Understanding the determinants of young commuters’ metro-bikeshare usage frequency using big data

Citation for published version (APA):

Document license:
TAVERNE

DOI:
10.1016/j.tbs.2020.06.007

Document status and date:
Published: 01/10/2020

Document Version:
Publisher’s PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher’s website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:
www.tue.nl/taverne

Take down policy
If you believe that this document breaches copyright please contact us at:
openaccess@tue.nl
providing details and we will investigate your claim.
Understanding the determinants of young commuters’ metro-bikeshare usage frequency using big data

Yang Liu, Yanjie Ji, Tao Feng, Harry Timmermans

1. Introduction

A large body of literature revealed that long commuting distance is the key factor creating the demand for car trips (Charypar and Nagel, 2005; Axiss et al., 2012; Beige and Axhausen, 2012; Rasouli and Timmermans, 2014; Auld et al., 2016; Zhao and Li, 2017; Liu et al., 2018a, 2018b). The reduction of travel distance not only decreases the use of cars but also improves individuals' subjective well-being (e.g., De Vos et al., 2016; Ettema and Schekkerman, 2016; Gao et al., 2017; Friman et al., 2017, 2018; Lorenz, 2018; Gan et al., 2019; De Vos, 2019). However, in China, relatively long-distance commuting is inevitable for most citizens (Wang et al., 2011; Zhu and Fan, 2018). It has become a pervasive phenomenon in Chinese cities because of changes in their physical and the urbanisation process since the economic reform in 1978. Danwei, a mixed land use pattern combining home and work-physical and the urbanisation process since the economic reform in 1978. Danwei, a mixed land use pattern combining home and work-place which basically shaped Socialist China's urban space in the Mao Era, have faded away after the reform (Li and Siu, 2001; Bray, 2005; Hsing, 2006). The diminishing influence of the Danwei compounds and the market-oriented reform have broken the balance between jobs and housing (Wang et al., 2011; Wang and Zhou, 2017). Moreover, as a result of the fast urbanisation process, the number of migrant workers that came to the major cities has dramatically increased over the past three decades (Chan and Ngai, 2009; Wong et al., 2010). They typically live at considerable distance from their workplace due to a lack of affordable housing in more central areas (Huang, 2012; Zou, 2014). Although many young migrants are not yet car owners because of economic constraints and policy interventions, they are potential car users. Therefore, to accommodate middle and long-distance travel needs and discourage car use, improving public transport services plays a primary role.

Urban metro systems have rapidly expanded in many Chinese cities. Nonetheless, the access and egress to metro is often problematic, limiting the intention to use the metro among commuters who do not have convenient access to and from metro stations. Fortunately, bike sharing systems provide a promising solution to overcome these barriers. There are two main bike-sharing systems in China, one of which is the traditional public bike with a fixed dock and the other is an innovative bike sharing system called free-floating bike-sharing (FFBS). Although
FFBS swept across many cities of China in 2016, serious problems, such as piled-up wrecked bikes and disordered parking that dominates sidewalks and other public spaces, have been widely reported (Li et al., 2018; Gu et al., 2019). After two years of excessive expansion, many companies offering FFBS services, such as Xiaoming Bike and Ofo, went bankrupt. Thus, this paper only focuses on the traditional public bike-sharing system due to its stable and sustainable development. The integration of public bike and metro can offer commuters a convenient and environment-friendly travel option. Thus, to encourage more young commuters to use bikes as a transfer mode, it is important to further understand their usage of the combined metro and public bike mode.

There are many studies on determinants of bike usage and frequency (e.g. Heinen et al., 2010; Damant-Sirois and El-Geneidy, 2015; Fu and Farber, 2017), but the usage frequency of metro-bikeshare has been rarely considered. The frequency, however, is important to capture routine behaviour of individuals. Considering that policies make sense only when the predictions of combined metro and bike use are reliable, traditional cross-sectional surveys can hardly yield useful insights about individuals’ repetitive behaviour. Therefore, multi-day smart card data which record every user’s complete travel information are useful to fill this research gap.

This study thus aims at investigating the determinants of young commuters’ usage frequency of metro-bikeshare using smart card data. It contributes to a deeper understanding of the role of bikesharing in encouraging young commuters to use public transport. Results will offer various policy recommendations to integrate metro and public bike.

In this study, aged 14–35 are considered to be youth, based on the China’s Middle-and Long-term Youth Development Plan (2016–2025) (The Chinese Government, 2017). Therefore, the youth adult/commuter refers specifically to individuals aged 18–35. The rest of the paper is organized as follows: Section 2 reviews the literature on factors influencing bike usage for metro access. Section 3 presents the data and study area. Section 4 describes the research methods and estimation results are presented in Section 5. Section 6 discusses the findings and policy implications. The conclusions and limitations of the paper are also provided in Section 6.

