International Journal of Transportation Science and Technology 9 (2020) 159–169

Contents lists available at ScienceDirect

International Journal of Transportation Science and Technology

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

Unlimited-ride bike-share pass pricing revenue management for casual riders using only public data

Gyugeun Yoon, Joseph Y.J. Chow *

C2SMART University Transportation Center, New York University, Brooklyn, NY, USA

ARTICLE INFO

Article history:
Received 28 April 2019
Received in revised form 23 December 2019
Accepted 14 January 2020
Available online 23 January 2020

Keywords:
Bike-share
Unlimited-ride pass
Bootstrap method
Revenue management
Logit model

ABSTRACT

Despite the proliferation of publicly available Big Data in Mobility-as-a-Service systems, few studies in the urban mobility service literature deal with unlimited usage price plan strategies. We conduct an experimental case study to design such a strategy: an unlimited-ride X-Day pass pricing for bike-share usage especially targeting short-term casual users. Public data from Citi Bike is used to estimate a pass choice model for bike-share services. As disaggregate data for riders have not been available due to privacy concerns, their travel behaviors are veiled and kept confidential. The estimation is made possible using bootstrap method to resample average numbers of daily trips for individuals from coefficients of a linear regression model that distributes daily trips to 1-Day and 3-Day Pass users. The analysis suggests 1-Day Pass users make 2.8 trips per day at an average cost of \$4.29/trip while 3-Day Pass users make 1.8 trips per day at an average cost of \$4.47/trip. The estimated pass choice model from the choice-based sample has a cost per trip coefficient of −2.884. The model is used in designing a new pricing plan of (\$18.50, \$12) for (3-Day, 1-Day) compared to a benchmark of (\$24, \$12). By using only public data, we can nonetheless show that this pricing plan should increase monthly revenue for Citi Bike by at least 5.5% and increase consumer surplus by at least \$0.09/trip/short-term pass customer.

1. Introduction

Bike-share systems have drawn the attention of the public, engineers, and decision makers as sustainable urban Mobility-as-a-Service (MaaS) alternatives that consume less space. Many of these systems use "unlimited-ride" pricing plans to increase ridership and to better capture detailed information from riders. However, plans vary from system to system. Table 1 explains pricing plans for bike-share systems in some cities around the world; those of Citi Bike in New York City (NYC) are relatively concentrated on short-term plans (e.g. 1-Day Passes, 3-Day Passes).

Among operational bike-share systems, Citi Bike is one of the most popular systems in North America, serving more than 274,000 annual members, or subscribers, with more than 12,000 bikes and 750 stations as of the end of June 2018. This number does not even include short-term pass customers, or casual users, who access from the kiosk on the street or via website or application. Citi Bike offers three pricing plans to users; a 1-Day Pass for 24-hour access, a 3-Day Pass, and an Annual

Peer review under responsibility of Tongji University and Tongji University Press.
* Corresponding author.
E-mail address: joseph.chow@nyu.edu (J.Y.J. Chow).

https://doi.org/10.1016/j.ijtst.2020.01.003
2046-0430/

Table 1
Pricing plans of bike-share systems in the world.

Pricing Plan	1-Day	3-Day	7-Day	30-Day	90-Day	1-Year membership
Citi Bike (New York City, NY)	$12	$24	–	–	–	$169
Indego (Philadelphia, PA)	$10	–	–	$17	–	$156
BLUEbikes (Boston, MA)	$10	–	–	$20	–	$99
Bixi (Montréal, Canada)	$5	–	–	$32	$57	$91
Captial Bikeshare (Washington D.C.)	$8	$17	–	$28	–	$85
Santander Cycles (London, UK)	£2	–	–	–	–	£90
Vélib' (Paris, France)	€5	–	€15	–	–	€37.20
Nubija (Changwon, South Korea)	₩1000	–	₩2000	₩4000	–	₩30,000
bike sampa (São Paolo, Brazil)	R$8	R$15	–	R$20	–	R$160

Note: Websites of systems were accessed on June 7th 2018.

Membership. Special offers are provided occasionally to certain customer groups. Table 2 provides detailed explanations of each plan. The system also imposes an additional fee for later returns that exceed the maximum ride time.

According to a sample data analysis for the first week of September in 2017, 10,480 bikes were on average used 5.36 times per day. The usage count of the most frequently used bike was 119 times during the same period, which averages 17 times per day. The bikes were ridden for 17.85 minutes per trip and were occupied for 1.60 hours per day, or 6.7% of a 24-hour period. This shows that the system can afford more bike users by inducing trip demand from other modes as its bikes stay docked most of the day.

A bike-share service needs to establish an effective pricing strategy to improve their revenue or increase bike ridership to promote usage of the entire system. It is obvious that the design of a pricing plan impacts membership demand and ridership because the effectiveness of a set of pricing plans depends highly on the behavior of the potential customer population. For example, in the third quarter of 2017, Citi Bike had average daily sales of 2465 1-Day Passes and 268 3-Day Passes, and these numbers may differ from other cities with different pricing options.

An effective pricing design depends on the availability of user information. Whereas a bike-share system can obtain and archive many sociodemographic details from annual pass customers' subscriptions, it is unable to collect such information from casual users who buy short-term passes while providing limited information. This deficiency makes it much harder to consider revenue management strategies for casual users, despite their large number.

Nonetheless, casual users should be separated from subscribers and analyzed independently due to their different behaviors as Kaviti et al. (2019) studied. The intercept survey they conducted at bike stations revealed that the frequency of cycling and trip purpose of both groups are distinguishable. Fig. 1 indicates the number of starting trips at Citi Bike stations in QGIS with the 20 most popular stations listed. Patterns of the short-term customers on the left clearly differ from the subscribers on the right. Pershing Square North station, which can be easily approached from Grand Central Terminal, processes the highest demand generated by annual subscribers. Stations in Lower Manhattan installed along Broadway also attract many long-term members. On the other hand, short-term pass users' favorite stations are usually adjacent to Central Park and Battery Park, two of the most popular tourist attraction points in NYC.

With the proliferation of publicly available Big Data for bike-share services, how can we use it to model the behavior of casual pass users to support revenue management strategies like price setting for different passes? Research on bike-share pricing plans is not frequently reported, and most of them focus on incentivization strategies for system balancing. We make three contributions to address this research gap in this case study. First, we design a new inference method for casual pass users of bike-share systems based on the publicly available Big Data. The method relates cost per trip to 1-Day and 3-Day Pass users and applies a bootstrapping method to resample those costs to the sample data. Second, we propose and estimate the first casual pass choice model for bike-share systems based on the choice-based sample of bootstrapped trip cost data. Third, we suggest new revenue management strategies and analytics for Citi Bike focusing on price adjustment from the current operating benchmark that should improve revenue and consumer surplus.

