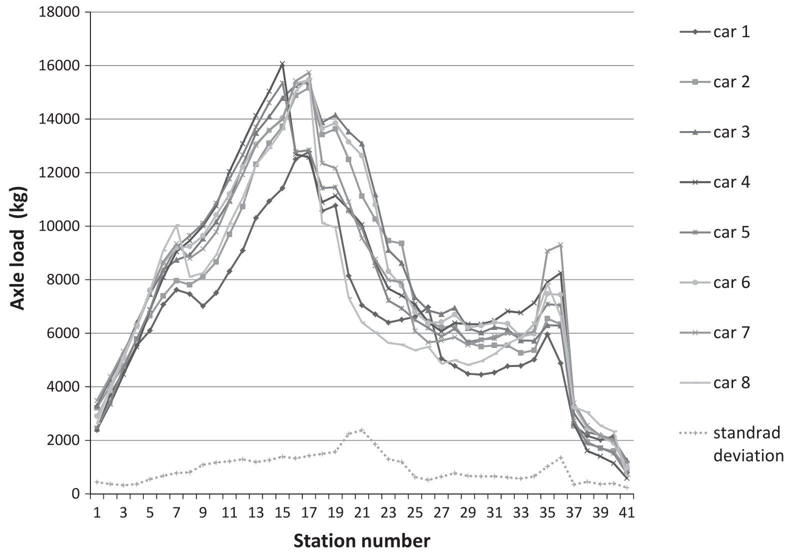
Unfortunately, data for axle loads measured in the peak hours was available only for 5 weekdays in a row that fell on the 1-month survey period. Fig. 1 shows the average axle load distribution for the available data across individual cars. As expected, a considerable difference across individual cars was confirmed. In particular, the discrepancy was exaggerated in consecutive stations with the same layout where the overall passenger load was high (stations 28–36 for south-westbound trains and stations 9–16 for north-eastbound trains). For convenience, rather than directly using the weight, the axle load variable took on a rank in the weight across individual cars, which ranged from 1 to 8 in a decreasing order of axle load. That is, the value of 1 was assigned to the most crowded (=heaviest) car of a single train, and the value of 8 to the least crowded (=lightest car). To constitute the axle load variable, surveyors assigned each respondent a predetermined rank corresponding to his/her chosen car.

3.3.4. Data for latent variables

Indicators of passengers' psychometric propensities were surveyed through a self-answering questionnaire. For convenience, available answers to the questions were confined to a five-point Likert scale that ranged from "strongly negative"



(a) Average axle loads for south-westbound trains



(b) Average axle loads for north-eastbound trains

Fig. 1. Axle load distribution across the individual cars of line 7 in the morning rush hours.

to "strongly positive." The scale included a neutral option but excluded an unsure option. Respondents had to select one of the five options for each indicator. To reduce the number of variables, a principal component analysis (PCA) was applied to the indicators. The PCA was useful in reducing many indicators into a tractable number of factors (=variables). By convention, each factor dictated its main indicators if the absolute value of a factor loading exceeded 0.5. Table 3 shows the resultant 7 variables and their respective indicators after completing the PCA. Each variable was titled based on our judgment after examining the corresponding indicators.

4. Modeling and conceptual framework

Two hierarchical structures were hypothesized to address the passenger preference for a specific car of a train. The first structure had a three-level hierarchy and the second had a two-level hierarchy (see Fig. 2). As for the three-level structure, the top level accommodated the passenger decision about whether to intentionally choose a specific car. At the middle level, the motivation behind the car choice was classified into either minimizing walking distance or maximizing comfort. At the

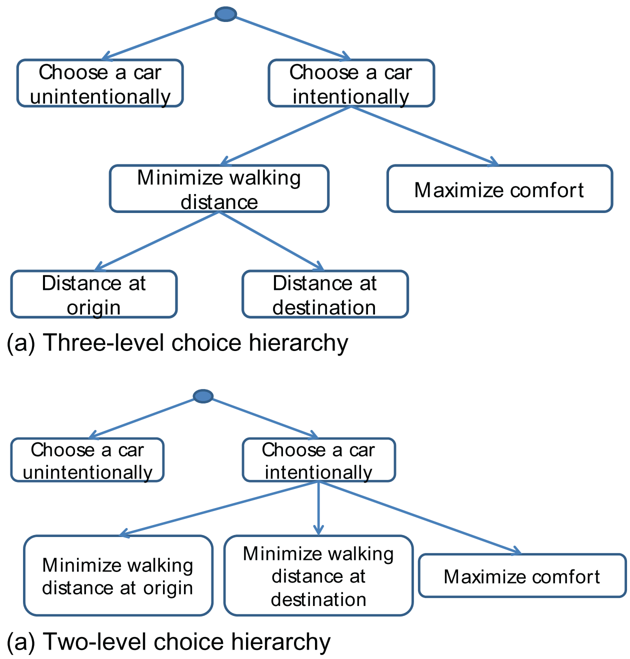


Fig. 2. Hypothesized choice structures.