2. Literature review

Since the first-generation of public bike systems was launched in the Netherlands in 1965, public bike systems have been widely developed in many cities around the world (DeMaio, 2009). Although the main idea underlying the development of public bike systems relates to energy saving, environmental protection and sustainable urban development, the function of bike-sharing in transportation systems may vary across cities due to different urban forms and social contexts (Zhang and Zhang, 2018). There are two main functions for bike-sharing in urban transportation systems in consideration of the relationship with other public transportation modes (Campbell and Braakwood, 2017). One function is to encourage short-distance travellers to use bikes instead of other public transport modes (Martin and Shaheen, 2014; Murphy and Usher, 2015; Scott and Ciuro, 2019). The other is as a feeder mode to other main transport modes (Shaheen et al., 2010; Fishman et al., 2014; Noland et al., 2016; Zhao and Li, 2017). Because the aim of this paper is to better understand the role of bikesharing in young commuters’ choice of transport mode in the context of long-distance commuting, the literature reviews is only concerned with the latter function of public bikes.

In the existing literature on bike as a feeder mode to transit, most scholars explored the factors influencing the choice of bike in access/egress trips to metro, bus or train. However, most of these studies are not aimed at public bikes. For instance, Krygsman et al. (2004) investigated the determinants of cycling and walking access and egress time for bus, tram, metro and train in the Netherlands. They found bike is usually used for longer access and egress times with a faster access/egress mode. Martens (2004) analysed bike-and-ride (B&R) systems in three European countries including the Netherlands, Germany and the United Kingdom. He found that although bike cultures and infrastructures of the three countries were different, the factors influencing the choice of the bike as an access or egress mode are similar, mainly including travel purpose, travel distance to the main transit mode and car availability. Using the data from the National Household Travel Surveys, Wang and Liu (2013) analysed transit-bike integrated trips in the United States from 2001 to 2009 and the characteristics of users. They found that commuters, younger travellers and individuals living in large and high-density urban areas are more likely to use the bike as the access/egress mode to transit. Similarly, Shelat et al. (2018) analysed user and trip characteristics of the combined transit and bike mode in the Netherlands and found that long-distance commuters tend to use the bike as an access mode to the train, while walking is used as an egress mode from the train. Focusing on metro, Chen et al. (2012) investigated the determinants of bike transfer demand at metro stations and found that land use of the surrounding neighbourhoods, bike rental services and lack of bus services affect the decision to use the bike to access/egress the metro. Recently, Zhao and Li (2017) also examined the determinants of bike use to access/egress metro stations in Beijing. They found that travel distance is the most important factor influencing commuters’ cycling rates to and from metro stations. Personal socio-economic characteristics such as age, income and education and structural context including built environment, transportation infrastructure, and transit services also had an impact on individual’s choice of the combination of metro and bike.

With the popularization of bikesharing systems, scholars began to focus on the integration of public bike and metro, especially in the developing countries. For instance, Yang et al. (2015) found that 62.8% of the commuters in suburban Nanjing intend to shift to using public bikes to access the metro. Ji et al. (2017) showed that male, younger, high-income transit commuters and those who have experienced bike theft are more likely to use public bikes to access the metro. Lin et al. (2018) conducted a comparative study to analyse the associations between built environment and public bike use for metro access in Beijing, Taipei, and Tokyo. They suggested that it is critical to consider local geographic and cultural proximity when developing cycling-friendly built environments. Unlike the above studies, Bachand-Marleau et al. (2011) used survey data collected from a public bike sharing system and Bixi (bike taxi) to identify current or potential users of transit-bike integration. They found those commuting by metro, or using car are most likely metro-bikeshare users. Tang et al. (2018) focused on the bikeshare pool sizing problem at public transit station for bikeshare commuters. They indicated that a bike sharing system helping commuters to solve their first/last mile travel would stimulate more people to use public transportation. Ji et al. (2018) and Zhao et al. (2018) recently identified metro-bike transfer trips from smart card data. Based on the identified results, Ma et al. (2018) further analysed the usage pattern of metro-bikeshare from four dimensions: transfer time, date, space and transfer modes including “return-enter” mode and “exit-lease” mode. They found that metro-bikeshare travel patterns vary across users with age and that the young users are the largest group.

Overall, research on the choice of metro-bikeshare using smart card data is still in its early stage. None of the aforementioned studies used smart card data to examine the determinants of using metro-bikeshare for commuting. In addition, because young people are the most critical group, further research on factors influencing their combined use of metro and public bike is warranted. This study addresses such relative gaps in the literature.

3. Study area and data

3.1. Study area

Nanjing, a representative Chinese central city, was adopted as the study area. As the capital of Jiangsu Province, Nanjing is an important
central city and commercial center in Eastern China. Nanjing’s residential population was 8.27 million in 2016, of which 6.78 million lived in the urban area (Nanjing Statistics Bureau, 2017). Nanjing’s built-up area was 1125.78 km² and the urbanization rate was 82% (Nanjing Statistics Bureau, 2017). The city can be divided into three regions, namely, core, inner and remote areas (Song et al., 2018). As one of the first transit metropolises in China, Nanjing currently has 7 metro lines with a total of 133 stations, covering 25 routes of 224 km, and 1024 bike stations with about 40,000 public bikes. It basically forms a metro-bikeshare system with metro as the mainstay and public bikes as the feeder mode. The spatial distribution of metro lines and bike stations is shown in Fig. 1. It should be noted that only the metropolitan core area and peripheral areas were selected as the study area since in 2016 the public bike system was not implemented in remote regions.