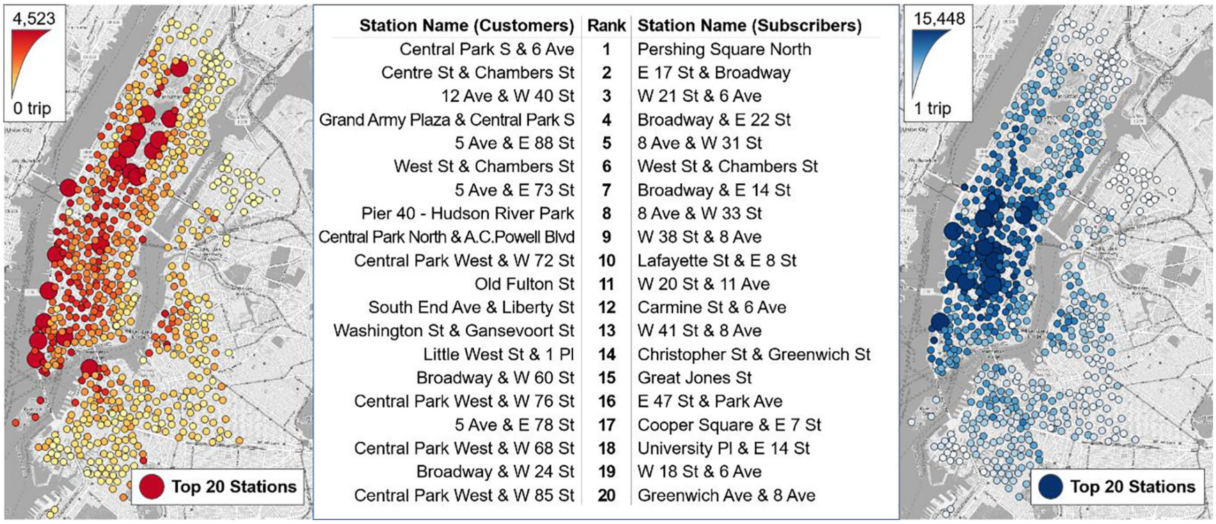
2. Literature review

Bike-sharing systems need to consider costs when they establish a pricing plan as general businesses do. According to the Institute for Transportation and Development Policy (2018), most systems decide their pricing structure based on operating

Table 2
Pricing plan for Citi Bike in New York City.

Plan	1-Day Pass	3-Day Pass	Annual Membership
Price	$12	$24	$169
Detail	• 24 hours of Citi Bike access • 30 minutes for each ride	• 72 hours of Citi Bike access • 30 minutes for each ride	• Unlimited ride for a year • 45 minutes for each ride

Source: Citi Bike website, https://www.citibikenyc.com/pricing.



Station Name (Customers)	Rank	Station Name (Subscribers)
Central Park S & 6 Ave	1	Pershing Square North
Centre St & Chambers St	2	E 17 St & Broadway
12 Ave & W 40 St	3	W 21 St & 6 Ave
Grand Army Plaza & Central Park S	4	Broadway & E 22 St
5 Ave & E 88 St	5	8 Ave & W 31 St
West St & Chambers St	6	West St & Chambers St
5 Ave & E 73 St	7	Broadway & E 14 St
Pier 40 - Hudson River Park	8	8 Ave & W 33 St
Central Park North & A.C.Powell Blvd	9	W 38 St & 8 Ave
Central Park West & W 72 St	10	Lafayette St & E 8 St
Old Fulton St	11	W 20 St & 11 Ave
South End Ave & Liberty St	12	Carmine St & 6 Ave
Washington St & Gansevoort St	13	W 41 St & 8 Ave
Little West St & 1 Pl	14	Christopher St & Greenwich St
Broadway & W 60 St	15	Great Jones St
Central Park West & W 76 St	16	E 47 St & Park Ave
5 Ave & E 78 St	17	Cooper Square & E 7 St
Central Park West & W 68 St	18	University Pl & E 14 St
Broadway & W 24 St	19	W 18 St & 6 Ave
Central Park West & W 85 St	20	Greenwich Ave & 8 Ave

Fig. 1. Spatial pattern of starting station chosen by short-term customers (left) and annual subscribers (right) overlaid on OpenStreetMaps network (data source: Citi Bike System Data).

cost per user per ride and the initial capital investment for infrastructure and hardware. Increasing ridership may raise both revenue and operating cost simultaneously, and it cannot be assured that the financial status of the operator will be improved. However, increasing short-term customers seems to be especially beneficial to operators due to their relatively high revenue per ride.

Kaviti et al. (2018) studied the impact of introducing a new pricing policy on ridership and revenue of Capital Bikeshare in the Washington metropolitan area. The implementation of the single ride fare (STF) promoted the usage of bike-share, which led to a 41% increase in casual riders. Kaviti and Venigalla (2019) examined the price sensitivities of bike-share users from the survey asking respondents to provide their demographic information and answer questions to measure their sensitivities. They concluded that the price reduction of STF and annual membership could benefit operator revenue.

Although research on pricing plans for bike-share services are limited, there are some studies looking at transit and taxi services. Nourinejad et al. (2017) analyzed pass programs and loyalty programs for public transportation riders. While pass programs ensure unlimited usage of systems, loyalty programs offer discounts when riders use systems more than a designated usage amount. Citi Bike has not provided a loyalty program, but there is a possibility that loyalty programs can be an alternative for customers who want to reduce the average travel cost. It may positively influence Citi Bike usage. According to Gandomi and Zolfaghari (2013), who studied the profitability of loyalty programs based on their analytic model, there are conditions which make a loyalty program more effective and benefit both users and firms. Chow (2014) also scrutinized the effect of providing electronic coupons on demand and the number of visitors at destinations, finding that it may not be beneficial to the profit of firms and urban economy but could increase ridership.

A few revenue management studies for bike-share systems focus on the cost for single trips but ignore unlimited passes. For example, Lin et al. (2017) conducted a stated preference survey to build a binary logit model of choosing a public bike system and found the best scenario to maximize the revenue. Ma et al. (2016) optimized usage and revenue using the average price for a trip. Chen et al. (2019) optimized the price and the availability rate of bikes to maximize the profit of the firm. Despite the existence of these studies, trip-level models are not appropriate to this research because passes allow limitless usage for a fixed period.

NYC Department of City Planning (2009) and Clayton et al. (2012) published bike-share feasibility study reports. They assumed specific demand levels and pricing plans to estimate revenue but did not consider responses of customers to various pass price levels. Great Rivers Greenway (2014) established a demand model based on observations from similar systems in other cities, reflecting residential, employment, transit, and recreational density.

As some reviewed studies demonstrated, revenue management strategies for bike-share systems tend to be based on individual trips, not multi-day plans. In fact, there are few studies in the urban transportation literature that deal with unlimited usage price plan strategies.