bottom level, the point at which a passenger wanted to minimize his/her walking distance was separated into his/her origin or destination stations. For the other two-level structure, the middle and bottom levels of the three-level structure were incorporated as a nest. Actually, after estimating two nested logit models, the three-level structure was collapsed to a two-level structure because the inclusive value (IV) parameter for the bottom nest of the three-level model did not significantly differ from unity.

Prior to modeling the two-level choice structure, a conceptual framework was set up, which reflected potential relationships between the selected variables and each end option in the structure, as shown in Fig. 3. Individual-specific variables were expected to affect the upper-level decision as to whether or not a passenger chose a specific car intentionally. Physical environments were assumed to be associated with the lower-level decision on where a passenger would minimize his/her walking distance given that he/she chose a specific car intentionally. Trip-related and psychometric variables were hypothesized to have an influence on decisions at both levels. As a whole, arrows that are augmented but toned down with a mild color in Fig. 3 indicate our own conceptual framework. However, numerous sub-combinations could be made with respect to the relationships between individual variables and each end option. Slender arrows in Fig. 3 show the final model specification after testing as many plausible models as possible within the conceptual framework. Dotted lines in Fig. 3 indicate a negative influence on the linked motivation. That is, a dotted arrow from a variable to its motivation corresponds to the inverse relationship between them.

The final model specification was statistically significant with an acceptable goodness-of-fit. The evaluation of a nested logit model was three-fold (Ben-Akiva and Lerman, 1985). First, the model structure can be validated based on the estimated coefficient for an IV. To rationalize a model structure, the coefficient should differ from unity with a statistical significance at the 0.05 level. The value for the second-level nest of the present model was 0.389 at a statistical significance level of 0.032. Second, the overall goodness-of-fit can be measured by ρ^2 or $\bar{\rho}^2$. The former reflects the contribution of information (=modeling) with respect to the log-likelihood without information. The latter is developed to adjust the impact of the number of coefficients to be estimated. The indices for the present model were 0.287 and 0.273, respectively, which did not deviate from other choice studies (Lee et al., 2012; Sohn and Yun, 2009). Last, the statistical significance of influential variables determined the usability of the model. The results from the present model will be discussed in the next section.

5. Results and discussion

On the whole, the hypothesized choice structure of the two-level hierarchy turned out to be sufficient to address passenger preferences for a specific car of a metro train, since the parameter for inclusive value (IV) was sufficiently different from unity and statistically significant at the 0.05 level. The estimation results from the final model specification offered useful results to identify potential motivations behind the passenger car choice in the morning peak hours. The estimation results are shown in Table 4 and the impact of each variable on each level of decision will be discussed as follows.

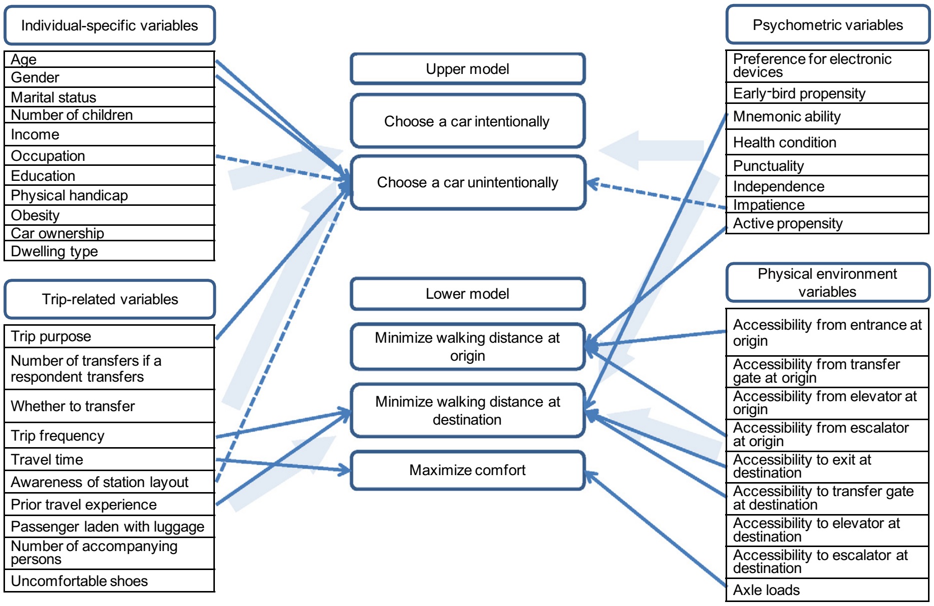


Fig. 3. Conceptual framework.