Nanjing was selected for two reasons. First, Nanjing public bike systems is one of the largest public bike operation systems in China, which was launched in 2013 by the government for public welfare. In 2016, the daily use of public bikes in Nanjing reached the peak of 250,000 trips, public bike has become one of important travel modes for Nanjing citizens (Nanjing Government, 2018). The number fell to 120,000 trips due to the emergence of FFBS systems in 2017. FFBS became hugely popular in Nanjing for its convenience and low cost at the beginning, with a daily usage of about 20 million trips (Hua et al., 2020). However, the prosperity of FFBS has only lasted for about one year. In 2018, the Nanjing government has controlled the number of free-floating bikes from 600,000 to 317,000 due to its excessive market launch and a series of negative problems thereby caused (CCTV News, 2019). At the same time, the average daily use of public bikes in Nanjing has begun to pick up. By the end of 2018, the city’s average daily usage of public bikes was about 200,000 trips, increased by nearly 20% compared with 2017 (Nanjing Government, 2018). More importantly, with the price of FFBS rising (1.5 yuan per half an hour usage), more and more FFBS users has begun to ride public bikes (the first two hours are free), especially those cycling for commuting trips.

Second, Nanjing is a typical Chinese single-center city. Most job opportunities are located within the metropolitan core area, while the residential density of the core area is relatively low (Fig. 2). With the continuous expansion of the urban area, the spatial-temporal distribution of travel has changed. The average travel distance has increased by 40% over the past ten years and currently exceeds 6.5 km (Nanjing Transport Bureau, 2018). It is predicted to further increase with the city expanding. As a result, increasingly more residents, especially young people will face the jobs-housing separation, and they are likely to use cars for long-distance commuting. In 2017, private car ownership in Nanjing has already broken through 2 million barriers. Due to an annual increase of about 7.9%, problems in the urban transport system become increasingly serious.

3.2. Data

In order to explore the relationship between the frequency of young commuters using metro-bikeshare and social demographics, travel-related characteristics and the built environment, four types of data are used in this study.

To identify combined metro-bikeshare trips, metro smart card data and public bike smart card data from the same time period are needed. Considering the impact of weather on ridership, smart card data from March 9 to 29, 2016, were used, and excluded weekends. The data were obtained from the Nanjing Metro Company and Nanjing Public Bicycle Company, respectively. Although the data were obtained from two different companies, the same ID is used in the two datasets and thus these datasets can be used to identify combined metro-bikeshare trips. In addition, the smart card includes the socio-demographic profile of
the cardholders including gender, age and hukou. Hukou institution is a special population policy used in mainland China to control population movement from the countryside to the city (Li and Zhao, 2015). Residents with local urban hukou can receive a higher level of the state’s welfare (such as the purchase qualifications of a house or car) than those without local urban hukou. In this study, the cardholders’ age was used to select young people. The recorded travel information includes boarding station, alighting station), time of boarding and alighting, and smart card type (e.g., regular card, student card and elderly card).

The 2015 Nanjing household travel survey data were also obtained and used to measure the neighborhood-level socioeconomic characteristics. The data included one-day trips of 92,147 individuals belonging to 31,351 households of Nanjing, representing a sampling rate of 1.14%. Respondents were asked to record their socioeconomic information (including household and individual) and activity and travel information during 24 h. For the purpose of the current research, only the socioeconomic information about households including household size, car ownership, e-bike ownership, bike ownership and annual income were extracted from the dataset. These socioeconomic characteristics were separately calculated at a bike station buffer level of individual trip endpoint (destination bike station).

As for the built environment characteristics, point of interest (POIs) data (residents, shops, parking lots, bus stations and bike stations etc.) were collected from Baidu Map (see Section 4.2 for a detailed description).

Note that we only extract the information necessary for the analysis of the paper from the original data as mentioned above, which will not disclose any personal identifiable information.

4. Methods

4.1. Extracting the dependent variable

The frequency of using metro-bikeshare during a single trip by young commuters is the dependent variable in this study. SC holders who transfer between metro and public bike can be matched on the basis of the same card ID.