Our research is structured as follows. First, the basic information of Citi Bike public data is summarized to determine behavioral patterns of casual users in the system. Second, the distribution of the number of trips that casual customers take is estimated. This is necessary to relate trip level attributes with casual pass choices because such data is not publicly available. Parameters of these distributions are derived from sample means and standard deviations using a linear regression of daily short-term pass sales and the ridership of casual customers. Third, a pass choice model is constructed using the variables resampled from the fitted distributions via bootstrapping method. A multinomial logit model is used as it can delineate not only individual choice probabilities but also the market share if aggregated. Consequently, the pricing combination for two plans is optimized to maximize revenue based on the result of the estimated model and the impacts on consumer surplus are quantified.

3. Data preparation for case study

We set out to assemble data needed to estimate a casual pass choice model that would allow us to evaluate different pass pricing strategies. The system data from the Citi Bike website is open to the public for analysis, development, visualization and other activities, and was collected for September 2017 (Citi Bike, 2017). The research boundary is shown in Fig. 2, including three boroughs of NYC – Manhattan, Brooklyn, and Queens. Although Citi Bike provides their service in Jersey City, NJ, it is left separate from the NYC area because there is no direct bikeway connection between the two cities as riders would need to take the bikes onto ferries to cross the river.

Citi Bike Daily Ridership and Membership Data (Citi Bike, 2018) are also available and inform on the sales volumes for Citi Bike passes and membership. This provides details on the daily number of passes sold to customers and distinguish seasonal trends by day of the week and month of the year. Short-term passes are sold well on weekends compared to weekdays, but the patterns of 1-Day and 3-Day Passes are different; the 1-Day Pass shows more variation between weekend and weekday sales.

Ridership demand is severely influenced by the weather: ridership dramatically decreases in the winter because of the cold weather and snow. Nankervis (1999) studied the relationship between bike ridership and the weather condition and found that both rain and temperature impact travel patterns. El-Assi et al. (2017) also illustrated that high humidity, snow, and precipitation hinder bike-share system users from riding bikes. Since weather conditions are random, we filter out rainy days in September 2017 to have a more controlled data set to evaluate pricing strategies, using historical weather condition data from the AccuWeather website (AccuWeather, 2017).

Citi Bike Trip Histories includes trip duration, start/end time and station, locations of stations, user type, and gender as illustrated in Fig. 3. Because the system does not require people to create accounts, the data has no personal information like payment method, phone/cell contacts, address, or personal identifications. In addition, among the seven variables shown in Fig. 3, birth year and gender are not available among casual users. Start/end station information and bike ID are not useful explanatory variables for the casual pass choice model either, since that decision is not made per trip. Trip duration is excluded because we assume people who decide to use this service and buy a short-term pass will conduct the same length of trip regardless of the type of pass they purchased. As such, trip duration will not impact the choice of short-term pass. With all the variables in the open Citi Bike trip dataset eliminated, it is essential to bring other major factor to build a casual pass choice model. One key variable is the average cost per trip, which is not directly available from the data set.

The central point of our argument is as follows. The data gap illustrated here makes a disaggregate analysis for Citi Bike harder because there is no direct way to link every single trip into a trip chain, arranging trips to an identical casual pass user who conducted them. This is likely the case for any other MaaS transport operator with similar service provision. Given this data gap, what methodologies can MaaS operators like Citi Bike use to provide decision support on pass pricing policies?

We propose a methodology to address this data gap in Section 4. The output of this process would be used to estimate the number of trips that each casual customer took, which is required to model the choice behavior of short-term pass users.

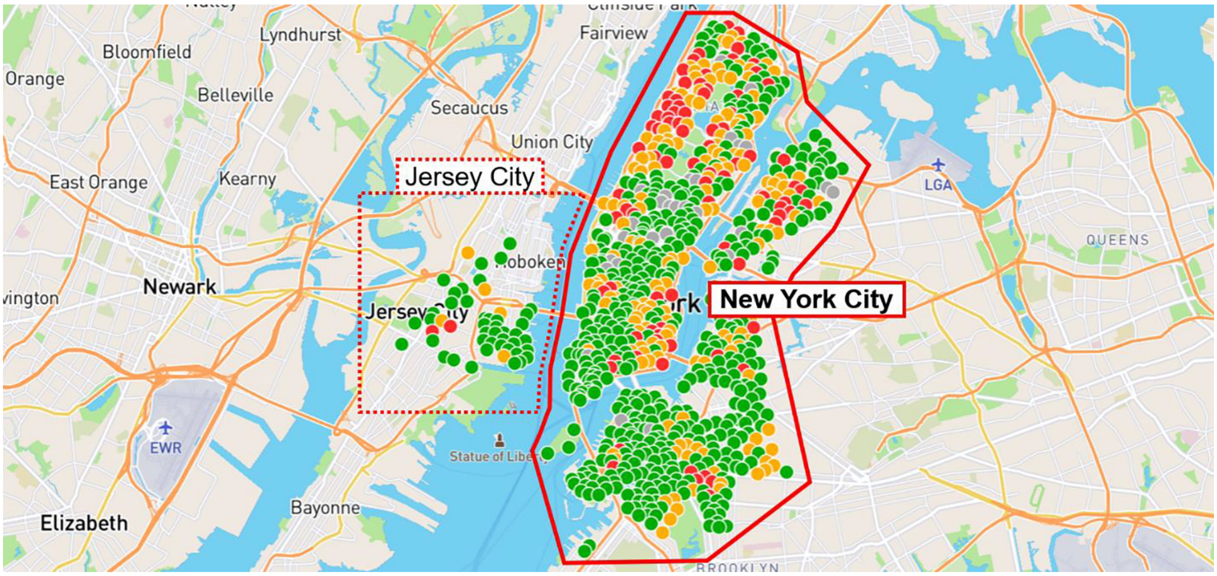
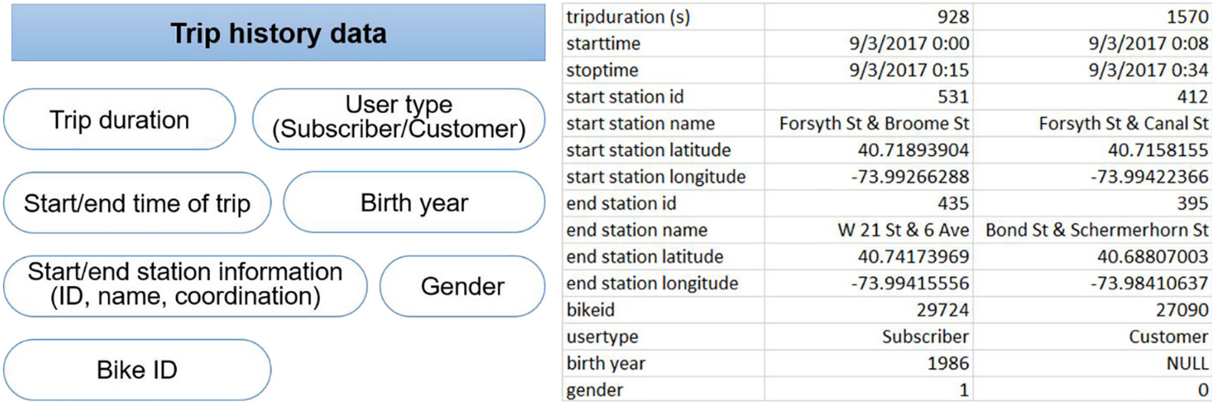


