Hindawi
Journal of Advanced Transportation
Volume 2019, Article ID 6873912, 20 pages
https://doi.org/10.1155/2019/6873912

Research Article

Measuring Retiming Responses of Passengers to a Prepeak Discount Fare by Tracing Smart Card Data: A Practical Experiment in the Beijing Subway

Qingru Zou,[1,2] Xiangming Yao,[2] Peng Zhao,[2] Zijia Wang,[3] and Taoyuan Yang[3]

[1]*School of Traffic and Transportation, Shijiazhuang Tiedao University, Shijiazhuang, Hebei 050043, China*
[2]*School of Traffic and Transportation, Beijing Jiaotong University, No. 3 Shangyuancun, Haidian District, Beijing 100044, China*
[3]*School of Civil Engineering, Beijing Jiaotong University, No. 3 Shangyuancun, Haidian District, Beijing 100044, China*

Correspondence should be addressed to Xiangming Yao; yaoxm@bjtu.edu.cn

Received 19 December 2018; Revised 28 March 2019; Accepted 15 May 2019; Published 4 June 2019

Academic Editor: Rocío de Oña

Copyright © 2019 Qingru Zou 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.

Understanding passengers' responses to fare changes is the basis to design reasonable price policies. This work aims to explore retiming responses of travelers changing departure times due to a prepeak discount pricing strategy in the Beijing subway in China, using smart card records from an automatic fare collection (AFC) system. First, a new set of classification indicators is established to segment passengers through a two-step clustering approach. Then, the potentially influenced passengers for the fare policy are identified, and the shifted passengers who changed their departure time are detected by tracing changes in passengers' expected departure times before and after the policy. Lastly, the fare elasticity of departure time is defined to measure the retiming responses of passengers. Two scenarios are studied of one month (short term) and six months (middle term) after the policy. The retiming elasticity of different passenger groups, retiming elasticity over time, and retiming elasticity functions of shifted time are measured. The results show that there are considerable differences in the retiming elasticities of different passenger groups; low-frequency passengers are more sensitive to discount fares than high-frequency passengers. The retiming elasticity decreases greatly with increasing shifted time, and 30 minutes is almost the maximum acceptable shifted time for passengers. Moreover, the retiming elasticity of passengers in the middle term is approximately twice that in the short term. Applications of fare optimization are also executed, and the results suggest that optimizing the valid time window of the discount fares is a feasible way to improve the congestion relief effect of the policy, while policy makers should be cautious to change fare structures and increase discounts.

1. Introduction

1.1. Motivation. The Beijing subway is suffering from serious congestion due to booming travel demand, especially during the morning peak hours (7:00 a.m.~9:00 a.m.). To relieve overcrowding, transit agencies implemented a discount pricing measure in 2016 that provided a 30% discount for passengers who checked in before 7:00 a.m. It was the first case in China to use time-dependent pricing policy in rail transit systems, and the measure was only a pilot policy at certain stations. After one year of the experiment, the discount rose from 30% to 50%, and the number of trial stations increased from 16 to 23. However, the policy is still in the pilot stage and is not expanding to a larger scale because the responses of passengers to the discount fare remain unclear, and no obvious reduction in peak demand is observed.

Differential pricing is an important and effective measure in traffic demand management (TDM). Whether pricing strategies work depends strongly on the responses of travelers to fare changes. Normally, the fare elasticity of demand is used to describe the relationship between demand changes and fare changes at an aggregate level. It is useful to estimate demand changes for system-wide and long-term pricing policies. However, for regional and short-term (valid time window is short) fare strategies, it is hard to capture the reactions just by the fare elasticity of demand. For instance, the prepeak discount price in the Beijing subway is valid

only before 7:00 a.m., which mainly influences passengers' departure time choices and will not result in obvious changes in demand. There is a great need to consider the microscopic travel responses of passengers to fare changes and then provide support to establish a precise pricing scheme.

Passengers' travel responses to fare changes are very complex and related to various external factors, such as service quality, travel preference, and socioeconomic factors. Previous works on travel responses are usually specific to a certain region or transit system and assume that the external factors remain the same before and after the policy. Unfortunately, few works on travel responses for transit systems in China are found, as the Beijing subway is the first case to apply time-of-day pricing policy in transit systems. Additionally, traditional works are mainly based on statistical methods using SP/RP (Stated Preference/ Revealed Preference) survey data, which greatly limits the potential to study passengers' travel responses in depth. The RP survey only creates a static and sampled picture of the traffic pattern in a short time period, which introduces errors as people's travel behavior changes over time [1]. The SP survey, based on fictitious situations, may have large errors because the choices of interviewees may be different from their actual behaviors. As a benefit of the application of the AFC system in rail transit, the smart card data produced from the AFC system provide valuable insights into the usage of public transport and help to better understand travel demand, behavior, and other purposes [2]. The AFC system not only records the accurate trip information of each passenger but also makes it possible to track behavior changes of passengers before and after a new policy.

Hence, this work aims to explore the retiming behavior of passengers for departure time choice and determine which passengers will travel early and how long they will shift from their original departure time to enjoy the discount fare. The main contributions of this work contain three aspects. (1) A new set of classification indicators, which is obtained directly from smart card data, is proposed to segment passengers into groups and then improve the objectivity of classification. (2) The shifted passengers who change their departure time are identified by comparing the expected departure times before and after the fare policy, and then the retiming elasticity of different passenger groups, retiming elasticity over time, and retiming elasticity functions of shifted time are measured for the Beijing subway travelers. (3) According to the results of the retiming elasticity, the active time threshold of the discount fare is very short, and valid time window optimization of the discount price is very important to improve the effect of the policy.

1.2. Literature Review. Passengers' responses to fare changes are usually indicated by demand changes. The fare elasticity of demand, which is widely used to measure passengers' responses to fare changes, is defined as the percentage change in travel demand due to a 1% change in fare. A large number of case-specific fare elasticities have been studied for traffic and transportation systems. Bresson et al. noted that the elasticity of transit demand can generally be regarded as -0.3 without considering case-by-case differences [3]. However, travel responses are complex and comprehensive behaviors of passengers, not only for price but also related to various external factors. Other works have indicated that demand elasticity varies greatly in terms of variables such as age, time span, transit mode, time of day, original fare level, income level, trip distance, data paradigm, type and direction of a price change, and demographic and geographic conditions [4–8]. To reduce the influence of external factors, the case-specific fare elasticity should usually be measured when designing new fare schemes for a certain traffic system. For rail transit systems, Holmgren studied 81 transit fare-change cases and found a price elasticity of demand of -0.38 [9], while Hensher extracted a value of -0.395 from 319 cases [7]. Schimek obtained an elasticity of -0.34 using data collected from 198 transit agencies during 1991 and 2012 [10]. According to the data collected in the UK, Paulley et al. showed that the price elasticities of the metro are -0.26 for peak hour demand and -0.42 for off-peak demand. Furthermore, they recommended a range of short-term transit fare elasticities of -0.2 to -0.5 (-0.15 to -0.3 for peak hours) and -0.37 for the metro in a predicted scenario [11]. However, considering the large variety in the values, it has been suggested that a range rather than an absolute value should be provided for fare elasticities [8]. Taking the fare change in the Beijing subway in 2014 as the background, in which a distance-based fare is applied to replace the flat fare structure, Wang et al. studied the price elasticities of demand by means of an SP survey with 4210 samples, and the values of the price elasticities ranged from -0.232 to -1.143 [12]. Furthermore, Wang et al. studied the price elasticities of different distances on weekdays and weekends using smart card data one week before and after a fare change [13]. Comparing the results of these two works, the elasticities from the SP survey are much larger than those revealed via the smart card data. In conclusion, the fare elasticities of demand differ greatly for different regions, transit systems, and time periods (such as peak and off-peak). It is necessary to measure the fare elasticities for a specific traffic and transit system when designing new fare schemes.

Another approach to measure the travel responses of passengers to fare changes is using travel behavior models. The fare is treated as a factor that affects the travel behaviors of passengers, such as the departure time choice, travel mode choice, and travel route choice. The most commonly used travel behavior models are the logit-based model and its derivative models. Mahmassani et al. [14], Lu et al. [15, 16], and Lu and Mahmassani [17] developed a multicriteria route and departure time user equilibrium model for dynamic traffic assignment applications with variable toll pricing. Aboudina et al. proposed an integrated model by considering departure time choice and travel route choice [18], while Habib et al. integrated departure time choice and travel mode choice [19]. Based on these travel behavior models, changes in demand can be estimated by a traffic assignment model, and fare policies can be evaluated through demand changes [20]. Simulations are usually used to examine the aggregate impacts of differential pricing measures on a transit or multimodal transport network. Liu and Charles noted that the functionality of a simulation system depends on its core choice model and its assumptions of individuals'

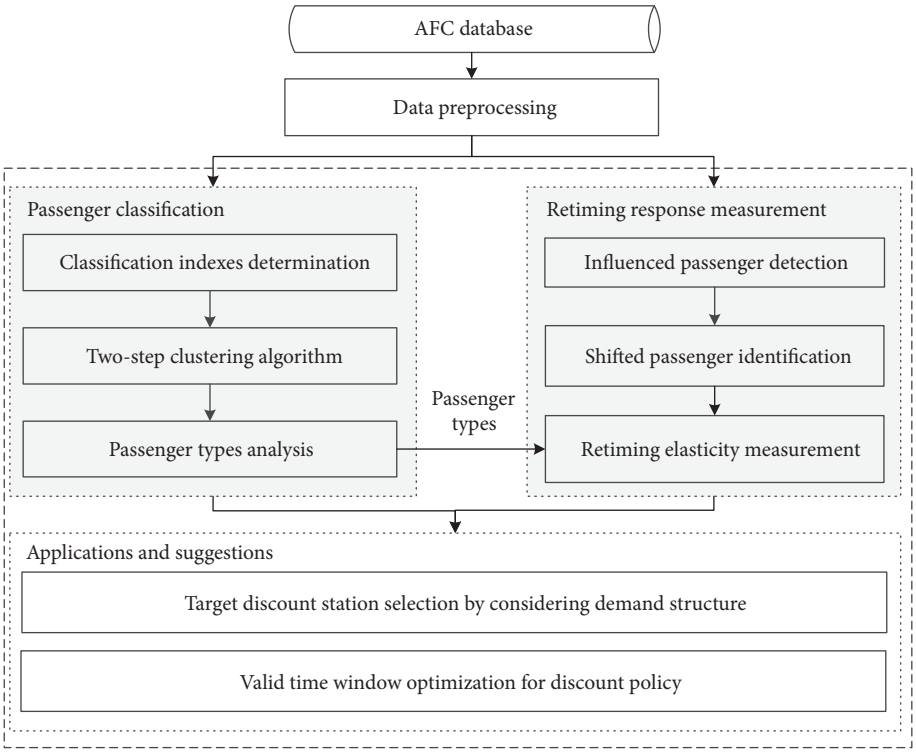


Figure 1: Framework for studying passengers' retiming responses to the prepeak discount fare.

"travel behavior," e.g., a traveler's response and sensitivity to a peak fare surcharge alone or combined with changes in other service attributes, such as onboard crowding and service frequency [21]. However, the heterogeneity and microscopic travel behaviors of passengers can be considered in these methods, and some necessary attributes in the model are hard to obtain (such as personal attributes and travel characteristics at the individual level). Due to the lack of advanced and intelligent data acquisition equipment, most studies use limited sample survey data or highly aggregated collection data. The available data greatly limits the potential to study the travel responses of passengers in depth.