5.1. Results from the upper model

White collar young females with a compulsory trip purpose were more likely to choose a specific car of a train during the morning peak hours, which was also consistent with the descriptive statistics in Tables 1 and 2. This result also support the notion that young females tended to avoid physical contact with strangers in a crowded car, which is consistent with the results from a focus group interview conducted by Hirsch and Thompson (2011), which found that young female passengers in particular regarded crowded train cars as opportunities for men to surreptitiously take advantage of them. The proclivity that women are averse to a crowded metro train was also confirmed in an effort by two Asian cities to separate women from men during the peak hours of metro operation. The Seoul metropolitan government has introduced metro trains with a specific car reserved only for women. Tokyo metro is now successfully operating such trains. This policy was indirect evidence showing that women feel some kind of fear when they are with strangers in a crowded train. As shown in Table 4, commuters with a compulsory trip purpose tried to minimize their walking distance by selecting a specific car of a train. Among the trip-related variables, a passenger's awareness of the layout of his/her origin and destination stations proved to expedite the preference for a specific car of a train. Unlike prior expectations, the influence of a passenger‘s psychometric propensities was minimal when addressing potential causes behind his/her car choice. Only a coefficient for the variable of impatience among psychometric variables took on an intuitively accountable sign for the upper model. The coefficient was, however, not backed up by statistical significance. A possible reason for the statistical insignificance of the variable, as well as other latent variables, might be due to the sequential estimation methodology, in which PCA was conducted to derive the latent variables independently of the choice model.

5.2. Results from the lower model

The lower-level decision identified the direct motivation behind passenger car choice. Physical environments were a key determinant for passenger car choice based on minimizing the walking distance. The accessibility to entrances/exits was positively associated with the car choice that depended on minimizing walking distance, even though the variable of the origin station had an insufficient statistical significance that slightly exceeded the marginal value (=0.1). It was interesting that at origin stations a transfer passenger's accessibility from a transfer gate to his/her boarding position had nothing to do with his/her car choice to minimize walking distance, while at destination stations a transfer passenger's accessibility from his/her alighting position to a transfer gate was positively associated with the car choice to minimize walking distance. The insignificant transfer accessibility at origin stations reflects the notion that the connection area between a platform and a transfer gate was so crowded in the morning peak hours that passengers from the transfer gate could not board a car stopping close to the area. The accessibility from escalators at origin stations was found to be positively associated with the motivation of car choice to minimize the walking distance. The reason that accessibility to escalators was insignificant at destination stations might also be due to the congestion around escalators when passengers disembark after the arrival of a train. Consequently, passengers tended to place a high priority on the walking distance at their destination station,

Table 4
Estimation results of the two-level nested logit model.

Variable	Coefficient	t-Value	p-Value
Choose a car unintentionally			
Age	0.025***	2.568	0.010
Gender (male: 1, female: 0)	1.143***	3.167	0.002
Occupation (white collar: 1, blue collar: 0)	−0.658**	−1.737	0.082
Trip purpose (discretionary: 1, compulsory: 0)	0.692**	1.695	0.090
Awareness of station layout	−1.985***	−4.663	0.000
Impatience	−0.193	−1.149	0.251
Minimize walking distance at origin			
Active propensity	0.352**	1.735	0.083
Accessibility from entrance at origin	0.112	1.493	0.136
Accessibility from escalator at origin	0.110**	1.792	0.073
Minimize walking distance at destination			
Mnemonic ability	0.323***	2.062	0.039
Accessibility to exit at destination	0.177***	2.577	0.010
Accessibility to transfer gate at destination	0.146***	2.605	0.009
Trip frequency	0.165***	2.244	0.025
Prior travel experience	0.586***	1.944	0.052
Maximize comfort			
Travel time	0.015**	1.617	0.104
Rank in axle load	0.056	0.769	0.442
Nesting parameter			
Inclusive value	0.389***	2.15	0.032
Goodness-of-fit			
Maximum log likelihood	−294.252		
Restricted log likelihood	−457.041		
ρ^2	0.287		
$\bar{\rho}^2$	0.273		

*** Significant at the 0.05 level.
** Marginally significant at the 0.1 level.

regardless of whether they transferred to other lines or terminated their trip there. In addition, the walking accessibility was more important for passenger car choice at destination stations than at origin stations, because at destination stations the variables were more statistically significant.

