A statistical analysis of bike sharing usage and its potential as an auto-trip substitute

Natalia Barbour, Yu Zhang*, Fred Mannering

Department of Civil and Environmental Engineering, University of South Florida, 4202 E. Fowler Avenue, ENG 207, Tampa, FL 33620, United States

ABSTRACT

Introduction: Bikesharing has become increasingly popular in urban areas as an alternative active transportation mode that can help relieve congestion, mitigate negative environmental impacts, and improve public health through increased physical activity. To understand the benefit of bikesharing, it is important to identify the factors influencing how often registered users use bikesharing, and assess whether and how much their bikesharing use is displacing an auto trip.
Methods: A survey was conducted, and random parameters logit models were estimated to study individuals' bikesharing usage rates and modal substitution. In addition to standard socio-de-mographic and travel behavior characteristics of the survey respondents, health-related questions were also included in the survey and health-related indicators were considered as explanatory variables in the estimated models.
Results: It was found that gender, age, income, household size, commute type and length, and vehicle ownership all played significant roles in bikesharing usage and modal substitution de-cisions. Regarding health measures, respondents' body mass index (BMI), one of the health-re-lated indicators, was also a significant predictor of bikesharing usage.
Conclusions: The outcomes of this research provide some initial insights into the bikesharing decision-making process that can help in the development of policies to improve the performance of bikesharing systems and making them a more viable transportation option.

1. Introduction

The concept of bikesharing has been around since 1960s, but only recently has it begun to receive large-scale acceptance as a viable transportation option (Fishman, 2016; Nikitas, 2018). However, the success of bikesharing systems can be highly variable because bikesharing is inherently tied to a geographical location that is defined by factors such as weather, urban density, local culture, or urban form design, and thus each bikesharing system has a set of unique characteristics. Depending on location, bike-sharing system could exhibit variable popularity, level of interest, and operational features. The level of utilization is generally estimated by trips per day per bike, in order to determine the number of bikes required in each system (Fishman, 2016). Smaller networks are at an inherent disadvantage relative to their larger counterparts because lower bicycle densities imply less convenient for potential users to find bikes. It has been shown that each location and bikesharing network attract different types of users, including commuters, students, local residents, or tourists (O'Brien et al., 2014). Depending on the trip purpose and type of user, the trip duration can vary. However, it has been found that, with the data from Melbourne, Brisbane, Washington, D.C., Minnesota and London, the average bikesharing trip lasts between 16 and 22 min (Fishman et al., 2014). Interestingly, many bikesharing users do

not seem to be regular users but rather use these programs as a complement to their primary mode of transportation (Fishman, 2016). Also, the factors influencing individuals' decision to participate in bikesharing can vary significantly. Because of the wide spectrum of factors (environmental and personal) influencing whether someone decides to use bikesharing, and the complexities of human decision making, understanding bikesharing use has been challenging.

In addition, the impacts of active transportation on public health have been extensively studied (Ewing et al., 2003; Frank et al., 2004; Giles-Corti et al., 2003; Saelens et al., 2003; Lopez, 2004). However, how the health conditions (or perceived health conditions) of travelers affect their decisions of travel and mode choice, especially regarding active transportation, is not yet fully understood. Past studies targeting cycling behavior indicated that physical capability is an implicit constraint in the choice of bicycle use (Stinson and Bhat, 2004; Smith and Kauermann, 2011; Ehrgott et al., 2012; Garcia-Palomares et al., 2012; Larsen et al., 2013; Habib et al., 2014; Wadud, 2014), although this has often not been considered explicitly (Philips et al., 2018). In recognition of the potential issues associated with not explicitly considering physical capability, Menghini et al. (2010) suggested the need to investigate the heterogeneity of cyclists in more detail. Further, McArdle (2010) pointed out that age, gender, body mass index, and levels of physical activity are all known to be key determinants of fitness and thus the capability to cycle, and Shaheen (2016) indicated that understanding the physical and behavioral casual factors associated bikesharing usage remains a key challenge. Thus, in the survey questionnaire designed for this research, the typical set of socio-demographic variables was expanded with health-related variables such as height, weight and self-reported health status. This data expansion was an attempt to minimize unobserved heterogeneity and potential omitted-variables bias in statistical-model estimation. It was hypothesized that body mass index (BMI) and overall well-being will be significant factors determining one's willingness to use an active transportation mode (such as bikesharing) on more regular basis, and that these factors will affect their likelihood of using bikesharing as a substitute for auto trips.

The survey questionnaire was disseminated through multiple channels such as CycleHop registered users list, University of South Florida mailing list, as well as social media. The questionnaires (distributed between February and April of 2018) incorporated a number of detailed questions relating to bikesharing, health, and socio-demographics. These collected data were then used to estimate two mixed logit models (random parameters logit models) addressing the frequency of bikesharing use and mode substitution while incorporating health-related factors.

The remainder of this paper begins with a literature review that focuses on various elements of bikesharing usage and modal substitution, followed by a detailed description of the survey, research design, methodological approach, and model estimation results. Finally, the paper concludes with a summary and discussion of key findings.

2. Bikesharing as a sustainable mode of transportation

Because biking does not involve harmful emissions while offering flexibility and convenience, it has become a valuable and environmentally friendly alternative to short auto trips in urban areas. The National Association of City Transportation Officials (2017) estimated that 25 percent more bikesharing trips were taken in 2016 than in 2015, and they also indicated that bikesharing growth is likely to continue in future years as more people recognize it as a low cost and health-inducing transportation option. Other research has found bikesharing to increase mobility, reduce transportation cost, mitigate traffic congestion, decrease fuel consumption, increase use of public transit, enhance environmental awareness, serve economic development, as well as improve health (Shaheen et al., 2016). Fishman et al. (2014) estimated that there was a significant reduction in motor vehicle use due to the presence of bikesharing systems. Their analysis was performed in the cities located in the United States, Great Britain, and Australia, and in every one of these cities a decrease in auto usage was found. Lu et al. (2018), who studied bikesharing in East Asia, reported benefits of bikesharing to include reductions in greenhouse gas emissions and fuel consumption, increased public transport use, improved accessibility, decreased traffic congestion and noise, lower travel cost, and increased physical activity and thus improved health and physical fitness. Such results have also been supported by other studies (Shaheen et al., 2010; Shaheen et al., 2013; Bauman et al., 2016; Pal and Zhang, 2017).