With the wide application of AFC systems in public traffic and transportation systems, massive and continuous trip information can be obtained. The unique ID of smart cards makes it possible to trace day-to-day behavior changes of passengers at the individual level. Pelletier et al. undertook a complete review of transit smart card research, and they suggested that there was great potential for demand management using card data [22]. Compared to survey data, smart card data have outstanding characteristics for travel behavior analysis, such as being accurate, continuous, and trackable. However, there are also intrinsic limitations of smart card data, including the absence of socioeconomic properties (age, career, income, home/work location, etc.) and travel information (trip purposes, trip chains, etc.) [2]. In recent years, many researchers have tried to mine indirect and potential information from smart card data, such as travel pattern analysis [23], trip purpose detection [24, 25], and travel behavior analysis [26]. Although various studies have been performed on smart card mining, the data have rarely been used to study a practical policy. Hence, taking the prepeak discount fare in the Beijing subway as the background, this work aims to study the travel responses of passengers for changing departure times for discount fares by tracing the behavior changes of individual passengers through the actual trip information from smart card data.

1.3. Organizations. Figure 1 provides the framework for analyzing the retiming responses of travelers under a prepeak discount price policy. There are two important parts of this work: passenger classification and retiming response measurement. Though detailed trip information (such as travel time, origin-destination, and fare) for each cardholder can be obtained from smart card data, it is still difficult to measure behavior changes at the level of individual travelers for a huge number of passengers. Hence, an unsupervised clustering approach is employed to classify passengers into groups according to their impersonal travel characteristics (such as trip times in a day or week). In regard to the retiming response measurements, the potentially influenced travelers who may be affected by the fare policy are detected first, and then a before-and-after method is used to determine which passengers shift their departure time. Once we obtain these shifted passengers, a deeper analysis of the retiming responses can be performed. Finally, applications of discount price optimization in two aspects are provided.

The remainder of this paper is organized as follows. Section 2 describes the background of the prepeak discount price policy implemented in the Beijing subway, and the data

Table 1: Development of the Beijing subway in the last decade.

Year	Number of lines	Network length (km)	Annual ridership (billion)	Daily ridership (ten thousand)	Growth rate (%)
2008	8	200	1.22	332	/
2009	9	228	1.42	390	16.39
2010	14	336	1.85	506	30.28
2011	15	372	2.19	601	18.38
2012	16	442	2.46	673	12.33
2013	17	465	3.20	878	30.08
2014	18	527	3.39	929	5.94
2015	18	554	3.32	910	-2.06
2016	19	574	3.66	1003	10.24
2017	21	588	3.78	1035	3.27

Table 2: Prepeak discount price scheme in the Beijing subway.

Line	Discount stations	Number	Discount percentage				Valid time window
			2016	2017	2018	2019	
Line BT	TQ, LHL, LY, JKS, GY, TZBY, BLQ, GZ, SQ, CMDX, GBD	11	30%	50%	50%	50%	check in before 7:00 a.m.
Line CP	NS, SHGJY, SH, GHC, ZXZ	5	30%	50%	50%	50%	check in before 7:00 a.m.
Line 6	BYHX, TZBG, WZXYL, CF, CY, HQ, DLP, QNL	8	/	50%	50%	50%	check in before 7:00 a.m.

source is supplied. The clustering-based passenger classification is shown in Section 3. Further, methods and rules for identifying the potentially influenced travelers and shifted passengers are provided first. Then, the retiming elasticity for different passenger types, changes in retiming elasticity over time, and retiming elasticity functions of shifted time are measured in Section 4. Afterwards, two examples of optimizing fare schemes and suggestions are given in Section 5. Finally, conclusions and future works are summarized.

2. Background

2.1. Prepeak Discount Price Policy in the Beijing Subway. The Beijing subway has been one of the largest and most congested transit systems in China. In the last decade, the length of the network expanded from 200 km to 588 km, and the average daily ridership grew to more than 10 million [27], as shown in Table 1. Although the train headway of certain lines has been decreased to the minimum time (less than two minutes), the limited transport capacity still cannot satisfy the booming demand, especially during peak hours.

To relieve the heavy congestion during the morning peak hours (7:00~9:00 a.m.), a pilot discount pricing strategy was first implemented in 2016 that provided a 30% discount for travelers who checked in before 7:00 a.m. This policy was tested at 16 stations located on Line BT and Line CP on normal weekdays. This measure was designed to encourage travelers to shift their departure time from peak hours to prepeak hours and thus spread the demand more evenly over the time period. Unfortunately, the effect of the policy was not noticeable after one year of the experiment. In 2017, the discount increased to 50%, and 8 new stations on Line 6 were added. At present, the same prepeak discount price scheme with a 50% discount is still implemented at these 24 stations. Detailed information for the prepeak discount price scheme is shown in Table 2, where the names of stations are represented by acronyms (the same acronyms are used hereinafter).

There exist many types of differential pricing patterns with specific objectives, such as fare increases to increase revenue. The main reasons for only adopting a discount price scheme in the Beijing subway have two aspects; the first and most important one is to relieve peak congestion, and the second is that managers do not want to shift congestion from rail transit to bus or road traffic. It is not a wise choice to ease subway congestion at the expense of congestion in other traffic modes.

2.2. Impacts of the Policy on Demand. It should be clear whether the discount policy has positive impacts (e.g., peak reduction or peak shift) on travel demand before exploring the retiming behavior in depth. If no obvious changes in demand are observed, there is no need for further study of the travel responses to fare changes. Considering that no

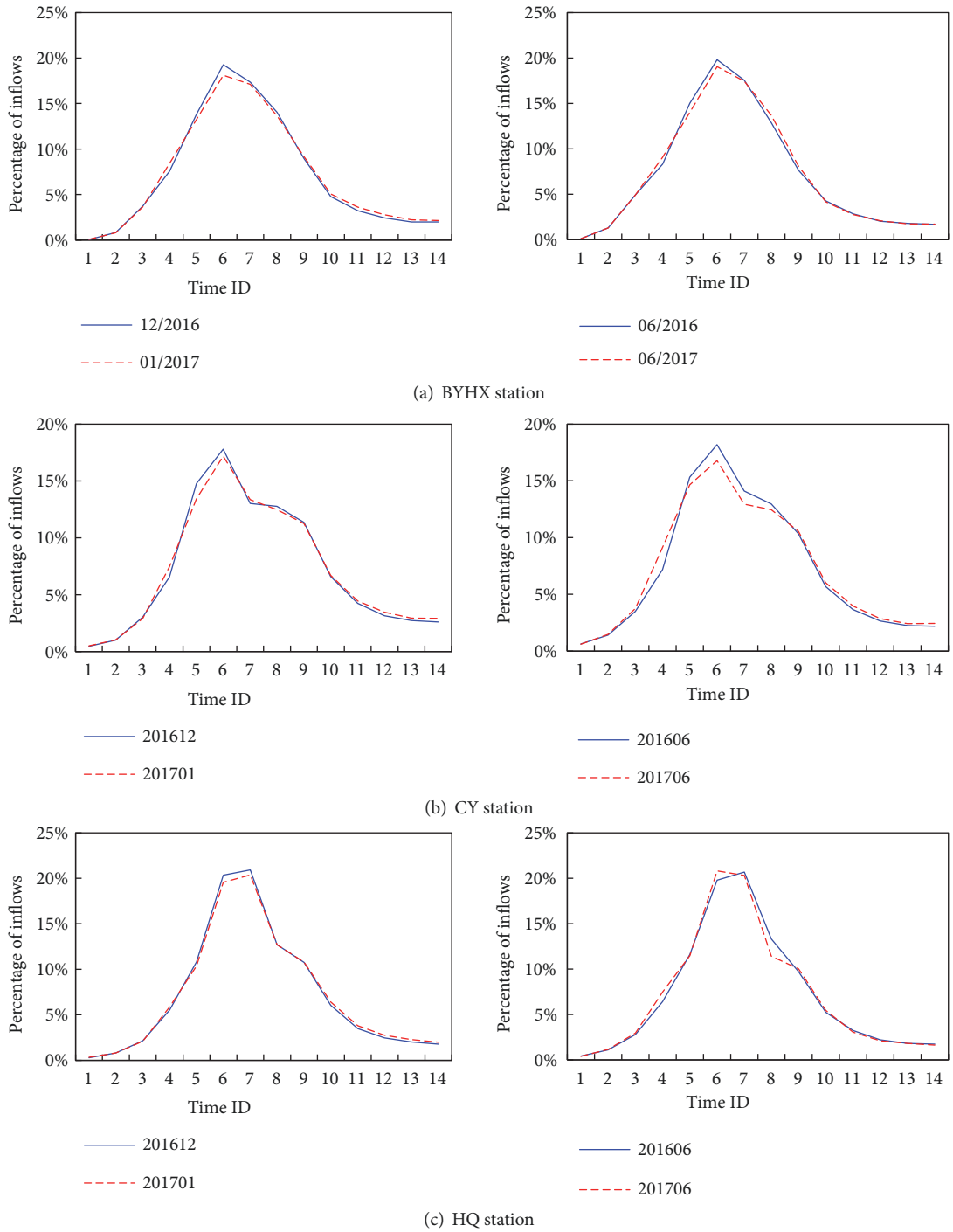


Figure 2: Demand distributions of stations before and after the fare policy.

significant effect was observed when the 30% discount pricing strategy was applied in the first year (2016), this work only focuses on the 50% discount fare strategy. A simple evaluation of the demand distribution is provided to show an overview of the influence of fare changes on demand. BYHX, CY, and HQ stations are selected as the representative stations to show the changes. The analysis time horizon (5:00 a.m. to 12:00 a.m.) is divided into equal lengths of 30 minutes that are consecutively numbered from 1 to 14.

Figure 2 presents the distribution of station inflows at the representative stations, in which the inflows are the average values in all weekdays of a month. The short term (one month after the policy) and the middle term (six months after the policy) are both explored. The following can be observed: (1) The percentage of peak demand exhibits a small decline during peak hours, indicating that the pricing strategy has a positive impact on managing ridership growth. (2) The peak time in CY and HQ stations moves forward, which means that

Table 3: Data structure of smart card transaction records.

Field Names	Comments
GRANT_CARD_CODE	Card ID
CARD_TYPE	Card type
TRIP_START_TIME	Check-in time
TRIP_ORIGIN_LOCATION	Origin station ID
DEAL_TIME	Check-out time
CURRENT_LOCATION	Destination station ID

the discount price can indeed influence travelers' departure times.

In summary, the prepeak discount pricing policy indeed has positive impacts on travel demand, though to an unequal extent on different stations. The reason for the discrepant impacts may be that the demand structure for each station is different, and passengers (such as commuters and irregular travelers) have differential sensitivities to fare changes.

2.3. Data Source and Data Processing. The AFC system was first equipped in Beijing on May 10, 2006, and has been used in both bus and subway systems. Travelers should swipe their smart cards when they check in and check out, and then two transaction records are archived for a trip. The primary information in the smart card transaction records used in this work is represented in Table 3.

Due to the large number of smart card transaction records, it is impossible to analyze travel behaviors by the original data directly. Two steps of data preprocessing are conducted before analysis. (1) The first step is data cleaning, which aims to exclude useless records. There are six types of smart cards in the Beijing subway, of which the one-trip temporary card and card for work cannot link to a constant cardholder; hence, the records related to these cards are neglected. Moreover, records on weekends and holidays are invalid because the prepeak discount price policy is only implemented on normal weekdays. (2) The second step is data conversion, which involves restructuring the trip information for each cardholder. A trip is defined as a base unit, which contains the origin and destination stations and the corresponding arrival and departure times. If all trips for a passenger in a period are extracted and sorted by the check-in time, the consecutive trip sequences can be acquired. The trip sequences can be easily used to compute travel indexes and analyze travel behaviors for the subsequent retiming analysis. An example of the trip sequences for a cardholder is shown in Table 4.

