Examining the Impact of Sample Size in the Analysis of Bicycle Sharing Systems

Ahmadreza Faghih-Imani
PhD Student
Department of Civil Engineering and Applied Mechanics
McGill University
Ph: 514-398-6823, Fax: 514-398-7361
E-mail: seyed.faghihimani@mail.mcgill.ca

Naveen Eluru*
Associate Professor
Department of Civil, Environmental and Construction Engineering
University of Central Florida
Ph: 407-823-4815; Fax: 407-823-3315
E-mail: naveen.eluru@ucf.edu

August 6, 2016

---

*Corresponding author

Abstract

Given the growing installation of bicycle-sharing systems (BSS) across the world, there is a gradual increase in research on BSS over the past few years. Research efforts examining BSS employed a wide range of sample size depending on the temporal or spatial aggregation. The main objective of this paper is to investigate the impact of sample size on BSS analysis using data from New York City's BSS (CitiBike). This paper proposes a systematic evaluation of the impact of sample size on model estimates, inference measures and predictive performance. We evaluate two major dimensions of BSS data: 1) system usage – impact of contributing factors on hourly arrival and departure rates at station level, 2) user destination choice – impact of factors on users' preference of destination station choice. To examine the system usage, we employ the linear mixed model methodology while the user destination choice is studied using the Multinomial Logit Model (MNL). The model estimation exercises for system demand and destination choice are conducted on several samples of data. The performance of these sample models in terms of parameters, inference statistics and predictions relative to a base sample data is observed. The results would help the analysts to make decisions on sample size for accurately examining BSS usage. The analyses show that the impact of sample size on parameters estimated is stronger than that of the impact on prediction performance.

Keywords: Sample size, bicycle sharing systems, CitiBike New York, linear mixed model, multinomial logit model, arrival and departure rates, destination choice, bicycle infrastructure, land use and built environment

1. INTRODUCTION

1.1. Background

In recent years, there has been growing attention on bicycle sharing systems (BSS) as an alternative and complementary mode of transportation (Shaheen et al. 2010; Faghih-Imani et al. 2014). A bicycle-sharing system provides increased flexibility to ride a bicycle without the costs and responsibilities associated with owning a bicycle (such as the need to secure their bicycles or perform regular maintenance). At the same time, the decision to make a trip can be made in a short time frame providing an instantaneously accessible alternative for a one-way or a round trip. These systems can enhance accessibility to public transportation systems by improving the last mile connectivity (Jäppinen et al., 2013). Moreover, BSS's implementation in the city can motivate new segments of the society to cycle resulting in an increase in the overall bicycling mode share while also enhancing physical activity levels to obtain better health outcomes (Fuller et al. 2011; Buck et al., 2013; Fishman et al. 2015). Further, earlier research efforts observed that BSS were successful in normalizing the image of cycling while increasing driver awareness towards cyclists improving the safety of cyclists (Goodman et al., 2014; Murphy and Usher 2015).

Cities, by installing bicycle-sharing systems, are focusing on inducing a modal shift to cycling, and subsequently, decrease traffic congestion and air pollution. There is significant evidence from the travel behavior data in the United States to support BSS installation in urban areas. According to data from the 2009 National Household Travel Survey (NHTS), about 37.6% of the trips by private vehicles in the United States are less than 2 miles long. The NHTS data also indicates that about 73.6% of bicycle trips in the US are less than 2 miles long. Even if a small proportion of the shorter private vehicle trips (around dense urban cores) are substituted with BSS trips it offers substantial benefits to individuals, cities, and the environment. Thus it is not surprising that more than 1000 cities around the world have installed or plan to install a bicycle-sharing system (Meddin and DeMaio, 2015). With the growing installation of BSS infrastructure across the world, there is a substantial interest in understanding how these systems impact the urban transportation system. Research efforts examining BSS employed a wide range of sample sizes depending on the temporal or spatial aggregation. While it is beneficial to use large sample sizes for analysis, an increase in sample sizes are associated with increased data preparation effort, and longer model run times. In this context, the main objective of this paper is to investigate the impact of sample size on BSS analysis using data from New York City's BSS (CitiBike). Specifically, the research evaluates the impact of sample size on model parameter estimates, inference measures and prediction capabilities. The findings provide analysts and planners guidelines on the "*minimum*" and "*ideal*" size of data necessary for examining BSS.