Fig. 2. Research boundary of Citi Bike stations.



tripduration (s)	928	1570
starttime	9/3/2017 0:00	9/3/2017 0:08
stoptime	9/3/2017 0:15	9/3/2017 0:34
start station id	531	412
start station name	Forsyth St & Broome St	Forsyth St & Canal St
start station latitude	40.71893904	40.7158155
start station longitude	-73.99266288	-73.99422366
end station id	435	395
end station name	W 21 St & 6 Ave	Bond St & Schermerhorn St
end station latitude	40.74173969	40.68807003
end station longitude	-73.99415556	-73.98410637
bikeid	29724	27090
usertype	Subscriber	Customer
birth year	1986	NULL
gender	1	0

Fig. 3. Citi Bike trip history data and example.

For the analysis, four weeks in September 2017, from the 3rd to the 30th, were sampled to excluded seasonal influence on bike usage to maintain a more controlled environment for demonstrating the methodology. During that period, precipitation of more than 1/4 inch was recorded on the 3rd and 6th and they were excluded, remaining 26 days in the sample data. The total number of passes sold was 68,714, of which 62,103 were 1-Day Passes, and 6611 were 3-Day Passes. Short-term customers made 208,363 trips in total.

4. Proposed cost per trip inference and bootstrapping procedure

To effectively estimate a casual pass choice model, we hypothesize that we need "average cost per trip" perceived by users. In this section we propose a method to derive this value. Clues for estimating average cost include daily sales of passes on Day i, $S_{i,1}$ for 1-Day passes and $S_{i,3}$ for 3-Day passes, and number of daily trips, N_i, conducted by customers. Customers who bought 3-Day Passes on a particular day i maintain access to the system until day $i+2$, which should be considered in the estimation.

An instrumental variable $A_{i,3}$ is introduced to represent the number of 3-day pass users with access to the system on day i, as defined in Eq. (1).

$$A_{i,3} = S_{i-2,3} + S_{i-1,3} + S_{i,3} \tag{1}$$

We propose the following linear regression model for N_i as a function of $S_{i,1}$ and $A_{i,3}$ shown in Eq. (2).

$$N_i = k_1 S_{i,1} + k_3 A_{i,3} \tag{2}$$

where k_1 and k_3 are coefficients of $S_{i,1}$ and $A_{i,3}$, meaning the marginal influence of each variable.

In this model, $S_{i,1}$ and $A_{i,3}$ are assumed independent from one another as the choice of customers are hardly affected by others who buy different passes on the same or different days. By estimating this model, we derive a key relationship used in our methodology. The coefficients k_1 and k_3 capture the average marginal increase in daily trips made by 1-Day ($S_{i,1}$) or 3-Day ($A_{i,3}$) pass users when the system sells one additional 1-Day *or* 3-Day Pass. The intercept of the model is not included because $S_{i,1} = A_{i,3} = 0$ will not generate any trips conducted by casual users.

We estimate the model using least squares and find that $\widehat{k}_1 = 2.797$ and $\widehat{k}_3 = 1.791$. The summary of the estimation is shown in Table 3 suggesting a good fit with an adjusted coefficient of determination of $R^2 = 0.9566$ and statistically significant coefficients.

Table 3
Linear regression result.

Regression Statistics	Multiple R	R Square	Adjusted R Square	Standard Error	Observations
	0.9992	0.9983	0.9566	377.9180	26
ANOVA	df	SS	MS	F	Significance F
Regression	2	2,015,523,760	1,007,761,880	7056.06656	8.53928×10^{-33}
Residual	24	3427729.162	142822.0484		
Total	26	2,018,951,489			
Model	Coefficients	Standard Error	*t*-Stat	*P*-value	
1-Day Passes	2.797	0.072	38.944	3.31669×10^{-23}	
3-Day Passes	1.791	0.256	6.988	3.16993×10^{-7}	

Fig. 4 indicates that the estimated regression model mimics the actual trip data well. As a comparison, the daily 1-Day pass sales curve in Fig. 5 presents a similar trend to Fig. 4 while the 3-Day pass remains uniform. Accordingly, we use the values of k_1 and k_3 as direct estimates of the average number of daily trips, which we define as $\widehat{\mu}_{1D,1}$ for 1-Day Pass users and $\widehat{\mu}_{3D,1}$ for 3-Day Pass users.

The dataset includes daily trips across 26 sunny days so $\sqrt{26}$ is multiplied with the standard errors to infer the sample standard deviations of the daily trips. This results in $s_{1D,1} = 0.072\sqrt{26} = 0.366$ and $s_{3D,1} = 0.256\sqrt{26} = 1.307$.

Note that $\widehat{\mu}_{3D,1}$ and $s_{3D,1}$ do not represent the whole 3-day period of Citi Bike use, but only one of the three days. We need to estimate the corresponding distribution parameters for number of trips made by users purchasing 3-Day passes across all three days, i.e. $\widehat{\mu}_{3D,3}$ and $s_{3D,3}$. These are estimated in Eqs. (3) to (5) assuming customers behave consistently across the three days. The calculated $\widehat{\mu}_{3D,3}$ and $s_{3D,3}$ are shown in Table 4.

$$\widehat{\mu}_{3D,3} = 3\widehat{\mu}_{3D,1} \tag{3}$$

$$VAR(3S_{i,3}) = 9VAR(S_{i,3}) \tag{4}$$

$$s_{3D,3} = \sqrt{VAR(3S_{i,3})} = 3\sqrt{VAR(S_{i,3})} = 3s_{3D,1} \tag{5}$$

Based on the estimated numbers of trips, the average cost is computed by dividing the price of the pass by the average number of trips in the period of the pass, resulting in \$4.29/trip for the 1-Day Pass and \$4.47/trip for the 3-Day Pass. The premium from the 3-Day pass likely reflects the added convenience of having to spend time purchasing other 1-Day passes if they need them. The values do not include extra time fees (\$4 for each additional 15 minutes) imposed on users who