Two different pricing schemes were implemented in the Beijing subway, namely, a 30% discount in 2016 and a 50% discount in 2017. Considering that there is no apparent effect of the 30% discount pricing scheme, only the new pricing plan (50% discount) started in 2017 is chosen for our analysis in this work. To avoid the overlapping influence of the 30% discount price in 2016, the 8 newly added stations located on Line 6 are selected as our target stations, which are BYHX, TZBG, WZXYL, CF, CY, HQ, YLP, and QNL stations.

Generally, time is needed for travelers to adjust their behaviors in response to a new policy. In addition, seasonal factors might affect travelers' timing behaviors, such as passengers being less willing to travel earlier in the dark winter. Four months of smart card data are used in this study, comprising one month before and after the policy (December 2016 and January 2017) for the short-term analysis and six months before and after the policy (June 2016 and June 2017) for the middle-term analysis.

3. Clustering-Based Classification of Passengers

Passenger classification or market segmentation provides an important method for travel behavior and traffic demand analysis, which can be utilized to explore the regularity and similarity for the same passenger group. Traditional classifications are usually according to passengers' travel or personal attributes, such as trip purpose, age, occupation, and income. However, the assumption that travelers with similar personal or travel attributes will exhibit similar travel behaviors is doubtful [28]. In this work, new segmentation indexes produced directly from smart card data and an unsupervised two-step clustering algorithm is employed to segment passengers objectively. Then, it is able to explore the responses of different passenger groups to the prepeak discount fare policy.

3.1. Classification Indexes. Three aspects of travel behaviors are considered for classification, travel intensity, temporal travel characteristics, and spatial travel characteristics. The fundamental rule for constructing indexes is that the variable can capture the travel behavior differences of different passenger groups.

(1) Travel Intensity. Travel intensity describes the usage of passengers traveling by subway. The greater the intensity is, the higher the loyalty of passengers on the subway. Travel intensity is represented by three variables as following. These travel indexes can be gathered directly from smart card records.

(a) *Number of trips per day* represents the daily usage of passengers traveling by subway.

(b) *Travel days per week* indicates the weekly usage of passengers traveling by subway. Usually, the number of travel days for commuters is larger than for other types of travelers.

(c) *Standard deviation of travel days per week* is used to measure the travel stability of passengers. Regular passengers have higher stability.

(2) Temporal Travel Characteristics. Normally, travel time can reflect the passenger types to a certain extent. For instance, if the first-trip departure time of a passenger is generally during the morning peak hours, he/she may be a commuter. Temporal travel characteristics explain travel behaviors related to time and are described by the three following indexes.

Table 4: An example of the trip sequences for a cardholder after data conversion.

Card ID	Card Type	Trip Sequence
28172180	normal card	{(ADM, 12/1/2016 8:12:00, BJZ, 12/1/2016 8:29:00); (JGM, 12/1/2016 17:44:00, ADM, 12/1/2016 18:00:00); (ADM, 12/2/2016 8:17:00, BJZ, 12/2/2016 8:34:00); (ADM, 12/3/2016 8:09:00, BJZ, 12/3/2016 8:26:00); (ADM, 12/4/2016 8:06:00, BJZ, 12/4/2016 8:23:00); (ADM, 12/5/2016 8:02:00, BJZ, 12/5/2016 8:18:00); (...) (...) (...)}

Table 5: Summary of classification indexes.

Dimension	Classification index	Unit	Range of value
travel intensity	travel days per week	days/week	$(0, 5]$
	standard deviation of travel days per week	none	$[0, 2.5]$
	number of trips per day	trips/day	$[1, 9]$
temporal travel characteristics	start time of the first trip	minute	$[300, 720]$
	standard deviation of the first-trip start times	none	$[0, 210]$
	activity duration per day	minute	$[0, 1200]$
spatial travel characteristics	OD coverage rate	none	$(0, 1]$
	trip distance (travel time)	minute	$[10, 240]$

(a) *Start time of the first trip* represents the median value of the first-trip departure times of a passenger over a long period. The reason for using the median is that it is more robust than the average value due to outliers. In addition, universal time is converted into minutes (start time is midnight) for easy calculation (the same for other time-related variables). In this work, the earliest first-trip start time is 5:00 a.m. (operation start time of Beijing subway), and the latest time is 12:00 a.m. Travel records later than 12:00 a.m. are not considered for statistics of first-trip start time. Hence, the range of the index is 300 to 720.

(b) *Standard deviation of the first-trip start time* measures the stability of the start time during a period. Combined with the standard deviation of travel days per week, it is easy to segment regular and irregular travelers.

(c) *Active duration time per day* is the duration time between the first-trip start time and the last-trip end time in a day. The duration time can be used to infer passenger types. For example, the duration of commuters and workers is usually larger than eight hours, while other types do not have apparent characteristics. The maximum duration time per day is the whole operation time, which is 1200 minutes (5:00 a.m. to 24:00 a.m.).

(3) Spatial Travel Characteristics. Spatial travel characteristics describe the discrepancies in travel space for different passengers. For example, commuters always travel within a constant origin-destination (OD) pair. Two variables are established to describe the spatial travel characteristics of passengers.

(a) *OD coverage rate* is the percentage of the number of traveled OD pairs to the number of trips. The smaller the OD coverage rate, the higher the travel stability in space. The range of coverage rate is from 0 to1.

(b) *Trip distance* is used to characterize the range of passengers' activities in space. Considering that the fare structure in the Beijing subway consists of a distance-based price pattern, the fare or travel time is nearly proportional to the trip distance. In this work, the travel time is used as a proxy variable for the trip distance. According to the statistics of smart card data, the maximum cost time of a trip for Beijing subway travelers is almost 4 hours, and the minimum is about 10 minutes. Hence, the range of the index is from 10 to 240.

Table 5 provides a summary of these classification indexes, in which the range of values is determined according to the statistical results from smart card data. The index of travel days per week, number of trips per day, activity duration per day, and trip distance are the average value, while the start time of the first trip is the median value.

3.2. Classification Method. Clustering is a kind of unsupervised classification method for segmenting samples according to the natural attributes of samples, without any a priori classification rules. There are many clustering algorithms, such as partitioning-based methods, hierarchical-based methods, density-based methods, and grid-based methods [29]. A clustering algorithm usually has distinctive characteristics and is suitable for different types of data structures. Considering that there are millions of travelers on the Beijing subway in a day, efficiency is an important factor for choosing a suitable clustering method. After a series of tests and analysis, a hierarchical-based two-step clustering algorithm is selected for its high computational efficiency and accuracy.

Table 6: Center point for each group.

Travel indexes	Group 1	Group 2	Group 3	Group 4	Group 5
average travel days per week	4.64	4.49	3.21	2.29	0.96
standard deviation of travel days per week	0.27	0.39	1.31	0.74	0.24
average number of trips per day	1.88	2.25	1.76	1.72	1.41
start time of the first trip	484.98	474.90	474.61	554.62	364.47
standard deviation of the first-trip start times	13.56	49.99	23.07	62.23	5.85
average activity duration per day	737.16	777.39	700.68	550.47	171.59
OD coverage rate	0.17	0.33	0.30	0.62	0.80

Table 7: Clustering results for different weekdays.

Date	Percentage of each passenger group in all passengers(%)				
	Group 1	Group 2	Group 3	Group 4	Group 5
2016.12.5(Monday)	32.24	12.18	13.42	17.53	24.62
2016.12.6(Tuesday)	32.10	12.27	14.12	17.68	23.84
2016.12.7(Wednesday)	33.03	12.18	13.00	17.39	24.41
2016.12.8(Thursday)	31.65	12.75	13.45	17.78	24.36
2016.12.9(Friday)	32.72	11.82	13.25	17.46	24.75
Mean	32.35	12.24	13.45	17.57	24.39

The two-step clustering algorithm is an improved method based on the BIRCH (Balanced Iterative Reducing and Clustering Using Hierarchies) model, whose prominent characteristics include the following: (1) the similarity within objects is measured by log-likelihood distance, which is suitable for numerical variables and categorical variables; (2) a clustering feature tree (CFT), which is a kind of data storage approach, is used to improve efficiency for big data; and (3) the optimal number of clustering groups is automatically determined according to the Akaike information criterion (AIC) or the Bayesian Information Criterion (BIC). The detailed information of this algorithm is provided in the previous work [30].

3.3. Classification Results. Approximately 6.79 million cardholders traveled by the Beijing subway in a month, and approximately 2.9 million passengers traveled in a day. Through data preprocessing and calculation, the travel indexes shown above can be obtained for each cardholder. The powerful statistical tool IBM SPSS Modeler is used to segment passengers and has higher computational efficiency for big data than other statistical tools. However, it is still impossible to cluster all cardholders in a month due to the large numbers. Here, two methods are used to improve the clustering efficiency. (1) The first method is to reduce the number of indexes and determine the key indexes. The "Feature Selection" module in SPSS Modeler provides a feasible method to achieve dimension reduction [29]. The Pearson correlation coefficient method in the "Feature Selection" module is used to reduce the number of variables, and the results show that the average trip distance can be removed due to its low degree of importance. Then, the number of indexes for classification is reduced to seven. (2) The second method is sample reduction. Considering the regularity and repeatability of subway travelers on different days, cardholders on a single day are sufficient to represent the characteristics of all travelers. All cardholders on June 12, 2017 (Wednesday), are chosen as our samples, with a total of 2,951,497 passengers.

The range of the number of groups is set to 2 to 15 for clustering, the similarity measurement within objects is the log-likelihood distance, and the clustering principle is the BIC rule. The optimal result is five clustering groups. The center point of each group is provided in Table 6.

Five consecutive weekdays are selected to validate the stability of the clustering results, as shown in Table 7. It can be seen that the differences of clustering results are small, which validates the stability of clustering.

Figure 3 shows the distributions of the classification indexes for each passenger group. It can be seen that there are great differences among different passenger groups, indicating that the indexes capture the heterogeneity of the passengers.

The objective of this clustering is to analyze the travel characteristics of similar passengers. According to the variable distributions of each clustering group and the travel characteristics of subway passengers, we attempt to identify the passenger types that correspond to each clustering group to better understand passengers' behaviors.