Bikesharing has also experienced significant growth in university-campus environments. This is because universities tend to have high population densities, large percentages of smart phone users, and extensive demands for shorter trips (between buildings on campus and to/from nearby student housing), all of which are potentially important ingredients for bikesharing success. Indeed, sustainability plans have become a concern in campus design and a bikesharing program is often a key element of such plans (Balsas 2003; Norton et al., 2007), because it can reduce traffic and parking congestion on and around campuses (Kaplan and Knowles, 2015).

With regard to the environmental factors, multiple studies have found that they play a significant role in willingness to use bikesharing (Nikitas, 2018). Some research revealed that proximity to the workplace or home tends to increase the usage of bikesharing systems (Shaheen et al., 2011; Molina-Garcia et al., 2015). In other work, Sun et al. (2017) studied the impact of environmental factors on bikesharing usage and found that traffic congestion did not influence the usage of bikesharing. On the other hand, bus accessibility was found to be positively associated with the usage of bikesharing, while metro accessibility was negatively associated with the usage. As expected, safety also plays a key role in bikesharing usage, and Sun et al. (2017) found that both on-street and off-street violent crimes tended to decrease the usage of bikesharing systems. Other studies found high population density, high levels of public transit accessibility, and the presence of upgraded facility types (bicycle lanes or bicycle paths), tended to increase the usage of bikesharing systems (Faghih-Imani et al., 2014; El-Assi et al., 2017).

Although, bikesharing systems in different locations could have different designs, scales, numbers and types of customers, they all share similarities with regard to user attitudes and perceptions. The discrepancy between the desire to use a bikesharing program and the actual use has to do with impediments, which could be self-imposed or based on factors out of the users' control, such as the

weather (Kaplan and Knowles, 2015). It should be noted that because the data in the current study will be drawn from bikesharing registrants in the state of Florida (with its highly favorable weather), the possibility to explore some of these impediments (such as weather) will be limited.

3. Socio-demographics of bikesharing users

Prior research has provided considerable insight into the relationship between socio-demographic characteristics and bikesharing usage. With regard to gender, Pucher et al. (2011) identified that about 65% to 90% of trips are done by men in countries where biking did not serve as a primary mode of transportation (US, UK, and Australia). In a study performed in London, less than 20% of bikesharing trips were made by females (Goodman and Cheshire, 2014). Akar et al. (2013) also found that women were less likely to ride a bicycle relative to men. In Netherlands, in contrast, more women than men use bicycles (Harms et al., 2014).

Where age is concerned, Buck et al. (2013) found that the users of bikesharing systems in Washington D.C. were, on average, younger than local cyclists. The average age for local cyclists was found to be 42 years old, whereas the average age for annual members of the bikesharing system and short-term users was 34 and 35 years old, respectively. In the U.S., Pucher et al. (2011), concluded that the number of 40 to 64-year-old cyclists increased the most of all the age groups that they studied between 2001 and 2009. During the eight years, cyclists in this age group doubled their share of bike trips.

Ethnicity has also been found to be an important factor determining whether an individual uses a bikesharing system. Studies in Washington D.C. and London found that the bikesharing population is not representative of the overall population composition of these cities (Buck et al., 2013; Fishman, 2016). Caucasians were over represented in the samples of bikesharing users relative to other ethnicities. Similarly, Borecki et al. (2012) found that bikesharing in Washington D.C. was largely undertaken by Caucasians.

With regard to income, prior studies found that people who use bikesharing had higher average income (Woodcock et al., 2014; Fishman et al., 2015; Fishman 2016). Also, Shaheen et al. (2014) found that bikesharing participants tended to be wealthier.

Another perspective on analyzing bikesharing adoption was undertaken by Gulsah et al. (2013). Their analysis was performed on the Ohio State University campus and was able to reveal some of the gender differences, as well as gender-based preferences and attitudes towards bikesharing. Although the surveyed population stayed in similar environments, women were found to feel less safe walking and biking (Gulsah et al., 2013). In other studies, traffic, lack of awareness of bike lanes, pedestrians, safety and campus design were found to be main impediments to bikesharing usage (Kaplan and Knowles, 2015). Similarly, Swiers et al. (2017) analyzed the cycling behavior of a university-student population and found that the two primary barriers to cycling were weather and safety.

Stinson and Bhat (2004) found a positive relationship between recreational cycling and cycling to commute. Moreover, Xing et al. (2010) found that 90% of those who cycled for commuting purposes were cycling for other purposes as well. These relationships in cycling behavior also suggest an association between cycling and other modes of transportation (Wuerzer and Mason, 2015).

There is also an extensive body of literature that links transportation and public health. The fact that active transportation modes help fight obesity and improve health has been addressed by many studies. Also, a correlation has been found between being overweight and living in less walkable communities (Ewing et al., 2003; Frank et al., 2004; Giles-Corti et al., 2003; Saelens et al., 2003; Lopez, 2004). Furthermore, Strum and Cohen (2004) found an association between urban sprawl in metropolitan areas and the prevalence of chronic diseases. Active transportation was also shown to significantly improve population health in California, with potential decreases in chronic diseases (Maizlish et al., 2017).

Other researchers have analyzed the connection between active transportation, health, and the usage of social networking services. For example, Hong et al. (2018) found that intensive users of social networking services were more likely to be obese, and tended to spend less time walking, making this group a natural target for interventions designed to increase physical activity.

Past research has shown that expanding the set of variables could be essential to more fully understand bikesharing behavior and developing strategies for bikesharing implementation and adoption. For example, Earl and Lewis (2018) suggested examining the role of context in health behavior and emphasized the importance of considering the environment while trying to influence health behavior.