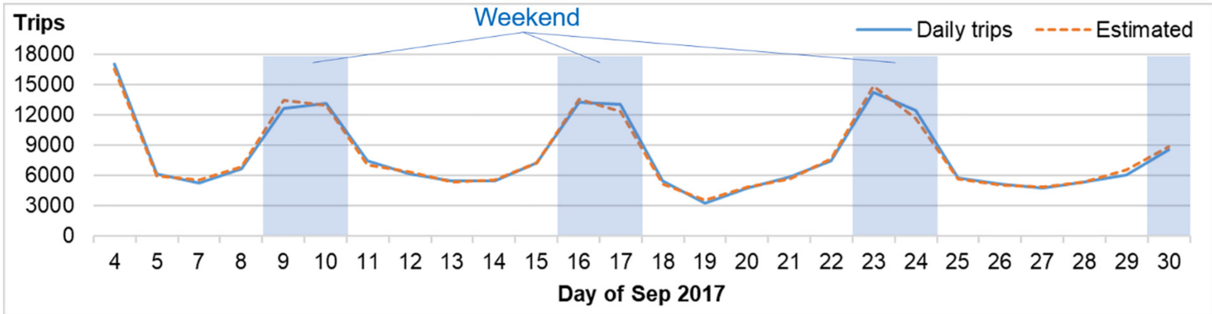


Fig. 4. Daily trips of short-term pass users and estimated trend by regression.

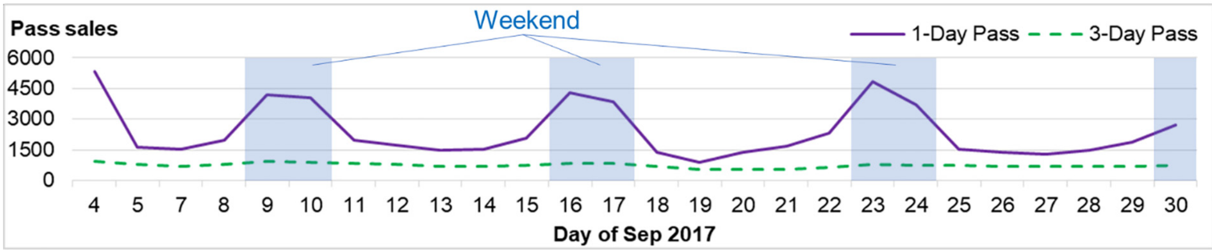


Fig. 5. Daily short-term pass sales.

Table 4
Statistic of average daily trips of pass users.

Statistic	1-Day Pass users	3-Day Pass users	Note
1. Mean of average daily trips	2.797 (p-value: 3.32×10^{-23})	1.791 (p-value: 3.17×10^{-7})	From regression
2. Standard deviation	0.072	0.256	(Adjusted R^2 = 0.957)
3. Mean of average trips during period of pass	2.797	5.373	Derived from 1.
4. Standard deviation of entire sample during period of pass	0.366	3.921	Derived from 2.
5. Average price per trip	\$12/2.797 trips = \$4.29/trip	\$24/5.373 trips = \$4.47/trip	-

exceed the 30-minute fixed period per trip. The exclusion of extra fee from average trip cost is reasonable because customers may not consider them when deciding between which pass to purchase. Moreover, the amount of the fee is the same for both 1-Day or 3-Day Pass.

With the estimated distribution parameters, we can now use bootstrap method to resample number of trips made. Resampled trip numbers are rounded to the nearest integer. Both groups of users are assumed to follow normal distributions due to huge sample sizes. For the 3-Day Pass users, a lognormal distribution is assumed to avoid negative values (for 1-Day Pass the variance is small enough not to sample negative values). Lognormal distribution parameters, a and b, have the following relationship with the mean $\widehat{\mu}_{3D,3}$ and variance $s^2_{3D,3}$ according to Forbes et al. (2010).

$$\widehat{\mu}_{3D,3} = \exp\left(a + \frac{b^2}{2}\right) \tag{6}$$

$$s^2_{3D,3} = \exp\left(2a + b^2\right)\left(\exp\left(b^2\right) - 1\right) \tag{7}$$

The resampled distributions of the numbers of trips are shown in Fig. 6. The total number of generated trips for each group are fixed to the number of samples, 62,103 and 6611. The numbers of trips are randomly distributed to customers of both passes, and average costs per trip are computed. A binary variable is included to the bootstrapped dataset to indicate the day that the passes are purchased, where 1 is used for weekend and 0 for weekdays.

The variables are prepared using the procedure described above – 1) the number of trips made during the pass periods, 2) average cost per trip, and 3) weekend indicator. For the 3-Day Pass alternative for users who chose 1-Day Pass, the number of trips is obtained by assuming that 1-Day Pass users would use Citi Bike 3 times more if one chose 3-Day Pass. One customer attribute that we exclude in our candidate variable set is the length of period staying in NYC. This information should impact a customer's likelihood of choosing 3-Day Pass if, for example, a customer stays shorter. Due to the lack of such data, however, we rely instead on the number of trips as a proxy of the stay time as there may be a positive correlation between two factors. The final bootstrapped data set is uploaded to https://github.com/BUILTNYU/CitiBikePass_Bootstrapped.

5. Pass choice model estimation from choice-based sample

A multinomial logit model (MNL) is adopted to model the choice behavior. Only the two types of passes are considered as choice alternatives conditional on the users choosing to purchase a pass. This is because no data is available from users who did not choose to purchase a pass for making trips, essentially resulting in a choice-based sample of only the two passes. Despite this limitation, the estimated model can always be adjusted to include other alternatives (see Ben-Akiva and Lerman, 1985) in a multinomial logit model if that data is made available. The model is shown in Eqs. (8) and (9).

$$P_n(1D) = \frac{\exp(V_{1D,n})}{\exp(V_{1D,n}) + \exp(V_{3D,n})} \tag{8}$$

$$P_n(3D) = 1 - P_n(1D) \tag{9}$$

where,
$P_n(1D), P_n(3D)$: probabilities of user n choosing 1-Day Pass (1D) and 3-Day Pass (3D)
$V_{1D,n}, V_{3D,n}$: representative utility functions of 1-Day Pass and 3-Day Pass for user n