(1) Group 1 (Regular Commuting Travelers). The passengers in group 1, who travel 4.64 days per week and 2 trips per day on average, have the most distinctive characteristics. In addition, the standard deviations of their first-trip departure times and OD coverage rates are small, indicating that they

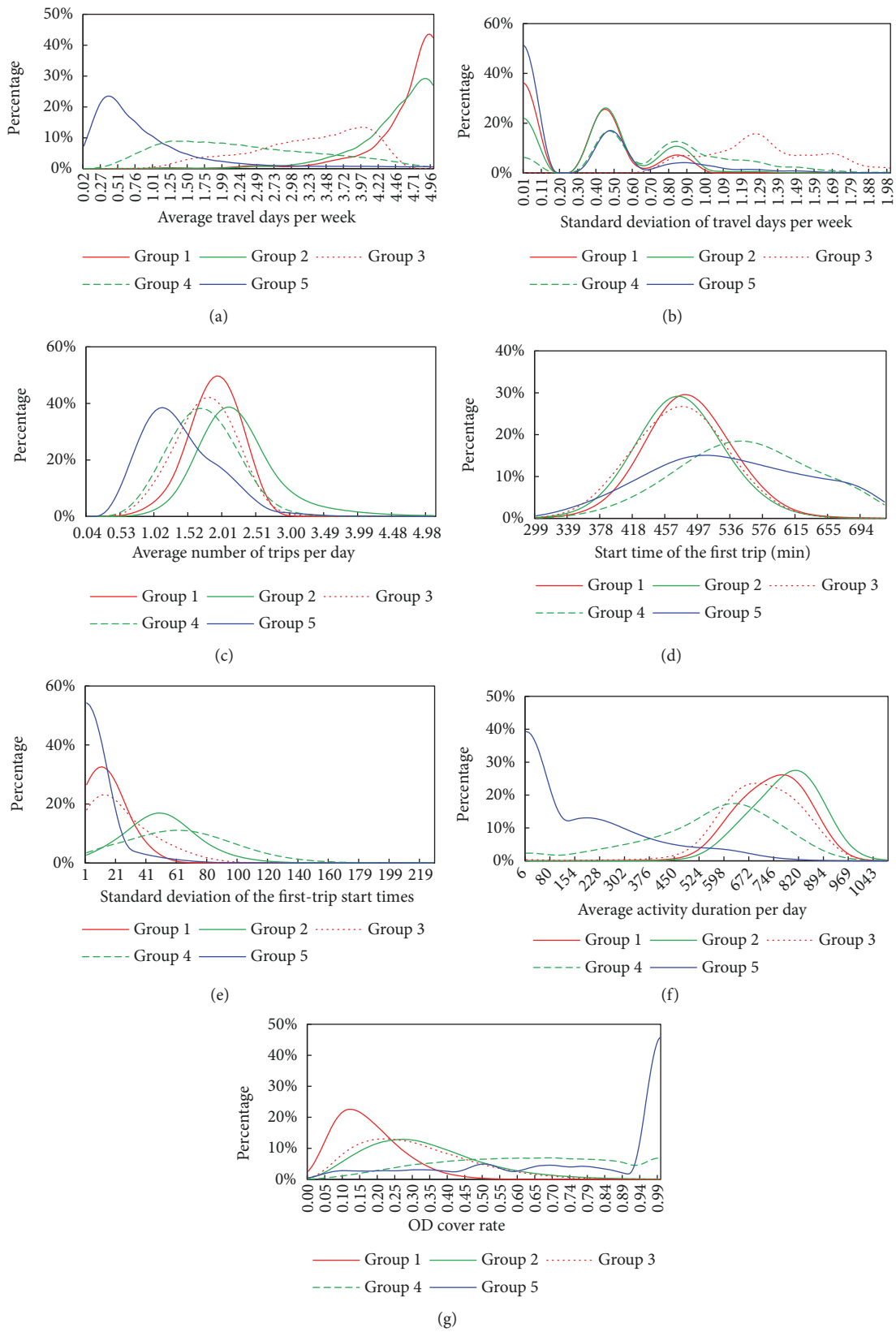


Figure 3: Distributions of the classification indexes for different passenger groups.

travel steadily in time and space (OD). These passengers are defined as regular commuting travelers.

(2) Group 2 (Working Travelers). The travelers in group 2 are similar to those in group 1, but there are slight differences. Group 2 has a larger standard deviation of the first-trip start times (49.99 minutes), indicating that these passengers may have flexible working times. Moreover, they have the largest trip counts per day (2.25), meaning they have high dependence on the subway. These passengers are defined as working travelers who travel not only for commuting but also for business.

(3) Group 3 (Frequent Travelers). The travel days per week for passengers in group 3 are smaller than those of the two prior groups, and the standard deviation of the travel days is widely distributed (Figure 3(b)). The temporal stability is weak, but the travel OD stability is strong (Figure 3(g)). These passengers can be inferred to be frequent travelers.

(4) Group 4 (Low-Frequency Travelers). Group 4 has a dispersed distribution on travel days, is unstable for the first-trip start times, and has a high OD coverage rate, which indicate that they use rail transit only for noncommuting trips. These passengers are considered low-frequency travelers whose travel purposes are daily living trips, such as visiting friends, entertainment, and shopping.

(5) Group 5 (Occasional Travelers). Compared to the other groups, group 5 has a low travel frequency in a day and a week and a large OD coverage rate and standard deviation of the first-trip start times. There is great randomness and variety for these travelers. Hence, the passengers of group 5 are defined as occasional travelers.

4. Retiming Response to Discount Price

Take the prepeak discount fare policy of Beijing subway in the year of 2017 as background. Firstly, this section introduces how to identify the potentially influenced passengers and departure time shifted travelers for the discount fare. Then, combining with the results of passenger classification in the above section, retiming elasticities for different passenger groups are measured and studied in depth.

Normally, passengers' responses to fare changes are indicated by demand change, and the fare elasticity of demand is used to capture demand changes to fare changes in transportation systems. However, the motivation in this study is not to examine the demand change but instead the passengers' departure time rescheduling behaviors due to the discount fare. The fare elasticity of departure time is defined to describe the sensitivity of passengers to shift their departure times due to fare changes and is named the retiming elasticity.

Note that the departure time in this work means the check-in time of passengers traveling by subway, not the actual departure time of passengers starting from their actual origin. Considering that the discount price measure in the Beijing subway is only valid in the early morning (before 7:00 a.m.), the retiming elasticity can be understood as the response of passengers by moving forward their first-trip departure times.

4.1. Influenced Passenger Detection. Because the discount price measure in the Beijing subway is only implemented in certain stations and valid in a short time window, only a small portion of travelers could be affected by the policy. Hence, identifying the potentially influenced passengers (target consumers of the policy) is the first step for further analysis. Three aspects are considered for detecting the potentially influenced passengers:

(a) *Passengers should travel from the target stations with discount prices.* In this case, only the 8 stations located on Line 6 are the target stations. Passengers who did not travel from these stations were excluded along with their corresponding smart card data.

(b) *Passengers should always start their first trip during peak hours.* The target consumers of the discount price are travelers during peak hours. The first-trip expected departure time can be utilized to check whether a passenger always travels during peak hours. According to the work of Peer et al., the median value of the departure times over a period can be regarded approximately as the expected departure time [31, 32].

(c) *Passengers should be residents who live around the target stations.* The discount price mostly influences the residents who live around the target stations, not occasional travelers who travel from those stations. Zou et al. [25] and Barry et al. [33] show that the first-trip origin station in a day is usually located near the home locations of passengers. The detection algorithm proposed in the work of Zou et al. [25] is used to identify the home location for each passenger, and then the residents who live around the target stations can be obtained.

Figure 4 shows the process for extracting the potentially influenced passengers (the smart card ID) from all travelers. Two scenarios of different time ranges after policy application are studied in this work, including the short term (one month later) and the middle term (six months later). Figure 5 provides the number of potentially influenced passengers. To analyze changes in passengers' travel behaviors, the selected passengers should travel in both months. Note that the selected passengers are not all influenced passengers because some passengers may lose their cards or change their home location during the study period. However, the selected sample covers almost 85% of all travel demand, and it is sufficient for our study. It can be found that there is a small increase in the influenced passengers in the middle term, which may be ascribed to seasonal influence.

4.2. Shifted Passenger Identification. Shifted passengers are travelers who change or move forward their departure times. The expected departure time (EDT), which describes the desired time that passengers want to travel, is utilized to

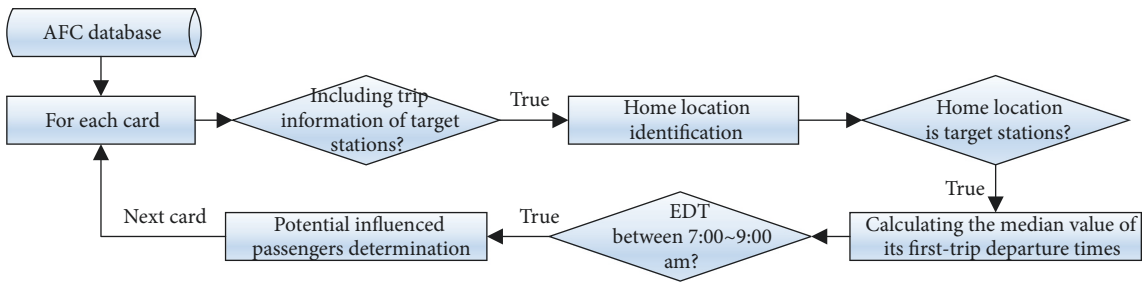


Figure 4: Process for extracting potentially influenced passengers due to the discount price.

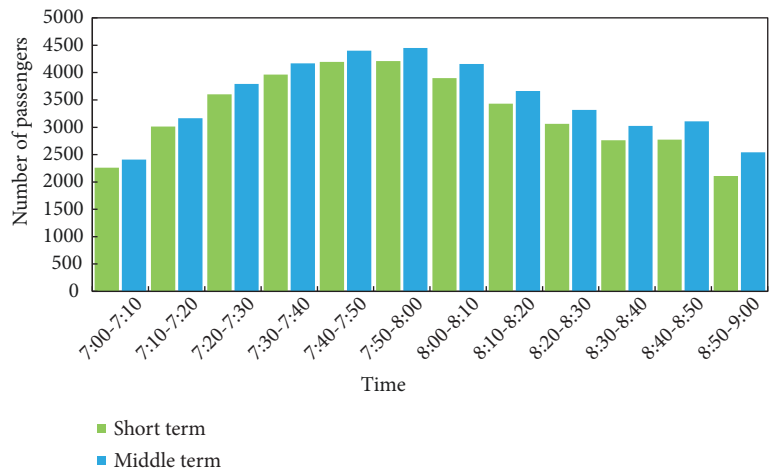


Figure 5: Distribution of potentially influenced passengers in the short term and middle term.

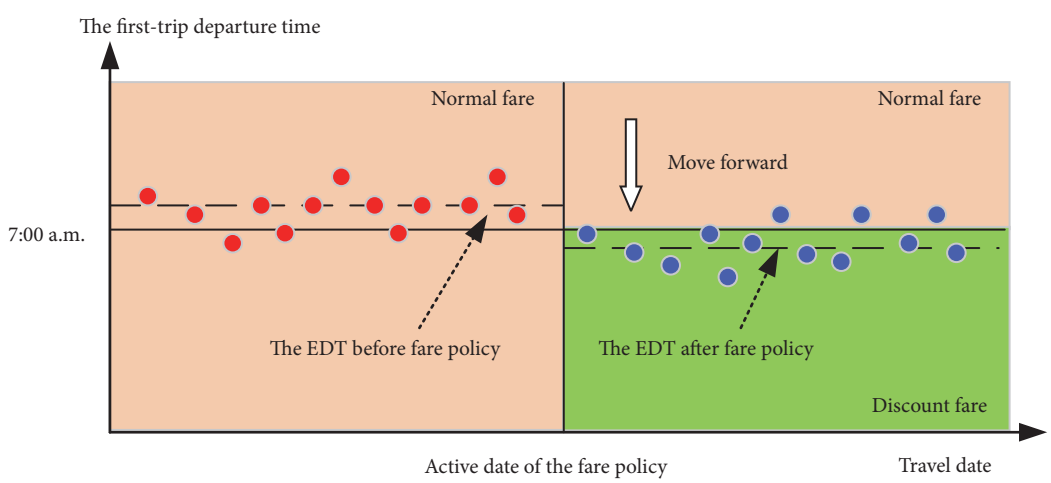


Figure 6: Schematic diagram of shifted passengers identification.

check whether a passenger is shifted. If the first-trip EDT of a passenger falls into the peak hours before the policy and shifts to the prepeak hours afterwards, then he/she can be regarded as a shifted passenger. The schematic diagram for identifying shifted passenger is provided in Figure 6. However, it is impossible to know the actual EDT for passengers. Referring to the work of Peer et al., the median value of departure times over a period can be regarded approximately as the EDT [32].

A shifted passenger is detected by the rule of

$$\delta_i = \begin{cases} 1, & \text{if } t_i < 7{:}00 \text{ a.m. and } t_i' \geq 7{:}00 \text{ a.m.} \\ 0, & \text{else} \end{cases} \quad (1)$$

where δ_i=1 means passenger (card ID), i is a shifted passenger, and t_i' and t_i are the corresponding EDT before and after fare changes, respectively.