In the current research, in addition to traditional socio-demographic characteristics, travel behavior, and travel history variables, the body mass index (BMI) will be considered as an explanatory variable. The BMI gives an estimation of excess body weight. Although it is not a direct estimate of body fatness, some studies have confirmed that BMI does correlate with body fat measurements, which may affect individual's physical mobility and furthermore the willingness to engage in active transportation. Body mass index does have its limitations because of natural variances across factors such as age, gender, ethnicity and body composition (BMI does not distinguish between excess fat, muscle, water or bone). However, BMI is non-invasive, easily calculated, and despite these limitations, it has been widely recognized as a good overall predictor of morbidity, mortality and a good assessment of individual's health status.

4. Survey and research design

A web-based survey was designed to collect the data on the bikesharing usage of registered bikesharing users. The survey dissemination took place between February and April of 2018. To make sure that a wide variety of demographic groups was reached, multiple distribution channels to disseminate the survey were used. CycleHop Bike Share Company, which operates bikesharing programs in Tampa, St. Petersburg, Orlando and the University of South Florida (Tampa campus) assisted in distributing the survey to its registered users as well as posting it on social media. To increase the number of responses, the survey was also sent to the students and faculty of the University of South Florida Tampa campus (where one of the bikesharing programs is operating) through an on-

campus email list. Respondents were asked about their use frequency of bikesharing; less than once a month, 1 to 3 times per month, 4 to 5 times per month, 6 to 10 times per month, and more than 10 times per month. Because the literature review concluded that bikesharing users in the United States do not generally use bikesharing on a regular basis, the survey did not ask about the actual number of uses but rather a usage category. Based on the number of observations in each category, the data were divided into two groups; one group indicating that they typically use bikesharing less than once a month, and the other group indicating that they typically use bikesharing once a month or more. Of the 301 registered bikesharing users, 165 fall in the first group and 134 the second.

The second question in the survey focused on mode substitution. That is, which mode of transportation would the respondent use if bikesharing was not available on a trip that they had chosen to ride a shared bike. Because the substitution of auto trips is critical in terms of environmental impacts and traffic congestion mitigation in urban areas, the focus was on identifying the characteristics of the group who would make their bikesharing trip by auto if bikesharing was not available on a trip they chose to use shared bikes. Out of 301 respondents, 140 indicated that they would use an automobile in the absence of bikesharing while the remaining 161 would use other modes including bus, personal bicycle, or walking.

The survey covered a variety of socio-demographic and household characteristics, as well as travel behavior and travel history characteristics (commute time and distance, traffic-crash history, parking time, grocery store proximity, total daily travel time, and so on). Furthermore, health-related questions such as weight, height and self-assessed health were added. Given the responses, the body mass index (BMI) was calculated using self-reported height and weight. In the sample, 175 respondents had a normal BMI (BMI equals to 25 or less), 81 people were classified as overweight (BMI between 25 and 30), and 45 respondents were classified as obese (BMI greater than 30). The respondents were asked to assess their health on the following scale: extremely bad, slightly bad, neither good nor bad, good, extremely good. In the collected sample only 1% reported their health as extremely bad, followed 3% as slightly bad, 4% as neither good nor bad, whereas 61% indicated good health and 31% extremely good. Respondents were also asked if they struggled with any illness or health condition on daily bases, and only 34 of the 301 respondents indicated such a struggle. Because of the exploratory nature of incorporating the health questions and potential issues with confidentiality regarding health information, the type of illness or health condition was not specified.

To get a sense of the respondent sample, Table 1 provides summary statistics for selected respondent attributes.

5. Methodological approach

In this study, two questions were considered; whether the survey respondent bikeshares one or more times per month (monthly usage), and whether the respondent would make an auto trip if bikesharing was not available on a trip they had chosen to bikeshare.

The above responses are discrete with a yes/no response indicating either monthly usage or auto trip substitution. To arrive at an estimable statistical model for both questions, a function that determines the probability of either using one or more times per month or substituting an auto trip (1 if the respondent is a monthly bikesharing user/substituting a bikesharing trip by an auto trip, 0 if not) was defined as,

$$F_n = \boldsymbol{\beta} \mathbf{X}_n + \varepsilon_n \qquad (1)$$

where $\mathbf{X}_n$ is a vector of explanatory variables that affect the probability of observation n being a monthly bikesharing user/substituting a bikesharing trip, $\boldsymbol{\beta}$ is a vector of estimable parameters, and ε_n is a disturbance term. If the disturbance term is assumed to be generalized extreme-valued distributed, a standard binary logit model results as (McFadden, 1981)

$$P_n(1) = \frac{1}{1 + \mathrm{EXP} - (\boldsymbol{\beta} \mathbf{X}_n)} \qquad (2)$$

where P_n (1) is the probability of the respondent being a monthly user/substituting a bikesharing trip, and other variables are as previously defined.

In model estimation, it is essential to account for the possibility of unobserved heterogeneity across respondents. That is, the possibility that different respondents will be affected by explanatory variables differently due to unobserved reasons (this is particularly likely while analyzing complex human decision-making processes). To account for the possibility of having one or more parameter estimates in the vector $\boldsymbol{\beta}$ vary across respondents, a distribution of these parameters can be assumed, and Eq. (2) can be rewritten as (Washington et al., 2011)

Table 1
Some key survey statistics.

Respondent Characteristic	Mean	Standard Deviation
Age (in years)	38	13.8
Height (in inches)	67.9	4.1
Weight (in pounds)	165	37.6
Body Mass Index	25.1	4.8
Household size (persons)	2.2	1.1
Household vehicle ownership (vehicles)	1.87	1
Annual household income (in dollars)	85,000	58,000

$$P_n(1) = \int \frac{1}{1 + EXP - (\boldsymbol{\beta X}_n)} f(\boldsymbol{\beta}|\boldsymbol{\varphi}) d\boldsymbol{\beta} \tag{3}$$

where $f(\boldsymbol{\beta}|\boldsymbol{\varphi})$ is the density function of $\boldsymbol{\beta}$, $\boldsymbol{\varphi}$ is a vector of parameters describing the density function (mean and variance), and all other terms are as previously defined. The resulting model is referred to as random parameters or mixed logit model (see Mannering et al., 2016, for a description of alternate methods of accounting for unobserved heterogeneity).