1.2. Bicycle-Sharing System Studies

There is a gradual increase in the research on bicycle-sharing systems over the past few years (see Ricci (2015) and Fishman (2016) for a review of recent literature on BSS). Table 1 presents a summary of earlier research efforts that employed system or user level data from BSS around the world. The table provides information on the study, temporal aggregation, spatial aggregation, BSS details, sample size and modeling approach employed. From the table, it is evident that based on the temporal or spatial aggregation, the sample sizes can vary substantially. The majority of the research efforts are conducted at the system level with only 2 out of 10 reviewed here exploring user level decisions. The methodologies considered for analysis include linear regression, linear mixed models, count models and multinomial logit models.

Under the systems perspective, quantitative studies employed actual bicycle usage data to capture the determinants of BSS usage (Nair et al. 2013; Rixey 2013; Gebhart and Noland 2014; O'Brien et al. 2014; Faghih-Imani et al. 2014; Faghih-Imani and Eluru 2014; Rudloff and Lackner 2014; Zhao et al., 2014; Wang et al. 2015). These studies typically postulate that BSS usage from a system perspective is influenced by various attributes such as BSS infrastructure (such as number of BSS stations and stations' capacity), transportation network infrastructure (such as length of bicycle facilities, streets and major roads), land use and urban form (such as presence of metro and bus stations, restaurants, businesses and universities), meteorological data (such as temperature and humidity), and temporal characteristics (such as time of day, day of the week and month). These studies mostly characterized usage as the number of trips originated and destined to one station or divided the usage of one station in two rates: arrivals (depositing bicycle) and departures (removal of bicycles). These studies employed various levels of aggregation both temporally such as hourly, daily or monthly usage and spatially such as station level or TAZ (traffic analysis zone) level.

The studies focussed on the user perspective contribute to the literature by studying user behavior in response to bicycle-sharing systems. These studies analyze how people integrate BSS with other urban transportation systems. Studies found that convenience offered by BSS is the main factor that significantly encouraged individuals to use the system (Fuller et al., 2011; Bachand-Marleau et al., 2012). Several studies highlighted the differences between BSS short-term users and BSS annual members' preferences towards the use of the system (Lathia et al., 2012; Buck et al., 2013; Faghih-Imani and Eluru, 2015). Studies also found that BSS users, in general, prefer shorter trips with all else same (Faghih-Imani and Eluru, 2015; Mahmoud et al., 2015). Further, research efforts demonstrated that BSS users prefer to use the existing bicycle facilities such as bicycle lanes and have a higher interest in stations closer to transit system such as subway stations (Faghih-Imani and Eluru, 2015; González et al., 2016). Several studies underscored the use of BSS for the daily commute to and from work especially by annual members (Faghih-Imani et al., 2014; Faghih-Imani and Eluru, 2015; Murphy and Usher 2015).

1.3. Data Employed for Analysis

The BSS operators provide system availability data to users on their websites. Through relatively simple scripting exercises, it is possible to build a database of bicycle availability across stations for the BSS system. The data thus obtained can provide a glimpse of how BSS usage varies across the day. By augmenting the BSS usage variables with land-use, temporal and meteorological variables, a demand model to study the number of bicycles arriving or departing from a station can be examined. The process, analogous to trip generation in the traditional four-step model provides the demand from and into a station. Transportation planners are interested in such analysis as it quantifies the impact of land use factors on BSS usage. An important question in the process of developing such BSS demand models is to choose the size of the data to be selected for the model estimation sample. As opposed to the traditional travel demand literature where sample sizes are quite limited, in the context of BSS, demand information is available for every minute for multiple days and months. Hence, the selection of appropriate sample for demand analysis is quite critical.