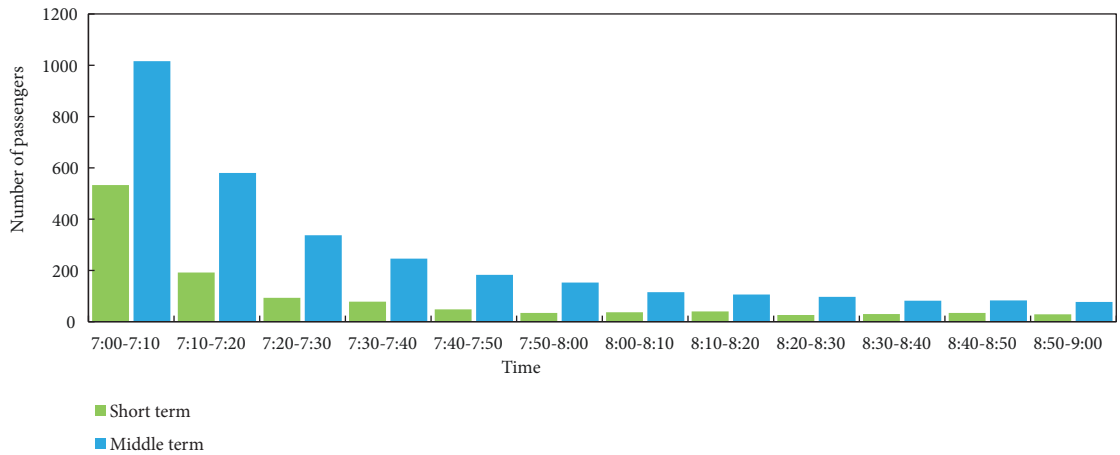


Figure 7: Number of shifted passengers in the short term and middle term.

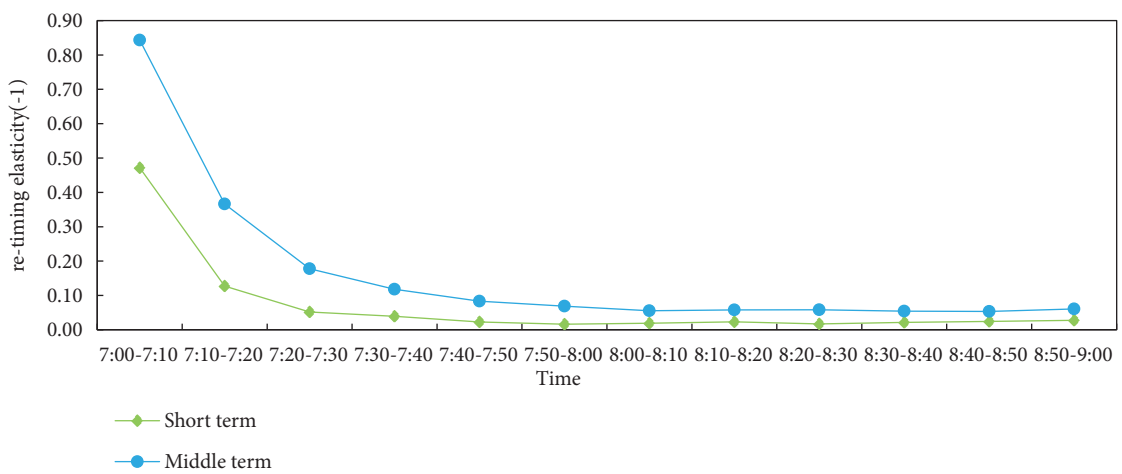


Figure 8: Time-varying retiming elasticities of all passengers.

Figure 7 shows the time-dependent distribution of shifted passengers. It can be concluded that (1) passengers with an EDT of 7:00~7:10 a.m. are mostly likely to travel earlier for the discount; (2) the number of shifted passengers decreases greatly with time going on (shifted time rising), which conforms with our intuition on travel time choices; (3) approximately 90% of these passengers travel of 7:00~8:00 a.m., which means that the effective time range of the fare policy is very limited; and (4) there is an obvious increase in shifted passengers in the middle term compared with the short term, indicating that time is an important factor in the effects of the fare policy.

4.3. Retiming Elasticity. Similar to the definition of the fare elasticity of demand, the retiming elasticity can be defined as the percentage change in ridership in response to a one percent change in fare. However, this metric uses a different measure of ridership, which is not the volume of demand but the number of passengers who change their departure times. By comparing the EDT of each passenger before and after the fare policy implementation, the shifted passengers can be determined. Then, the retiming elasticity can be formulated as a point elasticity, as shown in

$$\lambda = \frac{(d_1 - d_0)/d_0}{(f_1 - f_0)/f_0} \tag{2}$$

where λ is the retiming elasticity; d_0 is the ridership before the fare change, and d_1 is that after the change; and f_0 and f_1 are the corresponding fares. Here, the change in ridership is equal to the number of shifted passengers, and the fare change is 50%.

Figure 8 shows the time-dependent retiming elasticities without considering passenger types. The following can be observed: (1) Passengers travelling between 7:00 a.m. and 7:20 a.m. are more willing to change their departure times for the discount, but the elasticity decreases greatly with time (increase in shifted time). (2) It is difficult for passengers to depart 30 minutes earlier for the discount (the elasticity falls to -0.1); supposing that -0.1 is the lower bound of the elasticity to determine the acceptable transfer time, 30 minutes can be regarded as the maximal time that passengers can accept. (3) The retiming elasticity in the middle term is larger than that in

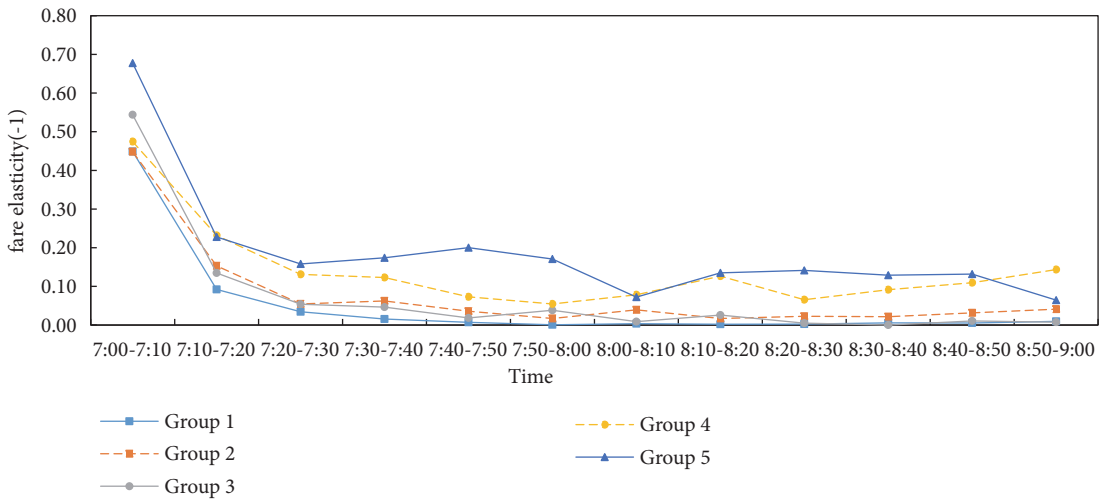


Figure 9: Distribution of retiming elasticities of different passenger groups in the short term.

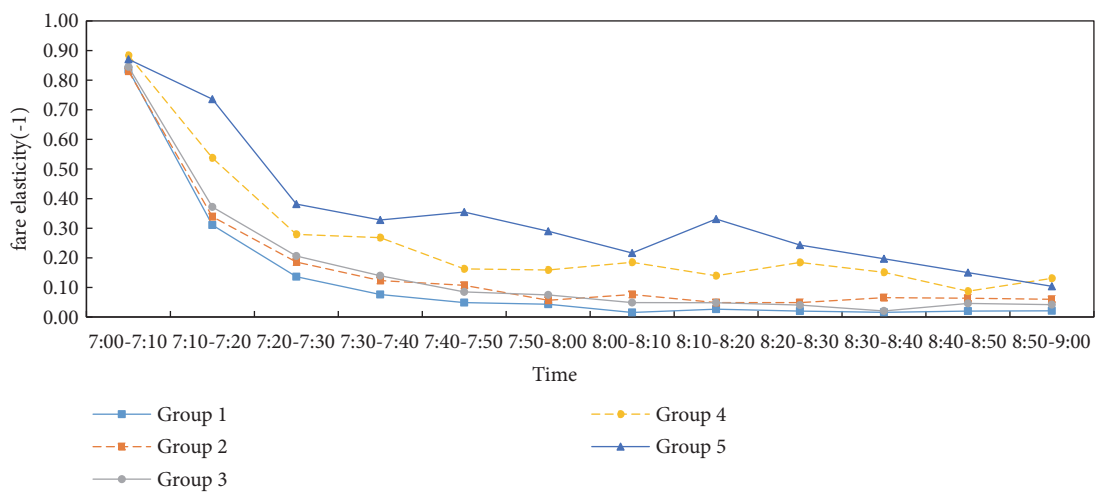


Figure 10: Distribution of retiming elasticities of different passenger groups in the middle term.

the short term, indicating that time is needed for passengers to adjust their travel behaviors.

4.3.1. Retiming Elasticities of Passenger Groups. Different types of passengers have diverse travel and personal characteristics and tend to have discrepant sensitivities to fare changes. Combined with the classification results, the retiming elasticities of different passenger groups can be explored at a finer resolution. The center point of each passenger group has been obtained in Section 3. Then, passenger type of a cardholder can be determined through the smallest Euclidean distance to center clustering points, as shown in

$$d_{ij} = \sqrt[2]{\sum_{k=1}^{K} \left(x_{ik} - \overline{x}_{jk}\right)^2} \tag{3}$$

where d_{ij} is Euclidean distance of passenger i to the clustering center point j; K is the set of classification indexes; x_{ik} is the value of index k for passenger i; and $\overline{x}_{jk}$ is corresponding value of center point j.

Figure 9 provides the retiming elasticities of different passenger groups in the short term. The following can be observed: (1) The retiming elasticity of all passenger types decreases greatly with time going on. (2) Passenger group 1 (regular commuters) has the minimum elasticity to change their departure times, while the elasticity of groups 4 and 5 are much larger than that of group 1. Passenger groups 1, 2, and 3 belong to high-frequency travelers, and groups 4 and 5 are low-frequency passengers. The retiming elasticity of low-frequency passengers is much greater than that of high frequency passengers. (3) For low-frequency passengers, the retiming elasticity remains at a relatively high level throughout the entire peak period, indicating that time is not an important factor in their departure time choice but fare is.

Figure 10 shows the distribution of retiming elasticities in the middle term, which has a similar trend as that of the short term. It can be found that the elasticities are obviously

Table 8: ATE of different passenger types in different time ranges.