In the model estimation, the possibility for the mean and variance of individual parameters to be a function of explanatory variables is also considered giving (Seraneeprakarn et al., 2017; Behnood and Mannering, 2017; Mannering, 2018),

$$\boldsymbol{\beta}_n = \boldsymbol{\beta} + \Theta \mathbf{Z}_n + \sigma_n EXP(\boldsymbol{\omega}_n \mathbf{W}_n) + \varphi_n \tag{4}$$

where β is the mean parameter estimate, $\mathbf{Z}_n$ is a vector of explanatory variables that influence the mean of β_n, Θ is a vector of estimable parameters, $\mathbf{W}_n$ is a vector of explanatory variables that captures heterogeneity in the standard deviation σ_n, $\boldsymbol{\omega}_n$ is the corresponding parameter vector, and φ_n is a randomly distributed term that captures unobserved heterogeneity across respondents.

Estimation of the random parameters logit model was undertaken by simulated maximum likelihood approaches because the required integration of the logit formula over the distribution of parameters is not closed form. Prior research has shown that Halton draws can deliver more efficient distribution of simulation draws than purely random draws (McFadden and Ruud, 1994; Bhat, 2003), and 1000 Halton draws were used in the estimation process. This is a number that has been shown to be more than enough to provide accurate parameter estimates (Bhat, 2003; Milton et al., 2008; Anastasopoulos and Mannering, 2009; Behnood and Mannering, 2016). In this study, the normal distribution was used for random parameters because it provided the best statistical fit for both response models (other distributions such as the log-normal, uniform, and exponential were not found to produce statistically better results than the normal distribution). It should be noted that additional approaches to address the unobserved heterogeneity have been widely applied in accident and injury-severity research (Behnood and Mannering, 2016; Osama and Sayed, 2017; Fountas and Anastasopoulos, 2018; Fountas et al., 2018a, 2018b; Marcoux et al., 2018; Balusu et al., 2018).

Marginal effects were calculated to determine the effect that individual explanatory variables have on response probabilities. The marginal effect of an explanatory variable gives the effect that a one-unit increase in an explanatory variable has on the response probabilities. For indicator variables (that assume values of zero or one), marginal effects will give the effect of the explanatory variable going from zero to one (Washington et al., 2011).

6. Model estimation results

Table 2 presents the summary statistics of variables found to be statistically significant in both models. Table 3 and Table 4 provide the random parameters logit model estimation results, including parameter estimates, t-statistics and marginal effects, for the usage of bikesharing and auto-mode substitution, respectively. The statistically significant explanatory variables in Table 3 and Table 4 were grouped into three categories; socio-demographic factors, travel behavior and history, and health indicators.

As shown in Tables 3 and 4, two variables in each model produced statistically significant random parameters. This significance was confirmed by conducting a likelihood ratio test to compare the random parameters logit model with fixed parameters model. For both models (as shown in Tables 3 and 4) the test rejected the null hypothesis that fixed and random parameters models are the same with over 95% confidence. Thus, only the results of random parameters models are presented.

All explanatory variables are in the "Yes" response functions (use bike sharing once a month or more/substituting a bikesharing trip by an auto trip) with the "No" response functions (for both models) implicitly set to zero. Also, estimation results indicate that no variables produce an estimated parameter with statistically significant heterogeneity in the means and/or variances, so Eq. (4) reduces to $\beta_n = \beta + \varphi_n$.

Table 2
Summary statistics for variables included in final model estimations.

Variable Description	Mean	Standard Deviation
Male indicator (1 if respondent is a male, 0 otherwise)	0.46	0.50
Caucasian indicator (1 if respondent is Caucasian, 0 otherwise)	0.76	0.43
Younger millennial indicator (1 if respondent is less than 30 years old, 0 otherwise)	0.53	0.50
Low annual household income indicator (1 if annual household income is less than $50k, 0 otherwise)	0.35	0.48
One-person household indicator (1 if respondent lives alone, 0 otherwise)	0.24	0.43
High annual household income indicator (1 if annual household income is more than $200k, 0 otherwise)	0.12	0.33
Drive-alone commute indicator (1 if respondent most often commutes to work by driving alone, 0 otherwise)	0.72	0.45
Lack of commute indicator (1 if respondent does not commute, 0 otherwise)	0.06	0.24
Daily travel time indicator (1 if respondent spends 90 min or more on total daily travel, 0 otherwise)	0.08	0.28
Higher vehicle ownership (1 if household owns or leases three or more vehicles, 0 otherwise)	0.21	0.41
Low average parking time indicator (1 if respondent spends less than 5 min total on finding a spot and walking to their destination, 0 otherwise)	0.66	0.47
Low parking time indicator (1 if respondent spends less than 3 minutes on finding a parking spot during a normal trip, 0 otherwise)	0.79	0.41
High BMI (body mass index) indicator (1 if respondent has BMI above 25, 0 otherwise)	0.42	0.49
Obese BMI (body mass index) indicator (1 if respondent has BMI above 30, 0 otherwise)	0.15	0.35

Table 3

Random parameters logit model estimation results for the probability of using bikesharing one or more times per month (all random parameters are normally distributed).