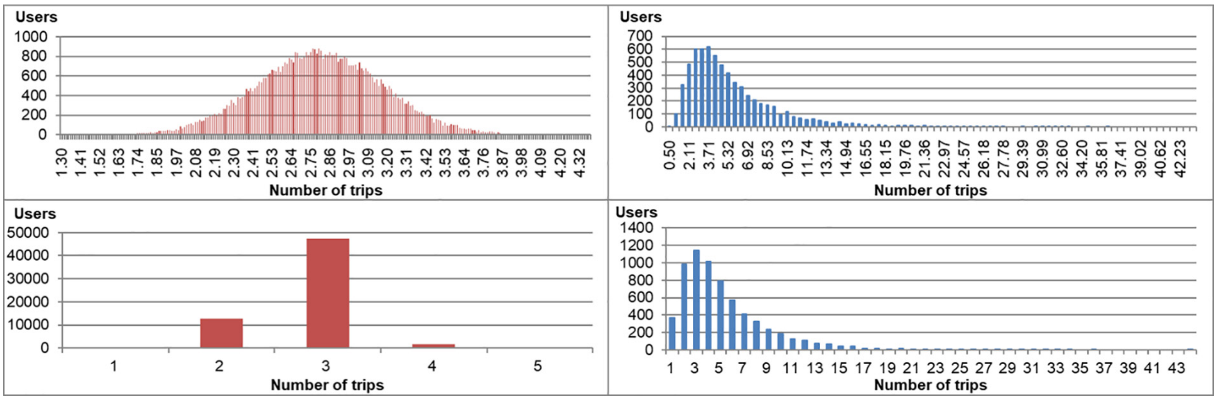


Fig. 6. Example histogram of bootstrapped distribution (up) and number of trips rounded to closest nonnegative integer (down) for 1-Day (left) and 3-Day (right) Pass users.

R, the free software for statistical analysis, is used to estimate the model. Because there are not many variables considered, all variable combinations with low correlations among variables are tried as candidates of the final model. Only one parameter is estimated for cost per trip and the weekend binary variable is included only in the 3-Day Pass alternative.

The estimated model presented in Table 5 is chosen as the best fitting model and the corresponding estimated utility functions are shown in Eqs. (10) and (11). Both the McFadden R^2 and the likelihood ratio test suggest a good fitting model. All coefficients are negative which means the increase in cost per trip or weekend use have negative effects on choosing the 3-Day Pass over the 1-Day Pass. The probabilities for choosing both alternatives may decrease as the average costs per trip for each alternative increase. Directions of these coefficients are reasonable, reflecting the basic relationship between price and demand. The tendency of choosing 3-Day Pass is reduced if a user accesses the Citi Bike system on weekends. This is also an acceptable result because weekend users expect to ride Citi Bike only a day when purchasing a 1-Day Pass rather than buying a 3-Day Pass and wasting one or two days.

$$V_{1D,n} = -2.884 \times \{\text{Cost per trip}_{1D,n}\} \tag{10}$$

$$V_{3D,n} = -7.384 - 2.884 \times \{\text{Cost per trip}_{3D,n}\} - 0.472 \times \{\text{Weekend}_n\} \tag{11}$$

This model is built to understand how short-term pass customers choose between the two passes. It assumes that users have prior knowledge of their variables when they make a choice. That is, when purchasing an unlimited-ride short-term pass, a customer will estimate the number of trips that they will make during a pass period and compare costs of passes to determine which one to buy.

To ensure reliability of the results, nine more models from different resampled datasets are estimated to compare the standard error from model estimation and standard deviation of estimated coefficients. As shown in Table 6, the discrepancies between them are negligible so the estimated coefficients are stable.

6. Revenue management application

Having established the pass choice model, the quantitative relationship between the average cost per trip and the probability of choosing an alternative is identified. As the short-term pass price affects users' behavior and market share of passes significantly, the fluctuation of the total revenue can be estimated under varying pass prices. Because the revenue is a nonlinear function of only two pass price variables, it is possible to graphically derive the optimal price set as shown in Fig. 6. As a benchmark, Citi Bike currently uses a price plan of \$12 for a 1-Day Pass and \$24 for a 3-Day Pass as shown in Table 2. For the price set of (\$24, \$12), the model forecasts demand of (6611, 62,103) and resulting revenue of \$903,900.

Fig. 7 is a 2-dimensional heat map of the revenue based on \$0.25 increments of each pass price based on the computed revenue. Darker cells represent an expected revenue that is higher than the midpoint while brighter cells are lower. The current benchmark prices are set as the upper bound prices for each pass since we do not have data on users who choose not to purchase the pass in the benchmark scenario. If we obtain a solution at the upper bound, it implies an optimal solution may be higher along that price dimension; otherwise, a discount is warranted.

Table 5
Estimation result of pass choice model by *R*.

Coefficients				
Number of samples: 68,714 (1-Day Pass users: 62,103, 3-Day Pass users: 6611)				
	Estimate	Standard error	*z*-value	Pr(>|z|)
Intercept ($V_{3D,n}$)	−7.384	0.071	−103.896	2.2×10^{-16}
Cost per trip ($V_{1D,n}$, $V_{3D,n}$)	−2.884	0.037	−78.746	2.2×10^{-16}
Weekend ($V_{3D,n}$)	−0.472	0.037	−12.681	2.2×10^{-16}
Significance				
Log-Likelihood (LLb)			−12,264	
Log-Likelihood with alternative specific constant (LLc)			−21,760	
Null Log-Likelihood (LL0)			−47,629	
McFadden R^2			0.436	
Likelihood ratio test			chisq = 18,992	

Table 6
Variation of model coefficients.

Variable	Set 1	Set 2	Set 3	Set 4	Set 5	Set 6	Set 7	Set 8	Set 9	Set 10	Standard deviation
Intercept ($V_{3D,n}$)	−7.384	−7.423	−7.378	−7.280	−7.477	−7.362	−7.509	−7.412	−7.252	−7.438	0.080
Cost per trip ($V_{1D,n}$, $V_{3D,n}$)	−2.884	−2.920	−2.891	−2.835	−2.934	−2.873	−2.957	−2.920	−2.837	−2.925	0.041
Weekend ($V_{3D,n}$)	−0.472	−0.475	−0.517	−0.471	−0.441	−0.460	−0.474	−0.525	−0.514	−0.505	0.028
McFadden R^2	0.436	0.431	0.440	0.429	0.437	0.436	0.438	0.434	0.423	0.438	-