Passenger type	ATE in the short term	ATE in the middle term	Change in ATE
Group 1	-0.10	-0.24	-0.14
Group 2	-0.13	-0.27	-0.14
Group 3	-0.14	-0.29	-0.15
Group 4	-0.18	-0.38	-0.20
Group 5	-0.27	-0.49	-0.22

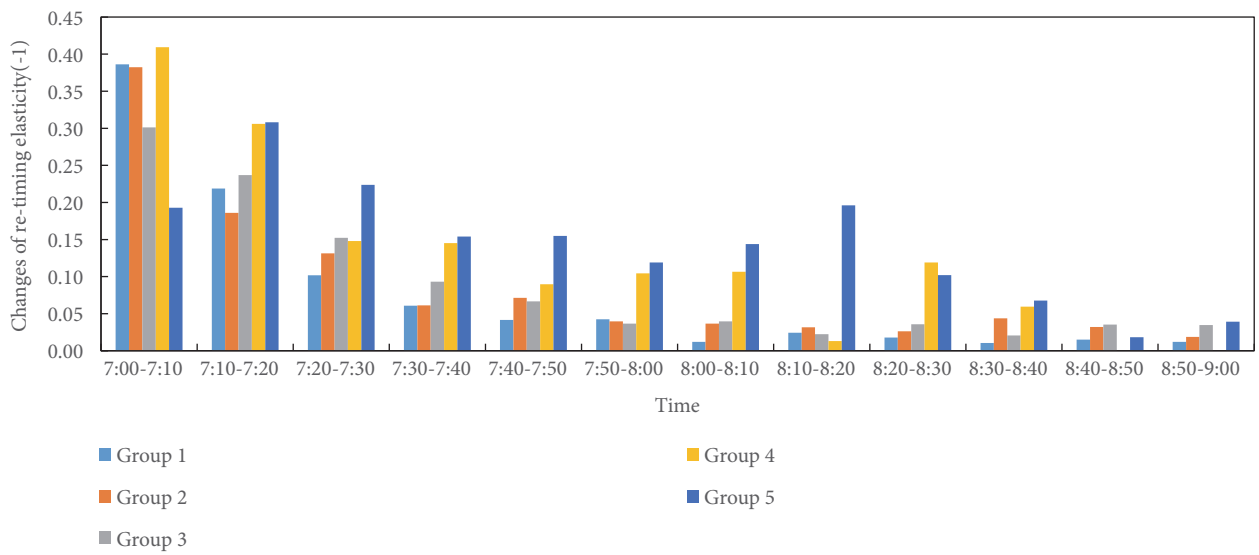


Figure 11: Differences in retiming elasticity between the short term and the middle term.

larger than those in the short term. The changes in retiming elasticity over time will be analyzed in the next section.

In summary, different types of passengers have substantial differences in their retiming elasticities for fare changes. Thus, the travel demand structure of stations can greatly influence the effect of the price policy. The demand at a station with a higher percentage of low-frequency passengers is influenced more by the discount fare because more passengers tend to shift their travel time.

4.3.2. Retiming Elasticity over Time. Passengers usually need a certain time to adjust their travel behaviors after a new policy implementation. This section explores the retiming response of passengers in different time ranges (the short term and the middle term).

Table 8 shows the average retiming elasticity (ATE) of passengers between 7:00 and 8:00 a.m. The following can be observed: (1) The elasticity in the middle term is obviously larger than that in the short term, indicating that time is a considerable factor in the effect of the policy. (2) Changes in the ATE of high-frequency travelers (groups 1, 2, and 3) are smaller than those of low-frequency travelers (groups 4 and 5), indicating that it is still difficult to change the behaviors of high-frequency travelers even after the policy is implemented for a certain amount of time.

Figure 11 provides the differences in retiming elasticity of different passenger groups between the short term and the middle term. It can be clearly seen that (1) passengers who travel between 7:00 a.m. and 7:30 a.m. have a greater increase in retiming elasticity, while there is limited change for passengers who travel later than 7:30 a.m., and (2) low-frequency passengers (groups 4 and 5) are more sensitive to the time than high-frequency travelers.

4.3.3. Retiming Elasticity of Shifted Time. It can be seen that the retiming elasticity of passengers strongly depends on the time span that passengers need to move up (shifted time) and is inversely related to the moving-up time. It is easy to understand that passengers always try to balance travel time costs and fare savings. In this section, the relationship between retiming elasticity and shifted time is studied by a data fitting method. Assume that the retiming elasticity of the shifted time satisfies a negative exponential distribution, as shown in

$$\lambda(\Delta t) = -ae^{\beta \cdot \Delta t} \tag{4}$$

where $\lambda(\Delta t)$ is the elasticity when the moving-up time is Δt, and α and β are parameters to be estimated.

Table 9 provides the fitting results of the elasticity for different types of passengers. The goodness of fit verifies our assumption.

For a passenger, the retiming elasticity can be understood as the probability that a passenger would like to change his/her departure time. Suppose the EDT of a passenger is

Table 9: Fitting results of retiming elasticity for different types of passengers.

Time range	Passenger Type	α	β	Goodness of fit R^2
The short term	Group 1	0.6145	-0.1144	0.9698
	Group 2	0.2827	-0.0588	0.9192
	Group 3	0.2474	-0.0555	0.8657
	Group 4	0.2904	-0.0410	0.9554
	Group 5	0.2350	-0.0205	0.8934
The middle term	Group 1	0.5422	-0.0598	0.9386
	Group 2	0.5196	-0.0494	0.9524
	Group 3	0.5342	-0.0484	0.9556
	Group 4	0.5283	-0.0349	0.9273
	Group 5	0.4921	-0.0224	0.8323

known, and the valid time of the discount fare is fixed; then, the departure time change probability can be calculated by the retiming elasticity function.

For a station, once the travel demand structure and time-dependent distribution of demand are determined, the number of passengers willing to change departure time can also be estimated. The passenger type and expected departure time of each cardholder can be easily obtained by collecting smart card history records; then, the number of such passengers can be estimated by

$$M = \sum_{n=1}^{N} \Delta f \cdot \lambda_n = \sum_{n=1}^{N} \Delta f \cdot \left(-a_n e^{\beta_n \cdot \Delta t_n}\right) \tag{5}$$

where M is the number of shifted passengers; n means passenger ID, and N is the total number of potentially influenced passengers; Δf is the percentage change of the fare discount; and α_n, β_n, and Δt_n are the corresponding parameters for passenger n (type).

5. Applications and Suggestions

In the section, two simple applications of retiming elasticity for optimizing discount price are presented. The first is to choose potential target discount stations by considering station demand characteristics, and the second is to analyze the valid time window for the effect of discount fares. Then, discussions and suggestions for improving the effect of the price policy are provided.

5.1. Potential Target Discount Station Selection. Suppose the percentage of the discount (50%) and the valid time window (before 7:00 a.m.) remain unchanged the same as the actual fare policy. This case aims to find potential discount stations with a large number of passengers being shifted by the discount. According to the retiming elasticity of different passenger groups, low-frequency passengers are more sensitive to fare changes than high-frequency passengers. Hence, the composition of station demand (flow) is important for the effect of the fare policy.

The total smart card records in June, 2017, are used. Classification indexes of each cardholder are computed from the records in a month, and then passenger type of each passenger can be determined through the smallest Euclidean space distance to the clustering center points. Also the first-trip EDT of each passenger can be gained. A representative day of June 12, 2017 (Wednesday), is selected to analyze the shifted passengers, which means only cardholders who travelled in the day are target passengers.

The potentially influenced passengers for the discount price during the morning peak hours (7:00~9:00 a.m.) can be collected through the process shown in Figure 4, and the results are shown in Figure 12. The number of shifted passengers for each station can be computed according to (5), in which Δf is -50% and the shifted time Δt for a passenger is the time span between his/her first-trip EDT and the end time of discount (7:00 a.m.).

Figure 13 shows the top 50 stations with the largest number of shifted passengers. These stations can be chosen as potential target stations for the next price scheme. However, it should be noted that peak demand shift is not the purpose of the price policy; the major goal is to relieve congestion. There are slight differences between the two. Stations with large numbers of shifted passengers may not have a substantial effect on reducing congestion because capacity supply is another important factor for congestion issues. Thus, capacity bottlenecks should also be considered when establishing a new pricing scheme. How to optimize fare schemes from the perspective of relieving system congestion will be provided in our future work.

5.2. Valid Time Analysis for the Effect of Fare Policy. For a prepeak discount fare policy, the valid time depends on the boundary (end) time. In the Beijing subway, all participating stations have the same end time (7:00 a.m.) of the discount. This is unreasonable because there are great differences in the temporal characteristics of station demand. Usually, the peak time of demand for suburban stations is much earlier than for urban stations, so different valid times for stations may perform better than a constant time.

Assume that the percentage of the discount remains the same (50%), while the end time of the discount can be changed. The same smart card data as in the case above is used. Figure 14 shows the demand distributions of two typical stations in the Beijing subway. CY (CHANG YANG) station on Line FS is chosen as a suburban station, while XZM (XI

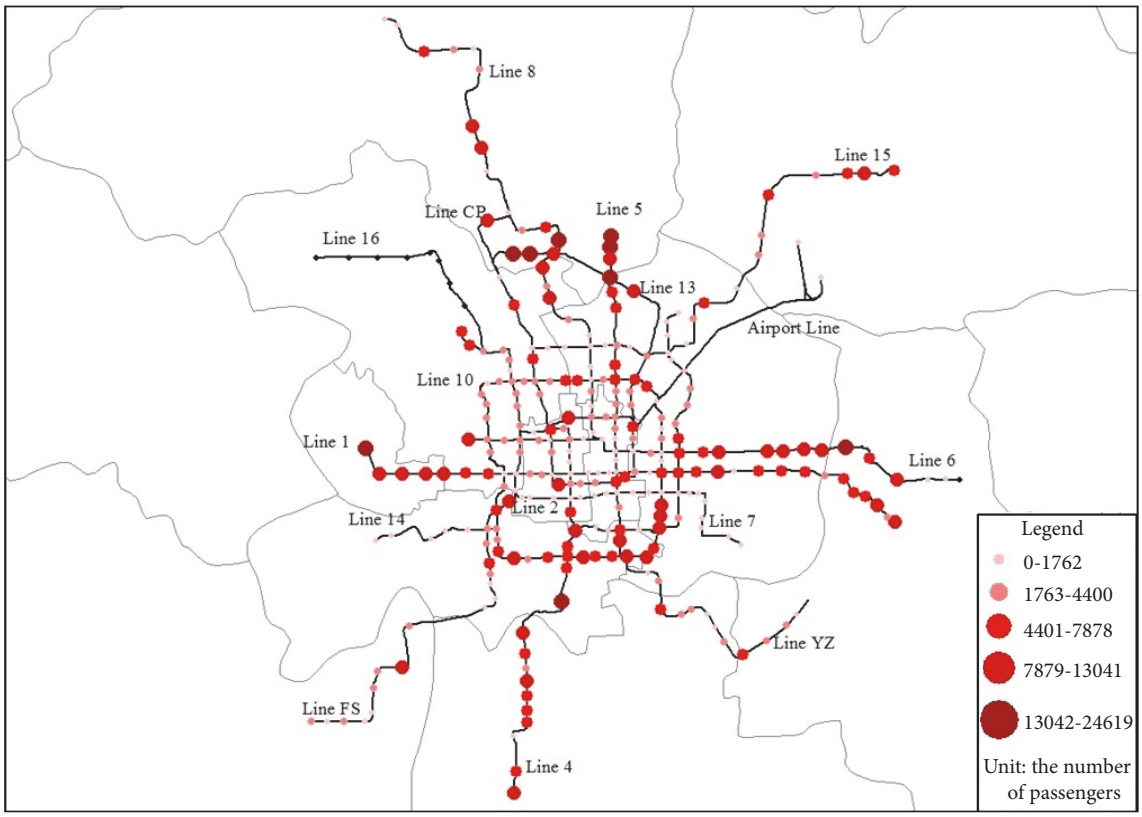


Figure 12: Distribution of potentially influenced passengers for the discount fare.