Variable Description	Estimated Parameter	t-Statistic	Marginal Effect
Constant	1.16	2.72	
Socio-demographic factors			
Male indicator (1 if respondent is a male, 0 otherwise)	0.72	2.77	0.17
Caucasian indicator (1 if respondent is Caucasian, 0 otherwise)	0.50	1.86	0.12
Low annual household income indicator (1 if annual household income is less than $50k, 0 otherwise) (*Standard deviation of parameter distribution*)	-1.03 (*4.60*)	-3.08 (*5.79*)	-0.25
One-person household indicator (1 if respondent lives alone, 0 otherwise)	0.53	1.91	0.13
Travel behavior and history			
Drive-alone commute indicator (1 if respondent most often commutes to work by driving alone, 0 otherwise)	-1.85	-5.30	-0.45
Lack of commute indicator (1 if respondent does not commute, 0 otherwise)	-1.51	-2.86	-0.37
Daily travel time indicator (1 if respondent spends 90 minutes or more on total daily travel, 0 otherwise)	1.19	2.69	0.29
Higher vehicle ownership (1 if household owns or leases three or more vehicles, 0 otherwise)	-0.64	-2.05	-0.16
Low average parking time indicator (1 if respondent spends less than 5 minutes total on finding a spot and walking to their destination, 0 otherwise)	-0.97	-3.55	-0.24
Health indicators			
High BMI (body mass index) indicator (1 if respondent has BMI above 25, 0 otherwise) (*Standard deviation of parameter distribution*)	0.30 (*1.87*)	1.21 (*5.08*)	0.07
Number of observations	301		
Log likelihood at zero	-232.20		
Log likelihood at convergence	-174.22		

Table 4

Random parameters logit model estimation results for the probability that a bikesharing trip would be substituted by an auto trip if bikesharing was not available (all random parameters are normally distributed).

Variable Description	Estimated Parameter	t-Statistic	Marginal Effect
Constant	-0.22	-0.59	
Socio-demographic factors			
Male indicator (1 if respondent is male, 0 otherwise) (*Standard deviation of parameter distribution*)	-0.82 (*5.30*)	-2.68 (*5.97*)	-0.20
Younger millennial indicator (1 if respondent is less than 30 years old, 0 otherwise)	0.55	2.17	0.13
High annual household income indicator (1 if annual household income is more than $200k, 0 otherwise)	-0.69	-1.84	-0.17
Travel behavior and history			
Drive-alone commute indicator (1 if respondent most often commutes to work by driving alone, 0 otherwise)	1.31	4.63	0.32
Low parking time indicator (1 if respondent spends less than 3 minutes on finding a parking spot during a normal trip, 0 otherwise)	-1.14	-3.59	-0.28
Higher vehicle ownership (1 if household owns or leases three or more vehicles, 0 otherwise)	-0.53	-1.86	-0.13
Health indicators			
Obese BMI (body mass index) indicator (1 if respondent has BMI above 30, 0 otherwise) (*Standard deviation of parameter distribution*)	1.40 (*2.32*)	3.18 (*3.29*)	0.34
Number of observations	301		
Log likelihood at zero	-208.6		
Log likelihood at convergence	-181.6		

6.1. Model estimation results: regular use of bikesharing

With regard to the socio-demographic factors affecting the probability of registered bikesharing users using shared bikes one or more times per month (see Table 3), it was found that male respondents were more likely to be regular bikesharers (use it once a month or more). This finding aligns with prior research stating that males are more likely to use bikesharing in general (Pucher and Buehler, 2012; Goodman and Cheshire, 2014; Akar et al., 2013). The average marginal effect indicates that males have a 0.17 higher probability of using shared bikes one or more times per month relative to females.

Respondents who identified themselves as Caucasian were found to be more likely to use bikesharing regularly. This result also aligns with prior studies that found Caucasians to be over represented in samples of bikesharing users relative to other ethnicities (Buck et al., 2013; Fishman, 2016; Borecki et al., 2012). Households with an annual income below $50,000 produced a normally distributed parameter with a mean of -1.03 and a standard deviation of 4.60. This results in 58.9% of respondents from these households being less likely to use bikesharing one time a month or more, and 41.1% respondents being more likely to do so (relative to respondents from households making $50,000 or more per year). This finding is important because it shows considerable variation

among lower-income households. While previous studies found that bikesharing membership and usage is usually associated with higher incomes (Fishman et al., 2015; Fishman 2016; Woodcock et al., 2014; Shaheen, et al., 2014), the variance that was found in this effect shows that some respondents in lower-income households have higher bikesharing usage than their higher-income counterparts. Thus, bikesharing in lower socioeconomic areas could be viable and help improve equality and mobility of the most vulnerable members of the society. The variation in this effect across low-income respondents also suggests that there are factors relating to low-income respondents that are not captured by income alone (reflected by the significant unobserved heterogeneity).

Respondents from single-person households were found to be more likely to be regular users compared to respondents from households with multiple occupants. This finding could be related to the presence of children in the household. Intuitively, the presence of children makes it harder for an individual to use a bicycle in general. Lack of the appropriate bike seats for small children coupled with the vehicle-dominant facilities do not encourage but rather discourage bikesharing use among individuals with small children. It is important to stress the fact that most environments and bikesharing systems do not cater to the caregivers of small children, especially in a context of equality and equity in transportation.

With regard to travel behavior and history, the indicator variable for commuters who mostly commute by driving alone and those who do not commute at all were found to be less likely to bikeshare regularly. For drive-alone commuters, the average marginal effect is quite large (in absolute terms) at -0.45 indicting that drive-alone commuters have a 0.45 lower probability of bikesharing one or more times per month than non-drive-alone commuters. Additionally, respondents who spent more than 90 min on total daily travel were found to be more likely to use bikesharing once a month or more, again with a relatively large average marginal effect of 0.29. As might be expected (reflecting the ease of vehicle access and usage), respondents from households with higher vehicle ownership (owning or leasing three or more vehicles) as well as those whose average parking time for their most regular trip that is less than 5 minutes (including finding a spot and walking to the destination) were less likely to use bikesharing one or more times a month.

Finally, respondents with BMI scores over 25 produced a normally distributed parameter with a mean 0.30 and a standard deviation of 1.87. This results in 56.4% respondents more likely to be a regular bikesharing user and 43.6% of respondents with high BMI being less likely, relative to their lower BMI counterparts. The fact that higher BMI respondents have higher usage probabilities than some of their lower BMI counterparts shows that bikesharing has some significant potential for improving public health. The fact that BMI was found to be significant factors in the model again underscores the importance of health-related factors in considering active-transportation modes such as bikesharing.

To assure that the model with inclusion of the high BMI indicator provides statistically better fit for the data, a model without this variable was estimated. A likelihood ratio test comparing models with and without the BMI variable indicates that the null hypothesis that the models are the same can be rejected with 93% confidence. Also, for bikesharing usage there is the possibility that BMI could be an endogenous variable. That is, respondents who have high bikesharing usage rates may lower their BMI. However, in this case it is unlikely that the bikesharing usage rates are high enough to directly affect BMI, although some caution should still be exercised when interpreting our results in this regard.