Following Ji et al. (2018), Ma et al. (2018) and Zhao et al. (2018), the combined metro-bikeshare trips were identified as follows: (1) Cycling trips that lasted less than one minute and metro transactions which start and end at the same station less than five minutes were deleted in consideration of the actual travel of users. (2) Metro transactions and bikeshare transactions were paired by one metro transaction with the subsequent bikeshare transaction, and the bikeshare transaction followed the metro transaction with the same card ID by time of day. (3) Interval distance and interval time of paired metro and bikeshare transactions were calculated, respectively. Interval distance is the Euclidean distance between the metro station and its paired bikeshare station. Interval time is the duration between return time for a bike and next check-in time at paired metro station or between checkout time at a metro station and next rent time for a bike. (4) 300 m and 10 min were used as the maximum value of the distance and time to identify transfers, since over 90% of transfer trips were finished within 300 m and 10 min. When the interval distance and time of paired metro transaction and bikeshare transaction were less than 300 m and 10 min, the paired transactions were identified as transfer trips. (5) The metro stations with only a few identified metro-bikeshare records (less than 30 times during continuous 15 workdays) were eliminated to make sure non-typical transfer patterns distort the results.

There were approximately 23.86 million and 1.91 million original transaction records contained in the metro- and public bike SC data, respectively. After identifying, the SC data produced a dataset of 23,213 combined metro-bikeshare trips made at 39 metro stations, with a total of 6,682 individuals. Fig. 3 shows the spatial distribution of the identified metro-bikeshare pairs and the number of transfer trips at each metro station. For each metro station, the darker the color red, the more the metro-bikeshare transfer trips it has. The minimum number of transfer trips at a metro station is 30, while the maximum number is 10,680 during continuous 15 workdays.

After the identification of metro-bikeshare trips, we need to identify commuters among the users. Given that commuting behavior has relatively regular spatial–temporal characteristics, metro-bikeshare commuters can be identified based on the users’ travel frequency, spatial constraints and temporal constraints. We assumed that a commuter is someone who frequently uses public bikes access/egress to the regular metro stations and travels for a fixed distance on the metro during the peak hours. To ensure that low frequency users who are being attracted by metro-bikeshare are not wrongly deleted, there are four steps to identify the commuters. First, we need to find out the regular metro stations. In the identified metro-bikeshare stations, 87.74% of individuals transferred at the same metro station and 11.48% at two different metro stations. It suggests that most of the commuters use metro-bikeshare only for the purpose of commuting. Since the transfer stations of metro-bikeshare commuters are usually no more than two (i.e. home-end station and work-end station), we deleted the individuals who transfer at more than two metro stations. Second, considering most commuting trips occur during peak hours, we only retained the metro-bikeshare trips during 7:00–9:00 and 17:00–19:00 as potential commuting trips. Third, we calculated Euclidean distances for individuals’ each potential commuting trip based on their metro origin–destination locations and deleted individuals with different travel distances for each trip. Fourth, the retained individuals evidenced metro-bikeshare trips less than three days in the fifteen workdays (less than one day per week on average) were removed, as the number of
trips is too small and there is a lack of regularity.

Finally, a total of 2,040 individuals were identified as metro-bikeshare commuters. Young commuters (18–35 years old) accounted for 56.71% of these commuters. A comparison of the usage frequency of metro-bikeshare between young commuters and other commuters during the selected continuous 15 days showed that young commuters’ usage frequency was 5.58 days/15 workdays, while other commuters’ usage frequency was 4.78 days/15 workdays. We further used the one-way analysis of variance (ANOVA) to test the usage frequency of the two groups, the p-value < 0.001 indicates that young commuters are more likely to choose metro-bikeshare than other commuters. The frequency of metro-bikeshare use of young commuters is shown in Fig. 4, which shows a declining trend. The vast majority of youths only used the metro-bikeshare occasionally on weekdays, with frequencies of 3 days/15 days. One possible reason is that young people who are being attracted by metro-bikeshare still prefer to choose other travel modes for their commuting trips. We further calculated the variance of the usage frequency and found its value (7.62) is larger than the mean.

4.2. Selecting the explanatory variables

As discussed, individual’s usage frequency of the combined metro and bikeshare mode is mainly influenced by four factors: individual and neighborhood-level socioeconomic characteristics, travel-related characteristics and built environment characteristics. Therefore, we extracted the corresponding indicators from the dataset. The statistical analysis results of the selected variables are shown in Table 1. Specifically, three characteristics of individuals were considered: gender, age, and hukou status.

Table 1 shows that the proportion of male and female users is almost the same (49.48% vs 50.52%). About 61% of the young users were younger than 30 and most of them were local people from Nanjing (72.10%). For this study, neighborhood-level socioeconomic characteristics were represented by the proportion of car ownership, bicycle ownership and high-income households within 400 m buffer of each bike station of individual trip endpoint (destination bike station).

Travel-related characteristics include four categorical variables: travel duration on metro, travel duration on public bike, transfer distance and transfer time including transfer during the morning peak, evening peak and at other times. Since commuters might have different travel-related characteristics during the fifteen workdays, the average travel duration on the metro and by public bike were calculated for each commuter separately. Besides, one transfer time can only be deemed as the main transfer time if the total number of that transfer exceeded 50% of the commuters’ total transfer time. Table 1 shows that more than 50% of the young commuters’ travel on metro takes between
15 and 30 min, while rides on a public bike last between 5 and 10 min.

Table 1 also shows that 65.33% of young commuters cycle to/from metro stations during the morning peak hours. Their average walking distance between metro station and the selected bike station along the road network is 89.77 m.