Among latent propensities and trip-related variables, only the activity was connected to passenger car choice at origin stations at the marginal significance level, but the rationale for this result eluded us. A latent variable and two trip-related variables were statistically significant for car choices based on minimizing the walking distance at destination stations, and thus could be additional grounds for separating the motivation of car choice into that at origin stations and that at destination stations. The mnemonic ability turned out to be positively linked to the passenger choice of a car to minimize the walking distance at destination stations. As expected, the trip frequency positively affected the passenger car choice to minimize the walking distance in destination stations. Passengers who had good memory and made a compulsory trip on a regular basis should be more familiar with the layout of their destination station and, thus, a good memory was more advantageous in minimizing the walking distance at destination stations. The travel experiences proved to be positively associated with the passenger car choice based on the minimization of walking distance at destination stations. This implies that, for discretionary trips, the prior travel experience could help passengers find the shortest walking distance at his/her destination station.

Unfortunately, estimated coefficients from the motivation of car choice to maximize comfort were either marginally or negligibly significant. The model estimation that car choice was based on comfort was expected to depend largely on travel times. That is, passengers with a long trip were assumed to be more likely to select a less-crowded car. The estimated coefficient for the travel time variable managed to match the expectations at the marginal significance level. A serious problem occurred when dealing with the surrogate variable for crowding. According to the sign of the estimated coefficient, the crowding rank variable derived from axle loads turned out to expedite car choice depending on maximizing comfort, but the result had little statistical significance. The reason for yielding an insignificant coefficient for the variable was addressable by the possibility that the axle load data might have lost its accuracy while going through the averaging process. The axle load data during the morning rush hours were available only for 5 days in a row, which were included in the entire 4-week survey period. If the exact crowded level within an actual car chosen by a respondent had been available, the coefficient would have been statistically significant.

6. Conclusions

The most important finding was that 76.6% of the respondents reported choosing a specific car intentionally. Among them, 69.7% stated that their motivation was to minimize the walking distance to an exit when they disembarked at a destination station, 16.6% reported that they sought to minimize the distance from the entrance when they boarded at an origin station, and the remaining 13.5% reported that they wanted to pursue comfort while traveling.

The present study also investigated the potential motivations behind these passenger preferences for a specific car of a metro train. A two-level hierarchical decision structure was hypothesized to account for passenger car choice. Four types of variables were selected to feed a nested logit model based on the two-level hierarchy: individual-specific characteristics, trip-related variables, physical environments, and psychometric propensities.

Regarding the upper-level decision, white collar young females with a compulsory trip purpose were more likely to choose a specific car of a train in the morning peak hours. A passenger's awareness about the layout of his/her origin and destination stations proved to expedite the intentional choice for a specific car of a train. For a lower-level decision, the location of physical environments such as entrances/exits, escalators, and transfer gates proved to have a considerable impact on the motivation for a passenger car choice based on minimizing the walking distance. Mnemonic ability and prior travel experiences were found to influence the car choice to minimize the walking distance at destination stations

The motivation to minimize the walking distance at destination stations was the most decisive determinant of passenger choice for a specific car of a train. This implies that the even utilization of train cars would be possible, at least for a projected metro line, if the precise forecast of travel demand were incorporated with an optimal station layout considering passengers' motivation to minimize the walking distance at destination stations. Furthermore, there can be an operational strategy to more evenly distribute the passenger load with no change in infrastructure. At least for metro lines that have a platform longer than a running train, there might be the possibility to accomplish a more evenly distributed passenger load by differentiating train-stop positions along a platform for each station (Sohn 2013).

Controlling individual-specific characteristics and psychometric propensities could be another key to a more even dispersion of passengers across individual cars. However, the present study does not provide sufficient grounds to foment policy to force passengers to scatter evenly among individual cars. Nevertheless, it might be possible to evenly distribute the passenger load without changes in both the structure and operation of a metro line. The promotion of walking could be introduced as an effective way to accomplish the load balance, particularly when associated with health enhancement benefits. Further research should focus on excavating more latent variables to accomplish an even utilization of specific cars of a train.