6.2. Model estimation results: auto-trip substitution

With regard to the socio-demographic variables influencing the probability that a bikesharing trip would be substituted by an auto trip if bikesharing was not available (see Table 4), it was found that people who identified themselves as male produced a normally distributed parameter with a mean -0.82 and a standard deviation equal to 5.30. This suggests considerable heterogeneity among male respondents. The estimation results imply that 56.1% of males who use bikesharing were less likely to make the trip by auto if bikesharing was not available and 43.9% being more likely to do so. Once again, gender was found to play a key role in bikesharing-related behavior. The findings are generally consistent with prior studies that found males to be more likely to use bikesharing in general (Pucher and Buehler, 2012; Goodman and Cheshire, 2014; Akar et al., 2013), but the considerable heterogeneity among male bikesharing users with regard to their substituting a bikesharing trip with an auto trip is an interesting finding. With regard to age, it was found that bikesharing respondents who are less than 30 years old were more likely to make an auto trip in the absence of bikesharing. It is noteworthy that other researchers have also found age to be a significant variable in bikesharing (Pucher et al., 2011; Buck et al., 2013). Respondents with annual household income above $200,000 were found to be less likely to substitute their bikesharing trip by an auto trip in the absence of bikesharing. This finding suggests that high-income households are less likely to increase their auto usage and they are more likely switch to another mode of active transportation in the absence of bikesharing. Prior studies also found income to be a significant variable while analyzing bikesharing. People who used bikesharing were found have higher average income (Woodcock et al., 2014; Fishman et al., 2015; Fishman, 2016).

With regard to travel behavior, respondents who commute by driving alone were found to be more likely to substitute their bikesharing trip by an auto trip in the absence of bikesharing. The high average marginal effect of this variable indicates that respondents that most often drive alone have a 0.32 higher probability of substituting their trip by auto relative to respondents that regularly commute by other means. This finding is like that found in the previous bikesharing usage model (see Table 3) and reflects the substantial residual effect of the auto culture among bikesharing registrants. Respondents who indicated a very low time (less than 3 minutes) to find a parking spot during their most regular trip and those whose households owned or leased three more vehicles, were found to be less likely to use an auto trip in the absence of bikesharing.

Regarding the health indicators, respondents who had body mass index in the obese range (BMI above 30) produced a normally distributed parameter with a mean 1.40 and a standard deviation 2.32. This shows considerable variation across the population with regard to the effect of BMI, with 72.7% of people with the obese BMI being more likely to substitute their bikesharing trip with an auto trip if bikesharing was not available and 27.3% being less likely.

Like the previous model on the usage of bikesharing, a separate model without the BMI indicator was estimated to underscore the statistical importance of the BMI indicator. A likelihood ratio test comparing the models with and without the BMI variable indicated that the hypothesis that the two models were equal could be rejected with over 99% confidence.

7. Summary and conclusions

This research focuses on exploring the determinants of bikesharing use, and its potential as an auto-trip substitute, by including self-reported health factors. Both estimated statistical models provide insights into how various survey respondents behave with regard to bikesharing decisions. For the frequency-of-use model it was found that Caucasian males, respondents from one-person households, and those with high total daily travel times (for all trips) were more likely to be a regular user of bikesharing (use it at least once a month). In contrast, respondents who drove alone for their commute trip and those who do not commute at all were less likely to bikeshare regularly. Also, respondents from households with higher auto ownership (leased or owned at least three vehicles) and low average parking time during their most regular trip were less likely to use bikesharing at least once a month. Variables that were found to vary across respondents included low annual household income (below $50,000) and the high body mass index (BMI) indicators.

With regard to the auto-mode substitution model (asking if a respondent would make an auto trip if bikesharing was not available), younger respondents (under 30 years old) were found more likely to make an auto trip in the absence of bikesharing. In contrast, those from households with annual household income more than $200,000 were found to be less likely to make an auto trip in the absence of bikesharing. Respondents who identified themselves as male were less likely to exhibit homogenous behavior and this parameter varied across population. With regard to travel behavior, it was found that respondents who commuted by driving alone were more likely to make an auto trip if bikesharing was not available. In contrast, those who spent less than 3 minutes to find parking for their most regular trip and those whose households owned or leased three or more vehicles were less likely to make an auto trip. Obese BMI indicator (BMI above 30) was found to vary across population, which reflected the willingness of a percentage of this group being less likely to make an auto trip in the absence of bikesharing. This is important because it suggests that some people with obese BMI are willing to improve their health through participating in active transportation.

The results of this paper can potentially help guide and develop our understanding of how bikesharing decisions are made. Household composition and vehicle ownership were found to be some of the key factors in decisions related to bikesharing behavior. It was also found that the lingering effects of auto reliance (reflected by respondents who indicated that most often commuted by driving alone) adversely affected the likelihood of a registered bikesharing user using bikesharing frequently or substituting their bikesharing trip with a non-auto mode. Finally, the model estimations did not show that self-reported health-related factors other than BMI played a significant role in bikesharing use and behavior. While the self-reported health question was unable to produce statistically significant results, variables derived from actual detailed health data may still prove valuable in future research on bikesharing behavior.