Five built environment characteristics surrounding metro stations were selected: distance to the CBD, distance to the nearest bus stop, number of parking lots, number of bike stations and number of bike docks. Table 1 shows the measurements and descriptive statistics of these variables.

- Gender: Male (571, 49.48%) vs. Female (583, 50.52%)
- Age: 18–30 (713, 61.79%) vs. 30–35 (441, 38.21%)
- Hukou: Local (852, 72.10%) vs. Non-local (322, 27.90%)
- Percentage of car ownership: Average within 400 m buffer of destination bike station (0.23, 0.18) vs. within 1 km buffer of destination bike station (0.29, 0.28)
- Travel duration on metro: Average travel duration on the metro for one commuting trip (< 15 min (260, 22.53%) vs. ≥ 30 min (304, 26.34%)
- Travel duration on public bike: Average cycling duration of public bike for one access/egress trip (< 5 min (270, 23.40%) vs. ≥ 10 min (302, 26.17%)
- Transfer time: Transfer time during the morning peak/evening peak (754, 400, 65.33% vs. 34.67%)
- Walking distance: Average travel distance between metro station and the selected bike station along the road network (89.77, 40.08)
- Distance to the CBD: Average distance from metro station to the CBD along the road network (209, 18.11% vs. 674, 58.40%)
- Distance to the nearest bus stop: Average distance from metro station to the nearest bus stop within 1.5 km buffer of metro station (278, 24.09% vs. 271, 23.48%)
- Number of parking lots: Average number of bus stops within 1.5 km buffer of metro station (338, 29.32% vs. 187, 16.20%)
- Number of bike stations: Average number of bike stations within 1.5 km buffer of metro station (212, 18.41% vs. 603, 52.27%)
- Number of public bike docks: Average number of bike docks at metro station/destination bike station (58.41, 12.19% vs. 12.19, 9.76%)
- Number of bus stops: Average number of parking lots within 400 m buffer of destination bike station (204, 17.68% vs. 413, 35.79%)
- Location of destination bike station: If the commuter lives/works in the core area; 0: otherwise (123, 10.66% vs. 1030, 89.34%)

Note: "*" represents the built environment variables around destination bike station.
4.3. Negative binomial regression model

The aim of this research is to examine the influence of individual and neighborhood-level socioeconomic characteristics, travel-related attributes and built environment factors that influence young commuters’ frequencies of the combined metro and bike mode use. Poisson and negative binomial (NB) regression models are commonly used for modeling count variables (e.g., crash frequency), as event occurrences are discrete, non-negative and random (Naznin et al., 2016). Accounting for reported over-dispersion inherent in the usage frequency data, a negative binomial regression model was estimated. Negative binomial (NB) regression is a widely used for modeling over-dispersed Poisson count data where the conditional variance exceeds the conditional mean. The NB model can be described as:

\[
\mu_i = \exp \left( \sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i \right)
\]

where \( \mu_i \) is the mean or expected usage frequency of the \( i \)th commuter (\( i = 1, 2, \ldots, n \)), \( x_{ij} \) represents the \( j \)th explanatory variable (\( j = 1, 2, \ldots, m \)) of the \( i \)th commuter and \( \beta_j \) is the corresponding parameter to be estimated. \( \varepsilon_i \) is the error term and \( \exp( \varepsilon_i ) \) is a gamma-distributed error term. According to this equation, the impact of the selected explanatory variables on the frequency of using metro-bikeshare can be expressed with the following incidence rate ratio (IRR):

\[
\frac{\exp(\sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i)}{\exp(\sum_{j=1}^{m} \beta_j x_{ij} + \varepsilon_i)} = \exp(\varepsilon_i)
\]

where \( \mu_i \) is the expected usage frequency of the \( i \)th commuter (\( i = 1, 2, \ldots, n \)) when the \( j \)th explanatory variable of her/him increases by one unit and \( \beta_j \) is the corresponding parameter.

The parameters of the model were estimated using maximum likelihood, while AIC (Akaike Information Criterion) was used to evaluate whether this model provides a better fit than the corresponding Poisson model. The smaller the AIC, the better the model is.

5. Results

Before estimating the model, multi-collinearity among variables was examined. Correlation among variables was tested using variance inflation factors (VIF). The correlation was considered high when the VIF was more than 7.5. Results indicated that none of the variables was highly correlated with other variables as the VIFs are less than 7.5. Therefore, all explanatory variables were included in the NB model. We retained the variables significant at the 95% level, and removed the other variables one by one to ensure the best fit of the model. The final parameter estimates, and model fit information are shown in Table 2. It is noted that the variables gender, age and hukou have been retained as it increases the goodness-of-fit of the model. The likelihood-ratio test of dispersion parameter alpha is equal to zero indicating that the dependent variable is over-dispersed and hence the NB distribution is appropriate. The p-value of the final model was < 0.001 (< 0.05) which shows that the model with individual and neighborhood-level socioeconomic attributes, travel-related characteristics and built environment characteristics fits better than the null model. Besides, the AIC of the model is 5225.914, which is smallest in all possible models.