The uneven distribution of passengers across cars has been a well-known problem the world over. As far as could be ascertained, however, the present study is the first to explore possible solutions for the problem in a systematic manner at the individual passenger level. It is uncertain that the findings of the present study are transferable to other contexts. The present study, however, could trigger ensuing studies to find answers to the questions of transferability.

References

Ben-Akiva, M., Lerman, S., 1985. Discrete Choice Analysis. The MIT Press, Cambridge, Massachusetts.
Ben-Akiva, M., Walker, J., Bernardino, A.T., Gopinath, D., Morikawa, T., Polydoropoulou, A., 2002. Integration of choice and latent variable models. In: Mahmassani, H.S. (Ed.), Perpetual Motion: Travel Behaviour Research Opportunities and Application Challenges. Elsevier, Amsterdam, pp. 431–470.
Ben-Akiva, M., Choudhury, C., Toledo, T., 2006. Modeling latent choices: application to driving behavior. In: Paper Presented at the 11th International Conference on Travel Behaviour Research, Kyoto, Japan, pp. 16–20.
Ben-Elia, E., Erev, I., Shiftan, Y., 2008. The combined effect of information and experience on drivers' routechoice behavior. Transportation 35 (2), 165–177.
Currie, G., 2010. Quick and effective solution to rail overcrowding: free early bird ticket experience in Melbourne, Australia. Transport. Res. Rec. 2146 (1), 35–42.
Hale, C., Charles, P., 2009. Managing peak demand for passenger rail: a literature review. In: 32nd Australasian Transport Research Forum (ATRF). Auckland, New Zealand, ATRF.
Hirsch, L., Thompson, K., 2011. I can sit but I'd rather stand: commuter's experience of crowdedness and fellow passenger behaviour in carriages on Australian metropolitan trains. In: Proceedings of Australian Transport Research Forum 2011, Adelaide, Australia.
Johansson, M.V., Heldt, T., Johansson, P., 2006. The effects of attitudes and personality trait on mode choice. Transport. Res. Part A 40, 507–525.
Jong, J.-C., Lai, Y.-C., Huang, S.-H., Chiang, P.-C., 2011. Development and application of rail transit capacity models in Taiwan. Transport. Res. Rec. 2216, 125–138.
Lee, I., Park, H., Sohn, K., 2012. Increasing the number of bicycle commuters. Proc. Inst. Civil Eng. – Transport 165, 63–72.
Morikawa, T., Ben-Akiva, M., McFadden, D., 2002. Discrete choice models incorporating revealed preferences and psychometric data. In: Franses, P.H., Montgomery, A.L. (Eds.), Econometric Models in Marketing, Advances in Econometrics, vol. 16. Elsevier, Amsterdam, pp. 29–55.
Otuzar, J., Willumsen, L.G., 2004. Modelling Transport, third ed. John Wiley, Chichester, UK.
Papinski, D., Scott, D.M., Doherty, S.T., 2009. Exploring the route choice decision-making process: a comparison of pre-planned and observed routes obtained using person-based GPS. Transport. Res. Part F 12 (4), 347–358.
Prato, C.G., Bekhor, S., Pronello, C., 2012. Latent variables and route choice behavior. Transportation 39, 299–319.
Pudney, P., Wardrop, A., 2008. Generating train plans with problem space search. Lecture Notes in Economics and Mathematical Systems 600 Part II, pp. 195–207.
Sohn, K., 2013. Optimizing Train-Stop Positions along a Platform to Distribute the Passenger Load More Evenly across Individual Cars. IEEE. T. Intell. Transp. 14 (2), 994–1002.
Qi, Z., Baoming, H., Li, D., 2008. Modeling and simulation of passenger alighting and boarding movement in Beijing metro stations. Transport. Res. Part C 16, 635–649.
Sohn, K., Yun, J., 2009. Separation of car-dependent commuters from normal-choice riders in mode-choice analysis. Transportation 36, 423–436.
Transit Co-operative Research Program (TCRP) Report 100, 2003. Transit Capacity and Quality of Service Manual. Transportation Research Board/National Academy Press.
Transit Co-operative Research Program (TCRP) Report 13, 1996. Rail transit capacity. Transportation Research Board/National Academy Press.
Vuchic, V.R., 1981. Urban Public Transportation Systems and Technology. Prentice Hall Inc., Englewood Cliffs, NJ.