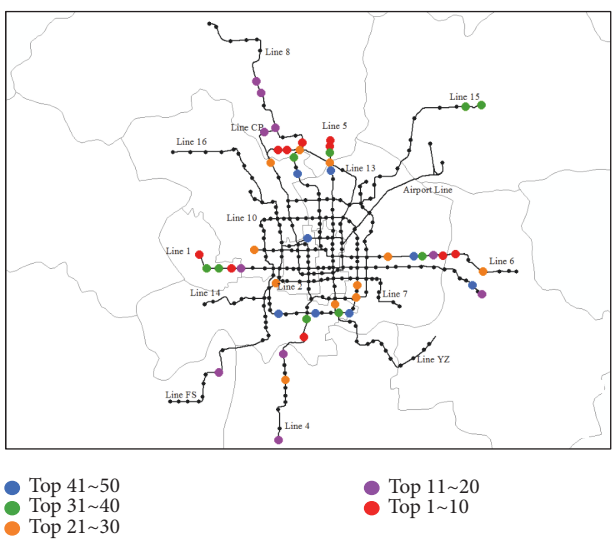


Figure 13: The top 50 stations with the largest number of shifted passengers.

ZHI MEN) station located on Line 2 is selected as a city center station. It can be clearly seen that the temporal distribution of demand differs distinctly, and the peak time of CY station is much earlier than that of XZM station.

Using the retiming elasticity functions and the detailed demand data, the number of shifted passengers can be estimated according to (5). Figure 15 shows the demand changes before and after the 50% discount price with the end time of 7:00 a.m. The effect of the discount price for the CY station is much greater than that for the XZM station for peak demand shift.

Taking XZM station as a case to further observe the influence of differential valid time for discount price, Figure 16 shows the performance of the discount price with differential

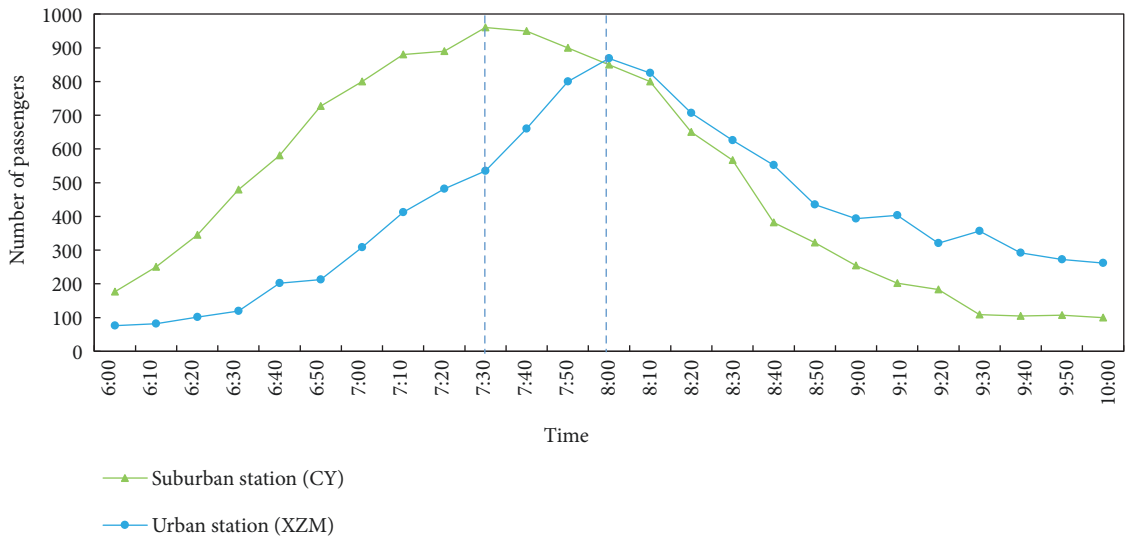


Figure 14: Distribution of station demand of two typical stations.

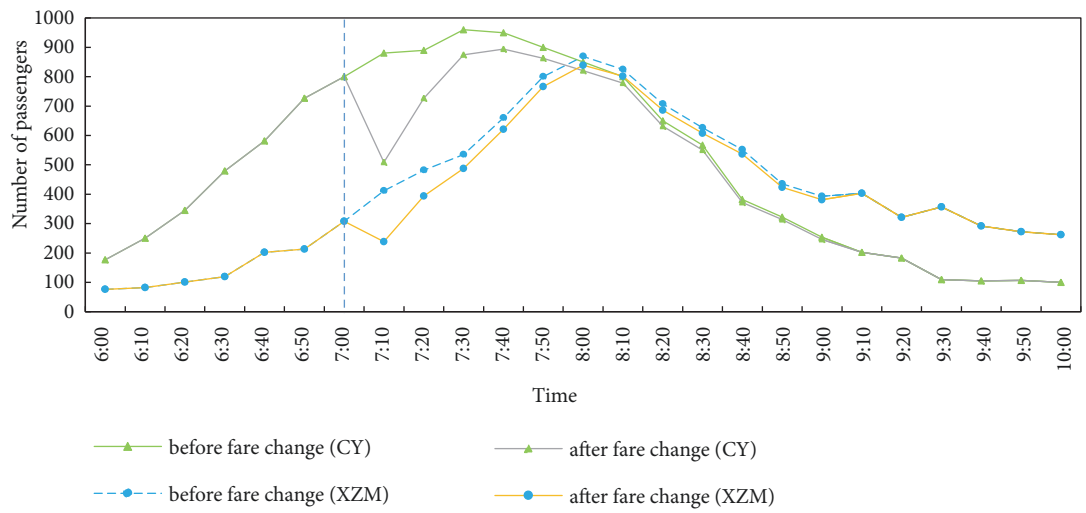


Figure 15: Estimated changes in peak demand after the 50% discount fare policy with the end time of 7:00 a.m.

end times. We can see that the effect of the discount price is strongly related to the end time of the discount price. The main reason is that the effective time range of the discount fare is very limited (almost less than 30 minutes). Hence, a proper valid time window of the discount has a significant impact on the effect of the policy.

However, it will be inconvenient for passengers if each station has a different valid time of the fare discount. In practice, the end time of the discount price can be determined region by region or line by line. Moreover, optimizations of valid time for the discount price should also consider the congestion and transport capacity in the view of network systems. It is interesting and complex work and will be addressed in our future works.

5.3. Discussion and Suggestions. According to the analysis results above, the discount price policy can indeed shift some passengers from peak periods to prepeak time. However, policy makers are still acting conservatively with this strategy and did not further implement it on a larger scale or on the whole network. The main reason may be that no significant reduction of peak demand can be observed in the pilot implementation, and the effect of congestion relief is limited (as shown in Section 2.2). How to improve the effect of the fare strategy? Three aspects of this issue are discussed as follows.

(a) *Change fare structures.* There are a variety of differential pricing structures (peak charge, ladder pricing, prepeak discount and peak congestion charge in combination, etc.). Undoubtedly, peak congestion charges could shift peak demand forward or backward. However, passengers may shift to bus or other road traffic modes. It is not a wise choice to ease subway congestion at the cost of increasing already serious congestion in other traffic modes.

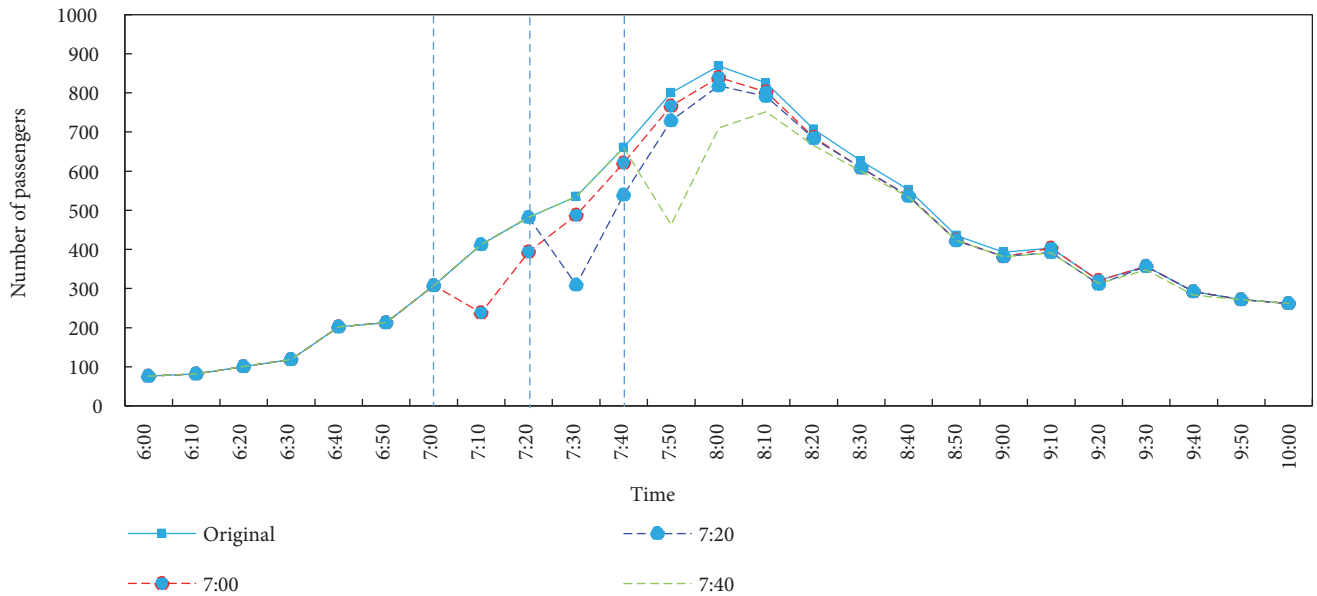


Figure 16: Estimated changes in demand for XZM station with different end times of the discount price.

(b) *Increase the percentage of the discount*. Usually, the greater the fare discount, the greater the demand changes. The most special and extreme situation is to offer free fares. The Melbourne subway has implemented a free fare strategy (early bird ticket) during morning peak hours and achieved good performance [34]. The Beijing subway can further try to increase the percentage of the discount fares but must consider revenue loss. Currently, the Beijing government provides large subsidies for public transit, including 50% of the operation costs for the subway and 70% for buses. Increasing revenue is another prominent issue for the Beijing subway, and revenue loss should be taken into account when providing a higher discount.

(c) *Optimize pricing schemes*. As the retiming elasticity of different passenger types is discrepant, the demand structure can be an important factor to optimize pricing schemes and thus select more suitable discount stations. Moreover, there is a maximum threshold of the influence time range of the discount fare policy (approximately 30 minutes). A case study also confirms that optimizing the valid time of the discount fare has a significant impact on reducing peak demand. Hence, valid time window optimization can be regarded as a feasible measure to improve the effect of fare policies.

6. Conclusions

An analysis framework is proposed to measure the retiming responses of travelers changing departure time due to a prepeak discount pricing strategy using a new data source of smart card transaction records. This work proves that smart card data have great potential and value for the assessment of traffic policies because of their traceable characteristics. The unique ID of the smart card makes it possible to trace behavior changes at the level of an individual passenger over a long time period, and the actual trip information gained from smart card data improves the accuracy of travel demand and behavior research. However, the drawbacks of smart card data are also quite demonstrable. Because the smart card is always anonymous, trip information cannot be linked to the personal and socioeconomic characteristics of card holders, which limits the potential to mine original and comprehensive reasons for behavior changes.

Some case-specific conclusions are also drawn. The detailed retiming elasticity of Beijing subway passengers to the discount fare is measured. Considering the heterogeneity of passengers, an unsupervised clustering approach is applied to segment passengers into groups, and the retiming elasticities of different passenger types are also obtained. The results show that low-frequency travelers are more sensitive to fare discounts than high-frequency travelers, and they are more willing to shift their departure time for the discount. It is found that the retiming elasticity decreases greatly with increasing shifted time, and it is difficult to change passengers' departure times by more than 30 minutes for the discount. Furthermore, the retiming elasticity functions of shifted time are obtained through a data fitting approach, which provides strong support for estimating demand changes according to the temporal distribution of station demand. Applications of fare scheme optimization confirm that the valid time window of the discount has a significant influence on the effect of discount fare policies. We suggest that optimizing the valid time window of the discount price could be a useful and feasible way to improve the policy effect, while changing fare structures and increasing discounts are not a wise choice for the Beijing subway.