Financial disclosure

Disclaimer

References

Akar, G., Fischer, N., Namgung, M., 2013. Bicycling choice and gender case study: the Ohio State University. Int. J. Sustain. Transp. 7 (5), 347–365.
Anastasopoulos, P., Mannering, F., 2009. A note on modeling vehicle accident frequencies with random-parameters count models. Accid. Anal. Prev. 41 (1), 153–159.
Balsas, C., 2003. Sustainable transportation planning on college campuses. Transp. Policy 10 (1), 35–49.
Balusu, S., Pinjari, A., Mannering, F., Eluru, N., 2018. Non-decreasing threshold variances in mixed generalized ordered response models: a negative correlations approach to variance reduction. Anal. Methods Accid. Res. 20, 46–67.
Bauman, A., Crane, M., Drayton, B., Titze, S., 2016. The unrealised potential of bike share schemes to influence population physical activity levels: a narrative review. Prev. Med. 103, S7–S14.
Behnood, A., Mannering, F., 2016. An empirical assessment of the effects of economic recessions on pedestrian-injury crashes using mixed and latent-class models. Anal. Methods Accid. Res. 12, 1–17.
Behnood, A., Mannering, F., 2017. Determinants of bicyclist injury severities in bicycle-vehicle crashes: a random parameters approach with heterogeneity in means and variances. Anal. Methods Accid. Res. 16, 35–47.
Bhat, C., 2003. Simulation estimation of mixed discrete choice models using randomized and scrambled Halton sequences. Transp. Res. Part B 37 (9), 837–855.
Buck, D., Buehler, R., Happ, P., Rawls, B., Chung, P., Borecki, N., 2013. Are bikeshare users different from regular cyclists? Transp. Res. Rec.: J. Transp. Res. Board 2387 (1), 112–119.
Borecki, N., Buck, D., Chung, P., Happ, P., Kushner, N., Maher, T., Buehler, R., 2012. Virginia TechCapital Bikeshare Study. Virginia Tech, Blacksburg.
Earl, A., Lewis Jr, N., 2019. Health in context: New perspectives on healthy thinking and healthy living. J. Exp. Soc. Psychol. https://doi.org/10.1016/j.jesp.2018.09.001. (in press).
Ehrgott, M., Wang, J.Y.T., Raith, A., van Houtte, C., 2012. A bi-objective cyclist route choice model. Transp. Res. Part A 46, 652–663.
El-Assi, W., Mahmoud, M., Habib, K., 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.

Ewing, R., Schmid, T., Killingsworth, R., Zlot, A., Raudenbush, S., 2003. Relationship between urban sprawl and physical activity, obesity, and morbidity. Am. J. Health Promot. 18 (1), 47–57.

Faghih-Imani, A., Eluru, N., El-Geneidy, A., Rabbat, M., Haq, U., 2014. How land-use and urban form impact bicycle flows: evidence from the bicycle-sharing system (BIXI) in Montreal. J. Transp. Geogr. 41, 306–314.

Fishman, E., 2016. Bikeshare: a review of recent literature. Transp. Rev. 36 (1), 92–113.

Fishman, E., Washington, S., Haworth, N., Watson, A., 2015. Factors influencing bike share membership: an analysis of Melbourne and Brisbane. Transp. Res. Part A 71, 17–30.

Fishman, E., Washington, S., Haworth, N., 2014. Bike shares impact on car use: evidence from the United States, Great Britain, and Australia. Transp. Res. Part D 31 (7), 13–20.

Fountas, G., Anastasopoulos, P., 2018. Analysis of accident injury-severity outcomes: the zero-inflated hierarchical ordered probit model with correlated disturbances. Anal. Methods Accid. Res. 20, 30–45.

Fountas, G., Anastasopoulos, P., Abdel-Aty, M., 2018a. Analysis of accident injury-severities using a correlated random parameters ordered probit approach with time variant covariates. Anal. Methods Accid. Res. 18, 57–68.

Fountas, G., Anastasopoulos, P., Mannering, F., 2018b. Analysis of vehicle accident-injury severities: a comparison of segment- versus accident-based latent class ordered probit models with class-probability functions. Anal. Methods Accid. Res. 18, 15–32.

Frank, L., Andresen, M., Schmid, T., 2004. Obesity relation- ships with community design, physical activity, and time spent in cars. Am. J. Prev. Med. 27 (2), 87–96.

Garcia-Palomares, J.C., Gutierrez, J., Latorre, M., 2012. Optimizing the location of stations in bikesharing programs: a GIS approach. Appl. Geogr. 35 (1-2), 235–246.

Giles-Corti, B., Macintyre, S., Clarkson, J., Pikora, T., Donovan, R., 2003. Environmental and lifestyle factors associated with over- weight and obesity in Perth, Australia. Am. J. Health Promot. 8 (1), 93–102.

Goodman, A., Cheshire, J., 2014. Inequalities in the London bicycle sharing system revisited: impacts of extending the scheme to poorer areas but then doubling prices. J. Transp. Geogr. 41, 272–279.

Gulsah, A., Fischer, N., Namgung, M., 2013. Bicycling choice and gender case study: the Ohio State University. Int. J. Sustain. Transp. 7 (5), 347–365.

Habib, K., Han, X., Lin, W., 2014. Joint modelling of propensity and distance for walking-trip generation. Transp. A: Transp. Sci. 10, 420–436.

Harms, L., Bertolini, L., Brommelstroet, M., 2014. Spatial and social variations in cycling patterns in a mature cycling country exploring differences and trends. J. Transp. Health 1, 232–242.

Hong, J., Sila-Nowicka, K., McArthur, D., 2018. Is the popularity of social networking services beneficial for public health? Focusing on active travel and BMI. J. Transp. Health. https://doi.org/10.1016/j.jth.2018.09.003. (in press).

Kaplan, D., Knowles, M., 2015. Developing a next-generation campus bike-share program. Plan. High. Educ. J. 44 (1), 63–75.

Larsen, J., Patterson, Z., El-Geneidy, A., 2013. Build it. But where? The use of geographic information systems in identifying locations for new cycling infrastructure. Int. J. Sustain. Transp. 7, 299–317.

Lopez, R., 2004. Urban sprawl and risk for being overweight or obese. Am. J. Public Health 94 (9), 1574–1579.

Lu, M., Hsu, S., Chen, P., Lee, W., 2018. Improving the sustainability of integrated transportation system with bike-sharing: a spatial agent-based approach. Sustain. Cities Soc. 41, 44–51.

Maizlish, N., Linesch, N., Woodcock, J., 2017. Health and greenhouse gas mitigation benefits of ambitious expansion of cycling, walking, and transit in California. J. Transp. Health 6, 490–500.

Mannering, F., 2018. Temporal instability and the analysis of highway accident data. Anal. Methods Accid. Res. 17, 1–13.