More recently, in addition to the system availability information, BSS operators release trip information containing details including origin and destination stations, start and end time of the trip for BSS users. The station usage may also be obtained from this information by aggregating trips originated or destined at one station. In addition to usage rate, the information is useful to understand destination choice behavior of BSS users. Developing destination choice models will allow us to undertake the trip distribution step for BSS users. The exercise would allow us to identify trade-offs between distance and several exogenous variables in determining the destination. Again, a challenge in this process is the determination of sample size for destination choice models.

For both demand analysis and destination choice modeling, the size of sample influences the complexity of the modelling process. Employing large samples requires substantial data preparation and model run times. For example, one month of data for a BSS with 300 stations results in 216,000 records of hourly arrivals or departures and about one million trips. The processing of usage or trip data and preparation of station level variables including built environment attributes and other variables such as weather characteristics or temporal attributes are substantially time-consuming. In addition to data preparation, a very large sample significantly increases the model run times. On the other hand, employing a smaller sample than appropriate would result in inaccurate and possibly even biased model estimates affecting the planning process. Hence, it would be useful to understand the sample size requirements for examining bicycle-sharing systems. Besides, the data is not always available; knowing the required appropriate sample size prior to collecting data would be beneficial. Due to the relative infancy of BSS, there is little to no guidance on the amount of data necessary for analysis.

1.4. Current Study in Context

It is evident from the discussion above that sample size requirements would assist transportation planners in developing reasonable models to study BSS trip generation (arrivals and departures) and distribution (destination). The current study proposes a systematic evaluation of the impact of sample size on model estimates, inference statistics, and predictive performance. Towards this end, we evaluate the BSS data from two perspectives: 1) system usage – what contributing factors influence hourly arrival and departure rates at a station level, 2) user destination choice – what factors contribute to users' preference of destination station choice.

To examine the system usage, we employ the linear mixed model methodology to determine the factors contributing to BSS usage. The usage is characterized as hourly arrival and departure rates of each station. The traditional linear regression model is not appropriate to study data with multiple repeated observations such as the hourly arrivals and departures for each station in our empirical context. Thus, we employ a linear mixed modeling approach that builds on the linear regression model while incorporating the influence of repeated observations from the same station. Further, the dependent variable is defined as the logarithm of hourly arrival and departure rates normalized by the station capacity (see Faghih-Imani et al., 2014). We analyze New York City's BSS (CitiBike) stations' hourly arrivals and departures for various samples.

To explore the user destination choice, we employ the Multinomial Logit Model (MNL) to examine the impact of individual bicyclist attributes (such as age and gender), trip attributes (such as time period of the day) and destination attributes (such as distance from the origin station, bicycle infrastructure variables and land use and built environment attributes) on destination choice. The most common methodology to study location choice in transportation and related literature is MNL (see Faghih-Imani and Eluru, 2015). We estimate the MNL model for CitiBike system in New York City.

The model estimation exercises for system demand and destination choice are conducted on several samples of data (separately for weekdays and weekend days). The performance of these sample models relative to a base sample data is observed. Further, the performance of these sample based models on a hold-out sample relative to the predictive accuracy of the base sample is also compared. In order to account for the randomness of selecting smaller samples, for each smaller sample size, we randomly select five sets of data from that large sample and report the range and the average results. The results would help the analysts to identify necessary sample size for accurately examining BSS usage.

The remainder of the paper is organized in the following order. Section 2 describes the research methodology, data, and the sample formation procedures. In Section 3, the models used in our analysis are described. Section 4 presents the model results and comparison. Finally, Section 5 summarizes and concludes the paper.