Note that the pseudo R² value reported in NB model is used as a reference for choosing the best model between possible models but as an absolute measure of model fit (STATACorp, 2015). In order to verify the actual goodness-of-fit of the model, we also compared the predicted results of the model with the actual values of the samples. We found that there are 61.26% of the total samples with the value of absolute error less than two, 33.02% of which with the value of absolute error less than one. This indicates that the regression performance of the model is good.

The z-statistic indicates that the percentage of bike ownership associated with neighborhood-level socioeconomic socioeconomic; travel time on bikeshare and transfer time associated with travel-related characteristics; built environments around metro stations, including the shortest distance from metro stations to the CBD, the shortest distance from metro stations to bus stops, the number of parking lots and the number of bike stations; built environments around destination bike stations, including the number of docks, the number of bus stops and destination bike stations location have a significant impact on the usage frequency.

As for travel-related characteristics, compared to commuters who travel by bikeshare between 5 and 10 min, while holding the other

---

Table 2
Negative binomial regression analysis for usage frequency.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Coef.</th>
<th>Std.Error</th>
<th>z-Statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (Male = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>0.015</td>
<td>0.027</td>
<td>0.579</td>
</tr>
<tr>
<td>Age (30-35 years old = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20-30 years old</td>
<td>0.040</td>
<td>0.028</td>
<td>1.435</td>
</tr>
<tr>
<td>40-49 years old</td>
<td>-0.039</td>
<td>0.030</td>
<td>-1.282</td>
</tr>
<tr>
<td>Percentage of bike ownership</td>
<td>0.263</td>
<td>0.058</td>
<td>4.500</td>
</tr>
<tr>
<td>Travel time on metro (115,30 min = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 15 min</td>
<td>0.014</td>
<td>0.034</td>
<td>0.421</td>
</tr>
<tr>
<td>&gt;= 30 min</td>
<td>0.056</td>
<td>0.032</td>
<td>1.784</td>
</tr>
<tr>
<td>Travel time on bikeshare (15,10 min = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt; 5 min</td>
<td>-0.020</td>
<td>0.034</td>
<td>-0.592</td>
</tr>
<tr>
<td>&gt;= 10 min</td>
<td>-0.114</td>
<td>0.034</td>
<td>-3.397</td>
</tr>
<tr>
<td>Transfer time (Evening peak = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Morning peak</td>
<td>0.126</td>
<td>0.029</td>
<td>4.380</td>
</tr>
<tr>
<td>Transfer distance</td>
<td>0.031</td>
<td>0.027</td>
<td>1.157</td>
</tr>
<tr>
<td>Distance to the CBD (Others = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>-0.254</td>
<td>0.069</td>
<td>-3.662</td>
</tr>
<tr>
<td>4th quartile</td>
<td>-0.181</td>
<td>0.072</td>
<td>-2.506</td>
</tr>
<tr>
<td>Distance to the nearest bus stop (Others = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>0.016</td>
<td>0.063</td>
<td>0.258</td>
</tr>
<tr>
<td>4th quartile</td>
<td>0.255</td>
<td>0.112</td>
<td>2.284</td>
</tr>
<tr>
<td>Number of parking lots (Others = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>0.051</td>
<td>0.109</td>
<td>0.469</td>
</tr>
<tr>
<td>4th quartile</td>
<td>0.266</td>
<td>0.104</td>
<td>2.564</td>
</tr>
<tr>
<td>Number of bike stations (Others = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>0.180</td>
<td>0.111</td>
<td>1.620</td>
</tr>
<tr>
<td>4th quartile</td>
<td>0.188</td>
<td>0.096</td>
<td>1.958</td>
</tr>
<tr>
<td>Number of docks at destination bike station</td>
<td>0.129</td>
<td>0.061</td>
<td>2.126</td>
</tr>
<tr>
<td>Number of bus stops (Others = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st quartile</td>
<td>-0.065</td>
<td>0.039</td>
<td>-1.682</td>
</tr>
<tr>
<td>4th quartile</td>
<td>-0.062</td>
<td>0.039</td>
<td>-1.595</td>
</tr>
<tr>
<td>Location of destination bike station (Inner area = ref.)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Core area</td>
<td>0.043</td>
<td>0.074</td>
<td>0.576</td>
</tr>
<tr>
<td>Intercept</td>
<td>0.825</td>
<td>0.285</td>
<td>2.900</td>
</tr>
<tr>
<td>Pseudo R²</td>
<td>0.050</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Log likelihood</td>
<td>-2589.957</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prob &gt; chi²</td>
<td>0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AIC</td>
<td>5225.914</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a significance at the 95% level.
b significance at the 99% level.
variables constant in the model, the usage frequency of those traveling by bikeshare over 10 min is expected to decrease by a factor of 0.892. However, no statistically significant difference is found between commuters who travel by bike less than 5 min and the reference group. This indicates that 10 min is a turning point, after which a significant decrease in use of metro-bikeshare is observable. By contrast, the estimated coefficient for the morning peak is 1.126, indicating that given the other variables are held constant, the expected number of usage frequency for commuters who travel during morning rush hours is 3.083 (IRR = exp (0.126)) times that of commuters who travel during the evening hours. This suggests that early rush hour has a greater impact on commuters’ use frequency of metro-bikeshare. Unexpectedly, the commuter’s travel duration on metro and walking distance between metro station and the selected bike station have no effect on the use frequency of the combined mode.