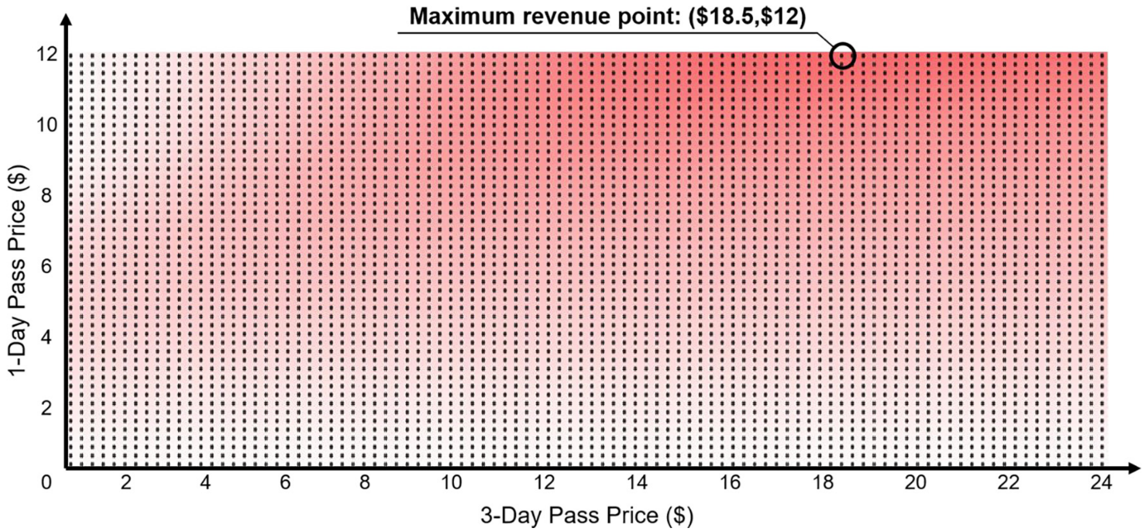


Fig. 7. Maximum revenue point as a function of prices.

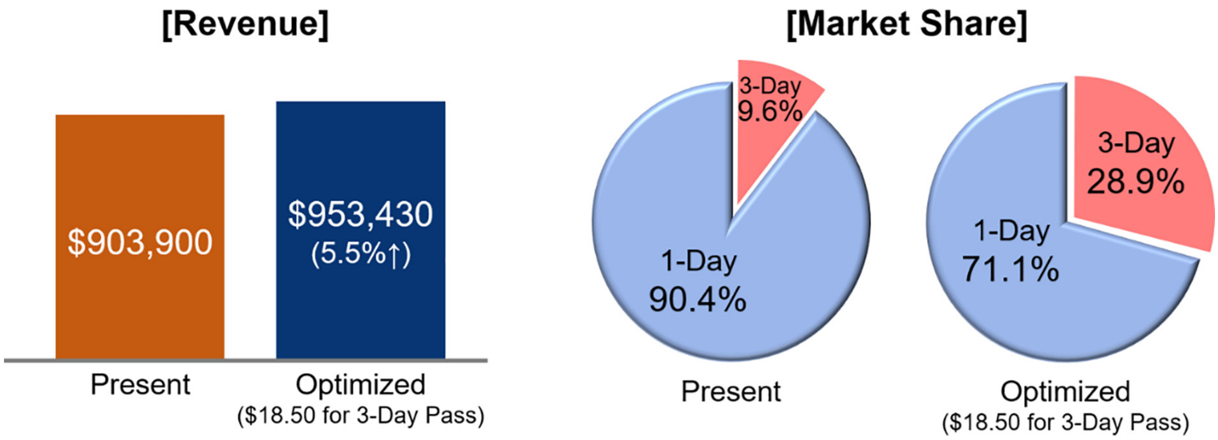


Fig. 8. Revenue and market share change estimation by price change.

The maximum revenue appears at the point (3-Day, 1-Day) = ($18.50, $12), which means that the revenue would be maximized for the existing customers if the Citi Bike system further applies a $5.50 discount on the price of 3-Day Pass from the current level of $24. This result may be acceptable because the average cost per trip paid by 3-Day Pass customers is a little higher than that of 1-Day Pass customers. The 1-Day Pass price of $12 is at the upper bound, so we can only conclude that the optimal price should be a minimum of $12.

The price set ($18.50, $12) can be compared to the benchmark of ($24, $12) for the existing customers, as summarized in Fig. 8. The revenue for 26 days, the number of days in dataset, is expected to increase from $903,900 to $953,430 for the sample, which reflects a 5.5% increase. This result is a lower bound increase because it does not consider the latent demand to Citi Bike system who can help raise even more revenue. The market share of the 1-Day Pass would shrink from 90.4% to 71.1%, which means 19.3% of 1-Day Pass customers would move on to 3-Day Pass and pay $6.5 more. In summary, despite the limitation of the public data, it is possible to show how revenue can be increased by at least 5.5% above the existing pricing policy, and potentially more if detailed trip surveys are conducted from potential users.

It is evident that lowering the price of the 3-Day Pass should benefit users due to the negative coefficients of the average cost per trip. We quantify the change in consumer surplus (Small and Rosen, 1981) as the expected maximum utilities among the choice set as shown in Eq. (12), where V_{in}^2 is the "after" scenario utility while V_{in}^1 is the "before" scenario utility.

$$\frac{1}{\mu_n}\ln\sum_{i\in C_n^2}e^{V_{in}^2}-\frac{1}{\mu_n}\ln\sum_{i\in C_n^1}e^{V_{in}^1} \tag{12}$$

where, μ_n: conversion of utility to a unit monetary value ($) for user n (absolute value), C_n^1, C_n^2: choice set of the user n before and after situation.

Since the choice model quantifies the choice relative to an average trip cost, the welfare impact can be measured in units of average trip cost using trip cost coefficient of 2.884. Using this coefficient with Eq. (12), the average of the expected consumer surplus over the 68,714 customers changes from −4.837 to −4.611, increasing by $0.23/customer. The average consumer surplus per trip per customer is also derived as $0.09/trip/customer.

Meanwhile, a discrete increment of $0.25 is chosen for pass prices to deliver simple numbers and make it more practical. If this discrete value is relaxed to allow continuous values, the exact solution is ($18.40, $12) with a revenue of $953,454. Furthermore, when optimizing the sum of total revenue and consumer surplus, the combination of pass prices leads to ($17.64, $12), earning a revenue of $951,386 and average consumer surplus of $0.10/trip/customer.

7. Conclusion

Presently, there is no literature on the choice behavior of short-term users to support design of pricing plans for bike-share systems around the world, and very limited numbers for MaaS systems. This study demonstrates how the availability of public data for these systems can help design pricing plans for them. Using data from Citi Bike in NYC, this research addressed three contributions.

1) Due to limited personal information of bike-share casual users, inference methods are introduced to make the data applicable for pass choice modeling. A linear regression model with an instrumental variable is estimated to break down number of total trips to trips by type of pass: 2.8 trips per day are made by 1-Day Pass users at an average cost of $4.29/trip, and 1.8 trips per day by 3-Day Pass users at an average cost of $4.47/trip. A bootstrap method is used to resample cost per trip per user given their observed pass choice.
2) A pass choice model is then estimated between the two passes, with ρ^2 = 0.44 and rejected likelihood ratio test of the model at 2.22×10^{-16} significance level. The cost per trip coefficient is −2.884 for utility functions of both passes. The purchase of 3-Day Pass is affected by a weekend usage variable with an estimated coefficient of −0.472.
3) A new pass pricing strategy that reduces the 3-Day Pass price from $24 to $18.50 is shown to increase revenue by at least 5.5% by modeling only the response from existing customers. Consumer surplus is expected to increase by at least $0.09/trip/customer.