Passengers' responses to fare changes are very complex and are related to various external factors. This work only studies the travel responses of Beijing subway travelers, and it

should be careful to apply these absolute values of elasticity to other transport systems. For further research, first, retiming responses of passengers in the long term (one year later or longer) could be explored. Second, based on the results of retiming elasticity, optimizing fare schemes to relieve congestion on the network is our final goal. That work is complicated and interesting and will be addressed in our future works.

Data Availability

The origin and detailed smart card data [Dump File (.dmp)] used to support the findings of this study were supplied by [Beijing Mass Transit Railway Operation Corporation] under license and cannot be made freely available, for possible issues of security and personal privacy. Requests for access to these origin data should be made to [Fei Dou, Email: heishenanhai@163.com].

Conflicts of Interest

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

Acknowledgments

The authors would like to acknowledge the Beijing subway for data support. This research is supported by the National Natural Science Foundation of China (71701011).

References

[1] F. Devillaine, M. A. Munizaga, and M. Trépanier, "Detection activities of public transport users by analyzing smart card data," *Transportation Research Record*, vol. 2276, pp. 48–55, 2012.

[2] M. Bagchi and P. R. White, "The potential of public transport smart card data," *Transport Policy*, vol. 12, no. 5, pp. 464–474, 2005.

[3] G. Bresson, J. Dargay, J.-L. Madre, and A. Pirotte, "The main determinant of the demand for public transport: a comparative analysis of England and France using shrinkage estimators," *Transportation Research Part A: Policy and Practice*, vol. 37, no. 7, pp. 605–627, 2003.

[4] B. E. Mccollom and R. H. Pratt, "Travel response to transportation system changes: chapter 12—transit pricing and fares," Transit Cooperation Research Program Report 95, Transportation Research Board, Washington, Wash, USA, 2004.

[5] Association of Train Operating Companies (ATOC), *Passenger Demand Forecasting Handbook*, Passenger Demand Forecasting Council, London, UK, 4th edition, 2002.

[6] J. M. Dargay and M. Hanly, "The demand for local bus services in England," *Journal of Transport Economics and Policy*, vol. 36, no. 1, pp. 73–91, 2002.

[7] D. A. Hensher, "Assessing systematic sources of variation in public transport elasticities: Some comparative warnings," *Transportation Research Part A: Policy and Practice*, vol. 42, no. 7, pp. 1031–1042, 2008.

[8] T. Litman, "Transit price elasticities and cross-elasticities," *Journal of Public Transportation*, vol. 7, no. 2, pp. 37–58, 2004.

[9] J. Holmgren, "Meta-analysis of public transport demand," *Transportation Research Part A: Policy and Practice*, vol. 41, no. 10, pp. 1021–1035, 2007.

[10] P. Schimek, "Dynamic estimates of fare elasticity for US public transit," *Transportation Research Record Journal of the Transportation Research Board*, vol. 2538, pp. 96–101, 2015.

[11] N. Paulley, R. Balcombe, R. Mackett et al., "The demand for public transport: the effects of fares, quality of service, income and car ownership," *Transport Policy*, vol. 13, no. 4, pp. 295–306, 2006.

[12] Z.-J. Wang, X.-H. Li, and F. Chen, "Impact evaluation of a mass transit fare change on demand and revenue utilizing smart card data," *Transportation Research Part A: Policy and Practice*, vol. 77, pp. 213–224, 2015.

[13] Z.-J. Wang, F. Chen, B. Wang, and J.-L. Huang, "Passengers' response to transit fare change: an ex post appraisal using smart card data," *Transportation*, vol. 45, no. 2, pp. 1–20, 2017.

[14] H. Mahmassani, X. Zhou, and C. Lu, "Toll pricing and heterogeneous users: approximation algorithms for finding bi-criterion time-dependent efficient paths in large-scale traffic networks," *Transportation Research Record*, vol. 1923, pp. 28–36, 2005.

[15] C. Lu, X. Zhou, and H. Mahmassani, "Variable toll pricing and heterogeneous users," *Transportation Research Record*, vol. 1964, pp. 19–26, 2006.

[16] C. C. Lu, H. S. Mahmassani, and X. S. Zhou, "A bi-criterion dynamic user equilibrium traffic assignment model and solution algorithm for evaluating dynamic road pricing strategies," *Transportation Research Part C: Emerging Technologies*, vol. 16, no. 4, pp. 371–389, 2008.

[17] C. Lu and H. S. Mahmassani, "Modeling user responses to pricing," *Transportation Research Record*, vol. 2085, no. 1, pp. 124–135, 2008.

[18] A. Aboudina, H. Abdelgawad, B. Abdulhai, and K. N. Habib, "Time-dependent congestion pricing system for large networks: Integrating departure time choice, dynamic traffic assignment and regional travel surveys in the Greater Toronto Area," *Transportation Research Part A: Policy and Practice*, vol. 94, pp. 411–430, 2016.

[19] K. M. Nurul Habib, N. Day, and E. J. Miller, "An investigation of commuting trip timing and mode choice in the Greater Toronto Area: Application of a joint discrete-continuous model," *Transportation Research Part A: Policy and Practice*, vol. 43, no. 7, pp. 639–653, 2009.

[20] C.-C. Lu and H. S. Mahmassani, "Modeling heterogeneous network user route and departure time responses to dynamic pricing," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 2, pp. 320–337, 2011.

[21] Y. Liu and P. Charles, "Spreading peak demand for urban rail transit through differential fare policy: a review of empirical evidence," in *Proceedings of the Australasian Transport Research Forum*, pp. 1–35, Brisbane, Australia, 2013, http://atrf.info/papers/2013/2013_liu_charles.pdf.

[22] M.-P. Pelletier, M. Trépanier, and C. Morency, "Smart card data use in public transit: a literature review," *Transportation Research Part C: Emerging Technologies*, vol. 19, no. 4, pp. 557–568, 2011.

[23] X. Ma, C. Liu, H. Wen, Y. Wang, and Y. Wu, "Understanding commuting patterns using transit smart card data," *Journal of Transport Geography*, vol. 58, pp. 135–145, 2017.

[24] A. Alsger, A. Tavassoli, M. Mesbah, L. Ferreira, and M. Hickman, "Public transport trip purpose inference using smart

card fare data," *Transportation Research Part C: Emerging Technologies*, vol. 87, pp. 123–137, 2018.

[25] Q. Zou, X. Yao, P. Zhao, H. Wei, and H. Ren, "Detecting home location and trip purposes for cardholders by mining smart card transaction data in Beijing subway," *Transportation*, vol. 45, no. 3, pp. 919–944, 2018.

[26] A.-S. Briand, E. Côme, M. Trépanier, and L. Oukhellou, "Analyzing year-to-year changes in public transport passenger behaviour using smart card data," *Transportation Research Part C: Emerging Technologies*, vol. 79, pp. 274–289, 2017.

[27] Research group of annual report of China urban mass transit, *Annual Report of China Urban Mass Transit (2006-2018)*, Press of Beijing jiaotong university, 2019.

[28] C.-Y. Tsai and C.-C. Chiu, "A purchase-based market segmentation methodology," *Expert Systems with Applications*, vol. 27, no. 2, pp. 265–276, 2004.

[29] C. C. Aggarwal and C. Reddy, *Data Clustering: Algorithms and Applications*, CRC Press, 2013.

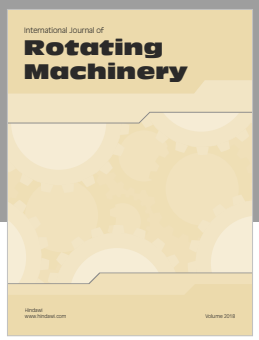
[30] L. Ding, F. M. Gonzalez-Longatt, P. Wall, and V. Terzija, "Two-step spectral clustering controlled islanding algorithm," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 75–84, 2013.

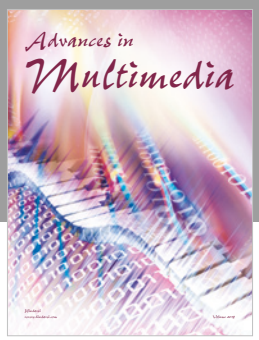
[31] S. Peer, J. Knockaert, and E. T. Verhoef, "Train commuters' scheduling preferences: Evidence from a large-scale peak avoidance experiment," *Transportation Research Part B: Methodological*, vol. 83, pp. 314–333, 2016.

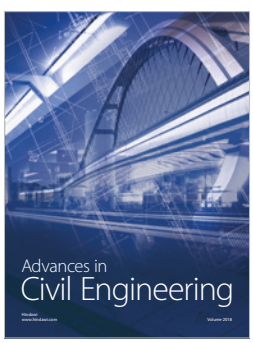
[32] S. Peer, E. Verhoef, J. Knockaert et al., "Long-run vs. short-run perspectives on consumer scheduling: evidence from a revealed-preference experiment among peak-hour road commuters," *International Economic Review*, vol. 56, no. 1, pp. 303–323, 2015.

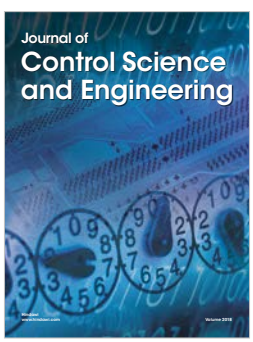
[33] J. J. Barry, R. Newhouser, A. Rahbee, and S. Sayeda, "Origin and destination estimation in New York City with automated fare system data," *Transportation Research Record*, vol. 1817, pp. 183–187, 2002.

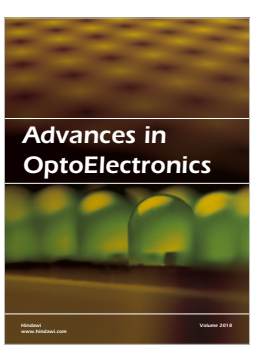
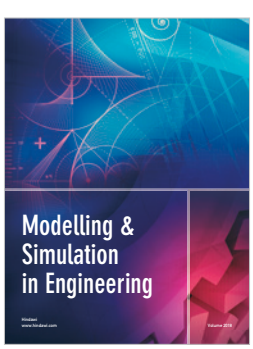
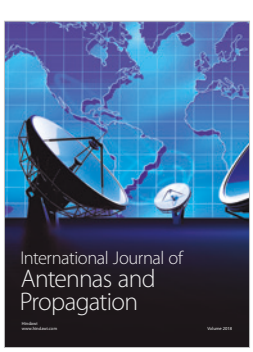
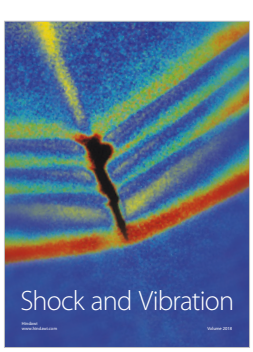
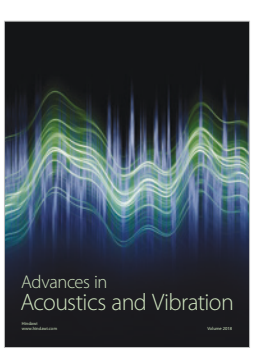
[34] G. Currie, "Quick and effective solution to rail overcrowding free early bird ticket experience in Melbourne, Australia," *Transportation Research Record*, vol. 2146, pp. 35–42, 2010.









Hindawi

Submit your manuscripts at
www.hindawi.com