For the built environment surrounding metro stations, the regression results displayed in Table 2 show that distance between metro station and the CBD is significantly related to cycling transfer frequencies. Compared with commuters who transfer at metro stations with the medium distance to the CBD, the usage frequency for those transferring at metro stations with short and long distances to the CBD would be expected to decrease by a factor of 0.776 and 0.834, respectively, while holding the other variables constant in the model. One possible reason for this result is that commuters use walking instead of bikes due to the short distance to metro stations in the CBD areas; there are fewer cycling facilities and services in the areas far from the city center. However, the estimated coefficient for long distance to the nearest bus stop is 0.255. It indicates that bus stops that are too far from a metro station increase the frequency of combined metro-bikeshare use. Interestingly, the significant positive coefficient of the number of parking lots suggests that the presence of more parking lots in metro station areas increases the cycling frequencies to/from metro stations. One of the main reasons might be that areas with high parking density are generally more commercially developed, thus resulting in more traffic congestion and parking costs. In addition, compared with commuters who transfer at medium-density bike stations, the usage frequencies for those transferring at high-density bike stations decreases by a factor of 0.829. This suggests that setting up too much bike stations in metro station areas is not conducive to metro commuters choosing public bike as a transfer mode.

In terms of the built environment surrounding destination bike station areas, if a bike station were to increase its docks by one unit, the usage frequency of the metro-bikeshare integration would be expected to increase by a factor of 1.138, holding the other variables fixed. However, the number of bike docks in metro station areas is not related to the frequency of using metro-bikeshare. This suggests that the increasing docks near residential and employment locations compared with those in metro station areas could encourage young commuters cycling as a transfer mode, even though the effect is small. The negative estimated coefficients of large number of bus stops indicate that the number of bus stops too much in residential and employment locations reduces commuter’s cycling transfer frequencies. It indicates that there exists a substitution relationship between bus and cycling. In addition, no statistically significant difference existed between commuters who live or work in the core areas and those live or work in the inner areas.

Unexpected, socio-economic attributes of young commuters including gender, age, and hukou status have no significant impact on their frequencies of using metro-bikeshare. Nevertheless, the estimated coefficients indicate that females, commuters younger than 30 years old and non-local residents use bike more frequently to transfer to/from a metro station than the reference group. It should be noted that the percentage bike ownership around individual trip endpoints has a significant positive effect on the usage frequency of the combined mode. If the percentage bike ownership increases by one unit, the frequency of commuters using metro-bikeshare is expected to increase by a factor of 1.301.

6. Discussions and conclusions

This research explored the influence of individual and household socioeconomic characteristics, travel-related and built environment attributes on the usage frequency of cycling as a feeder mode to/from metro station. Some interesting results were found. Firstly, findings show that young commuters are more likely to use public bikes as a feeder mode to/from metro stations than other age groups, and the frequency of metro-bikeshare use of the 18–30 group is significantly higher than of the 30–35 group (Table 1). In Nanjing, most young commuters under 30 who have been just entered the workforce tend to be poorly paid, stressed, and unmarried (the average age of marriage of Nanjing residents was 32.6 in 2017). Therefore, their choice of travel mode is probably more the result of a lack of opportunity to own or use cars instead of a preference for cycling. In other words, although some of these commuters may not be willing to choose metro-bikeshare, they have no better options at this stage. Therefore, transport policies should attempt to discourage these young people from converting to car users by promoting public transport services. Besides, offering a discounted public bike deposit and rental price may be appropriate (Ji et al., 2017). Unlike the 18–30 group, most people in the 30–35 group are building their careers. These up-and-comers are experiencing a series of important events, such as career promotions, increasing income, marriage, and childbirth. They are highly likely to shift from public transport users to car users because the likelihood of car purchasing and needs usually increase with increasing income and/or having children. However, most of these commuters have only one car per household. Hence, there are still many young people in this age group using the metro-bikeshare option. A second car purchasing limitation may have a short-term effect, but the limitation can never be effective without promoting metro-bikeshare services.