This research serves as a case study in MaaS revenue management. It provides some very interesting insights on what can and cannot be done with Big Data availability. For example, casual users in MaaS systems will tend to fall into the category where personal information and trip chain information will not be available without additional survey collection. As such, the use of instrumental variables and bootstrapping can help to synthesize such attributes for each purchase. Because no survey data is available on potential customers who chose not to purchase the day passes, we cannot obtain an optimal pricing plan. Nevertheless, we show that it is possible to improve upon the benchmark pricing plan by at least 5.5% in revenue and $0.09/trip/customer in consumer surplus. These insights the value of this case study.

Several directions can be taken for future research. With more data provided by a bike-share company with stated preference surveys for alternative plans, the pass choice model can be expanded to more alternatives and a more comprehensive demand model can be estimated. Other bike-share systems can also be modeled to compare the estimated parameters. The pricing revenue management decisions can be combined with capacity and service coverage expansion decisions in more sophisticated network design optimization models. Alternatively, a symbiotic network design problem delivered by Chow and Sayarshad (2014) can be considered for a public agency to optimize bike lane investments considering their impact on bike-share ridership as Xu and Chow (2019) attempted.

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.

Acknowledgments

The authors were partially supported by C2SMART University Transportation Center.

References

AccuWeather, New York September Weather, 2017. https://www.accuweather.com/en/us/new-york-ny/10007/september-weather/349727?monyr=9/1/2017&view=table. Accessed on Nov 15, 2017.
Ben-Akiva, M.E., Lerman, S.R., 1985. Discrete Choice Analysis. MIT Press, Boston, MA.
Citi Bike. System Data. Retrieved from https://www.citibikenyc.com/system-data. Accessed on Nov 15, 2017.
Citi Bike. Citi Bike Membership & Pass Options. Retrieved from https://www.citibikenyc.com/pricing. Accessed on Jun 7, 2018.

Chen, Y., Wang, D., Chen, K., Zha, Y., Bi, G., 2019. Optimal pricing and availability strategy of a bike-sharing firm with time-sensitive customers. J. Cleaner Prod. 228, 208–221. https://doi.org/10.1016/j.jclepro.2019.04.151.

Clayton, S., Farber, C., Green, S., Kitzerow, E., Markfield, M., Song, I., White, C., Yang, Y., Zavacky, G., 2012. Feasibility Study for a Pittsburgh Bike Share. Heinz College-Public Policy and Management Report. Retrieved from https://bikepgh.org/wp-content/uploads/2012/11/BikeShareFinalReport.pdf. Accessed on May 28, 2018.

Chow, J.Y.J., 2014. Policy analysis of third party electronic coupons for public transit fares. Transp. Res. Part A: Pol. Pract. 66, 238–250. https://doi.org/10.1016/j.tra.2014.05.015.

Chow, J.Y.J., Sayarshad, H.R., 2014. Symbiotic network design strategies in the presence of coexisting transportation networks. Transp. Res. Part B 62, 13–34. https://doi.org/10.1016/j.trb.2014.01.008.

El-Assi, W., Mahmoud, M.S., Habib, K.N., 2017. Effects of built environment and weather on bike sharing demand: a station level analysis of commercial bike sharing in Toronto. Transportation 44 (3), 589–613. https://doi.org/10.1007/s11116-015-9669-z.

Forbes, C., Evans, M., Hastings, N., Peacock, B., 2010. Statistical Distributions. John Wiley & Sons Inc, Hoboken, NJ.

Gandomi, A., Zolfaghari, S., 2013. Profitability of loyalty reward programs: an analytical investigation. Omega 41 (4), 797–807. https://doi.org/10.1016/j.omega.2012.10.003.

Great Rivers Greenway, 2014. St. Louis Bike Share Study. Retrieved from https://altaplanning.com/wp-content/uploads/St.-Louis-Bike-Share-Report.pdf. Accessed on May 28, 2018.

Institute for Transportation Development Policy, 2018. The Bike-Share Planning Guide. (2018 ed.). Retrieved from http://bikeshare.itdp.org/. Accessed on July 28, 2018.

Kaviti, S., Venigalla, M.M., 2019. Assessing service and Price sensitivities, and pivot elasticities of public Bikeshare system users through monadic design and ordered logit regression. Transp. Res. Interdiscip. Perspect. 1,. https://doi.org/10.1016/j.trip.2019.100015 100015.

Kaviti, S., Venigalla, M.M., Lucas, K., 2019. Travel behavior and price preferences of bikesharing members and casual users: a Capital Bikeshare perspective. Travel Behav. Soc. 15, 133–145. https://doi.org/10.1016/j.tbs.2019.02.004.

Kaviti, S., Venigalla, M.M., Zhu, S., Lucas, K., Brodie, S., 2018. Impact of pricing and transit disruptions on bikeshare ridership and revenue. Transportation, 1–22. https://doi.org/10.1007/s11116-018-9904-5.

Lin, J., Wang, N., Feng, C., 2017. Public bike system pricing and usage in Taipei. Int. J. Sustain. Transp. 11 (9), 633–641. https://doi.org/10.1080/15568318.2017.1301601.

Ma, Y., Qin, X., Xu, J., Zou, X., 2016. Research on pricing method of public bicycle service: a case study in Guangzhou. In Service Operations and Logistics, and Informatics (SOLI). In: 2016 IEEE International Conference on July 10th, 156–161. https://doi.org/10.1109/SOLI.2016.7551679.

Nankervis, M., 1999. The effect of weather and climate on bicycle commuting. Transp. Res. Part A 33 (6), 417–431. https://doi.org/10.1016/S0965-8564(98)00022-6.

New York City Department of City Planning, 2009. Bike-Share Opportunities in New York City. New York, NY. Retrieved from https://www1.nyc.gov/assets/planning/download/pdf/plans/transportation/bike_share_complete.pdf. Accessed on May 28, 2018.

Nourinejad, M., Gandomi, A., Chow, J.Y.J., Roorda, M.J., 2017. Optimal discount policies for transit agencies: the case of pass programs and loyalty programs. In: Presented at First Triennial Conference of INFORMS Transportation Science and Logistics, Chicago, IL, Jul. 28, 2017.

Small, K.A., Rosen, H.S., 1981. Applied welfare economics with discrete choice models. Econometrica 49 (1), 105–130. https://doi.org/10.2307/1911129.

Xu, S.J., Chow, J.Y.J., 2019. A longitudinal study of bike infrastructure impact on bikesharing system performance in New York City. Int. J. Sustain. Transp., 1–16 https://doi.org/10.1080/15568318.2019.1645921.