Mannering, F., Shankar, V., Bhat, C., 2016. Unobserved heterogeneity and the statistical analysis of highway accident data. Anal. Methods Accid. Res. 11, 1–16.

Marcoux, R., Yasmin, S., Eluru, N., Rahman, M., 2018. Evaluating temporal variability of exogenous variable impacts over 25 years: an application of scaled gen- eralized ordered logit model for driver injury severity. Anal. Methods Accid. Res. 20, 15–19.

McArdle, W.D., 2010. Exercise Physiology: Nutrition, Energy and Human Performance, 7th ed. Lippincott Williams Wilkins, Philadelphia, P.A., London (International ed. ed.).

McFadden, D., 1981. Econometric Models for Probabilistic Choice, Structural Analysis of Discrete Data Using Econometric Applications. MIT Press, Cambridge, MA.

McFadden, D., Ruud, P., 1994. Estimation by simulation. Rev. Econ. Stat. 76 (4), 591–608.

Menghini, G., Carrasco, N., Schussler, N., Axhausen, K., 2010. Route choice of cyclists in Zurich. Transp. Res. Part A 44 (9), 754–765.

Milton, J., Shankar, V., Mannering, F., 2008. Highway accident severities and the mixed logit model: an exploratory empirical analysis. Accid. Anal. Prev. 40 (1), 260–266.

Molina-Garcia, J., Castillo, I., Queralt, A., Sallis, J., 2015. Bicycling to university: evaluation of a bicycle-sharing program in Spain. Health Promot. Int. 30 (2), 350–358.

National Association of City Transportation Officials, 2017. Bike share in the US: 2010–2016. ⟨https://nacto.org/bike-share-statistics-2016/⟩ Retrieved: November 2017.

Nikitas, A., 2018. Understanding bike-sharing acceptability and expected usage patterns in the context of a small city novel to the concept: a story of 'Greek Drama'. Transp. Res. Part F 56, 306–321.

Norton, R.K., Brix, A., Brydon, T., Davidian, E., Dinse, K., Vidyarthi, S., 2007. Transforming the university campus into a sustainable community. Plan. High. Educ. 35 (4), 22–39.

O'Brien, O., Cheshire, J., Batty, M., 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. J. Transp. Geogr. 34, 262–273.

Osama, A., Sayed, T., 2017. Investigating the effect of spatial and mode correlations on active transportation safety modeling. Anal. Methods Accid. Res. 16, 60–74.

Pal, A., Zhang, Y., 2017. Free-floating bike sharing: solving real-life large-scale static rebalancing problems. Transp. Res. Part C 80, 92–116.

Philips, I., Watling, D., Timms, P., 2018. Estimating individual physical capability (IPC) to make journeys by bicycle. Int. J. Sustain. Transp. 12 (5), 324–340.

Pucher, J., Buehler, R., 2012. Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. Transportation 39 (2), 409–432.

Pucher, J., Buehler, R., Seinen, M., 2011. Bicycling renaissance in North America: an update and re-appraisal of cycling trends and policies. Transp. Res. Part A 45 (6), 451–475.

Saelens, B.E., Sallis, J.F., Black, J.B., Chen, D., 2003. Neighbor- hood-based differences in physical activity: an environment scale evaluation. Am. J. Public Health 93 (9), 1552–1558.

Seraneeprakarn, P., Huang, S., Shankar, V., Mannering, F., Venkataraman, N., Milton, J., 2017. Occupant injury severities in hybrid-vehicle involved crashes: a random parameters approach with heterogeneity in means and variances. Anal. Methods Accid. Res. 15, 41–55.

Shaheen S., Cohen A., Zohdy, I., 2016. Shared mobility: Current practices and guiding principles. U.S. Department of Transportation. Report No. FHWA-HOP-16-022.

Shaheen, S., Guzman, S., Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia. Transp. Res. Rec. 159–167.

Shaheen, S., Martin, E., Cohen, A., 2013. Public bikesharing and modal shift behavior: a comparative study of early bikesharing systems in North America. Int. J. Transp. 1, 35–54.

Shaheen, S., Martin, E., Chan, N., Cohen, A., Pogodzinski, M., 2014. Public Bikesharing in North America During a Period of Rapid Expansion: Understanding Business Models, Industry Trends and User Impacts. Mineta Transportation Institute, San Jose.

Shaheen, S., Zhang, H., Martin, E., Guzman, S., 2011. Hangzhou public bicycle: understanding early adoption and behavioral response to bike sharing in Hangzhou, China. Transp. Res. Rec. 2247, 33–41.

Smith, M.S., Kauermann, G., 2011. Bicycle commuting in Melbourne during the 2000s energy crisis: a semiparametric analysis of intraday volumes. Transp. Res. Part B 45 (10), 1846–1862.

Stinson, M., Bhat, C.R., 2004. Frequency of bicycle commuting: internet- based survey analysis. Transp. Res. Rec. 1878, 122–130.

Strum, R., Cohen, D.A., 2004. Suburban sprawl and physical and mental health. Public Health 118 (7), 448–496.

Sun, Y., Mobasheri, A., Hu, X., Wang, W., 2017. Investigating impacts of environmental factors on the cycling behavior of bicycle-sharing. Sustainability 9 (6), 1060.

Swiers, R., Pritchard, C., Gee, I., 2017. A cross sectional survey of attitudes, behaviours, barriers and motivators to cycling in University students. J. Transp. Health 6, 379–385.

Wadud, Z., 2014. Cycling in a changed climate. J. Transp. Geogr. 35, 12–20.

Washington, S., Karlaftis, M., Mannering, F., 2011. Statistical and Econometric Methods for Transportation Data Analysis, Second Edition. Chapman & Hall/ CRC, Boca Raton, FL.

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., Goodman, A., 2014. Health effects of the London bicycle sharing system: health impact modelling study. Br. Med. J. 348.

Wuerzer, T., Mason, S., 2015. Cycling willingness: investigating distance as a dependent variable in cycling behavior among college students. Appl. Geogr. 60 (6), 95–106.