Secondly, our results show that commuters are expected to use public bikes to transfer to/from metro as a feeder mode when the cycling time is less than 10 min. Considering that the time they spent on picking up/dropping bikes, 2 km is a tolerance threshold of the distance for cycling between homes/workplaces to metro stations. However, over 20% of young commuters cycle over 2 km (Table 1). Thus, it is necessary to shorten the travel distances between residential/employment locations and metro stations. The minimum standard of transit-oriented development policies should ensure that individuals can access public transport within an acceptable cycling distance to/from residential, commercial or workplace. Besides, as He et al. (2016) showed, younger commuters possess a shorter tolerance threshold commuting time, and 40% has a tolerance threshold of commuting time < 30 min. In other words, when commuting time exceeds young people’s tolerance threshold, they are highly likely to use cars to save time. In our study, the average time young commuters spent on metro was 25.12 min and the average commuting time was 35 min. This implies that some of the metro-bikeshare users are potential car users. Therefore, policies that reduce young people’s commuting time under their tolerance threshold can considerably encourage them to use public transport. Providing staff dormitory or low-rent homes near young people’s workplaces could be an effective short-term solution. In the long-term, the current job-housing separation must be mitigated by developing multiple intensive land use, especially in inner and remote areas.

Thirdly, different measures and policies should be implemented respectively for core areas, inner areas and residential/employment locations. Our results revealed that high-density bike stations in metro station areas are not conducive to metro commuters choosing public bike as a transfer mode. On the one hand, the majority of high-density bikes stations are distributed in the metropolitan core area where the density of metro stations and the distances between these stations are short. Thus, commuters may have more alternatives other than public bikes, such as walking. On the other hand, since commuters’ trips are mostly undertaken during rush hours, commuters living in core areas
may give up using public bikes due to the crowded metro. Therefore, it is more important to improve the level of public transport services and implement effective travel demand management for commuters in the core areas. However, the supply of public transit facilities is inadequate in inner areas and residential /employment locations. Therefore, adding bike parking facilities in these areas could encourage metro passengers to choose cycling as a transfer mode.

Furthermore, the impact of the proportion of private bike ownership near residential /employment locations on young commuters’ usage frequency of metro-bikeshare is also eye-catching. Our results indicate that with the increase in private bike ownership, young commuters are more willing to frequently choose metro-bikeshare. This result is seemingly in contrast with our common understanding that there is a substitution relationship between private bike and public bike. One reason may be that owning a bike can directly reflect the commuter’s attitude towards bike using. Those who own bikes are probably more likely to adopt metro-bikeshare as a travel mode. This means that they are more willing to choose public bike as a feeder mode to metro. Another reason may be as Ji et al. (2017) pointed out that commuters, especially those who have bicycle theft experiences are more willing to choose public bikes for metro access due to a low costs and responsibilities for owning a bicycle. Therefore, it is important to change the image of cycling and revive the bicycle culture among the younger generation.

The study has some limitations. Due to the limitation of the dataset, household characteristics were not considered. However, commuters travel mode choice might be influenced by household characteristics, such as household size, household car ownership, the number of children in a household and household time use (Zhang and Timmermans, 2017). Therefore, a questionnaire containing household factors that may affect metro-bikeshare travel behavior needs to be designed for further analysis. Besides, since Nanjing’s public bike scheme has not been implemented in suburbs, commuters in suburb areas (Luhe, Jiangning, Pukou district) are not included in this study. Commuters living/working in core areas and suburbs may have different travel behavior, their heterogeneous choice of metro-bikeshare integration needs to be investigated.

Nevertheless, this study contributes to the literature on metro-bikeshare integration by exploring the determinants of young commuters’ usage frequency of metro-bikeshare. Firstly, although there is a large body of literature about metro-bikeshare integration, young commuters, as the biggest user group of metro-bikeshare integration system have been paid few attentions. Thus, it is difficult to know these young commuters’ commuting patterns by metro-bikeshare integration and their possible changes in these patterns. This study focuses on the use of public bike as a feeder mode to/from metro for young commuters between 18 and 35 years old, which may help for filling a gap in the existing literature. Secondly, young commuters who use bikes as an access/egress mode had been extracted from three weeks of metro smart card data and public bike smart card data with the same ID. Hence, our result can show young commuters’ usage frequencies of metro-bikeshare integration, which is difficult to obtain by a traditional survey data. Thirdly, relevant policy recommendations for young metro commuters who cycle as a feeder mode are proposed from the aspects of individual attributes, travel-related characteristics, the built environment and cycling preference. These recommendations are intended to encourage young commuters, especially those travel relatively long distance, to use the combined metro and bike system.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This research is funded by the National Key R&D Program of China (Grant No. 2018YFB1600900), the Research Foundation of Graduate School of Southeast University (Grant No. YBJJ1842), the Fundamental Research Funds for the Central Universities (No. 2242020K40063) and the China Scholarship Council (CSC) (Grant No. 201806090205). We are grateful for the comments and suggestions from the editor and the anonymous reviewers who helped improve the paper.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tbs.2020.06.007.

References


Damaio-Sirois, G., El-Genaidy, A.M., 2019. To be or not to be dockless: empirical analysis of dockless floating bike sharing in Nanjing cut down and the use of public bicycles has picked up. Acces.


STATACorp., 2015. Stata Statistical Software: Release 14 [Online]. StataCorp LP, CollegeStation, TX.