2. RESEARCH METHODOLOGY

2.1. Data Source

New York's CitiBike system is the latest major public bicycle-sharing system around the world and the largest in the United States. The service was launched in May 2013 with 330 stations and 6,000 bicycles in the lower half of Manhattan and some part of northwest of Brooklyn (Figure 1). The system covers the city's major commercial business districts and some residential areas with an average daily ridership of 34,000 trips. New York City is the most populous city in the US and a host to millions of visitors every year. In 2013, the mode share of cycling in New York City reached 1% from about 0.5% in 2007 (Kaufman et al., 2015). The city's dense and walkable urban form provide a good opportunity for the success of a well-designed BSS.

The data used in our research was obtained from CitiBike website (https://www.citibikenyc.com/system-data). The CitiBike website provides trip dataset for every month of operation since July 2013. The trip dataset includes information about origin and destination stations, start time and end time of trips, user types i.e. whether the user was a subscriber of the system with annual membership or a customer with a temporary pass, and the age and gender for members' trips only. Additionally, the stations' capacity and coordinates, as well as trip duration, are also provided in the dataset. The built environment attributes such as bicycle routes and subway stations are derived from New York City open data (https://nycopendata.socrata.com/) while the socio-demographic characteristics are gathered from US 2010 census and the weather information are for Central Park station from National Climatic Data Center.

2.2. Sample Formation

The main objective of this study is to examine the impact of sample size on BSS analysis. For this purpose, we look into two sets of models: hourly arrival and departure rates and a destination station choice model. In this section, we explain the study methodology on how the various samples are generated. Findings from earlier research have indicated that travel patterns of BSS are different in weekdays and weekends (Faghih-Imani et al. 2014). Hence, the research exercise is conducted separately for weekdays and weekends. Moreover, earlier studies showed that there is a significant difference between the behavior of annual members' and the behavior of customers with temporary passes in using BSS (Lathia et al., 2012; Buck et al., 2013; Faghih-Imani and Eluru, 2015). In this paper, we distinguish between the trips made by annual members and daily customers. For the sake of brevity, we only focus on trips, arrivals, departures and destination choices made by annual members[1].

A sample formation was necessary in order to obtain the arrivals and departures. We aggregated the number of trips originated from/destined to one station by the different type of users at an hourly level to obtain hourly arrivals and departures by members and daily customers at a station level. Further, we normalized stations' arrivals and departures with station capacity to account for the influence of station capacity on demand. Station capacity is defined as the total number of dock spaces at each station. In our modeling efforts, we employ logarithm of the hourly normalized arrivals and departures as the dependent variable. We focused on the month of September 2013; i.e. the peak month of the usage in 2013 for our base analysis. We separated weekdays and weekends. This would give us a base sample consisting of 166,320 records (330 stations × 24 hours × 21 days) for weekday models and 71,280 records (330 stations × 24 hours × 9 days) for weekend models. We estimate our base model for arrivals and departures and assume the estimate results as the true (base) values and compared the rest of models with these results. Then, from this base sample we randomly select a series of smaller samples. For this purpose, we select random weekday/weekend days in the month of September and assign them to each station. For weekdays, we choose 10 days, 5 days, 3 days, 2 days and 1 day randomly for each station to create smaller samples while for weekends we choose samples for 5 days, 3 days, 2 days and 1 day. It must be noted that the random days assigned to each station are different from random days assigned to other stations; thus the sampling approach covers the whole month across the urban region. To account for the impact of randomness, we generate five sets of these random days, estimate the arrival and departure models with both, and then obtain the average results (while also providing the range of the estimates).

For destination choice model, to be consistent with the usage analysis, again we focus on the trips in the month of September. The sample formation exercise also involved a series of steps. First, trips with missing or inconsistent information were removed. Second, trips longer than 90 minutes in duration (only 0.5% of all the trips) were deleted considering that the trips longer than 90 minutes are not typical bicycle-sharing rides and could also be a result of misplacing the bicycle when returning it to the station. At the same time, trips that had the same origin and destination were also eliminated. For trips with the same origin and destination, it is possible that the bicycle was not functioning well and the users returned them to the origin station. Also, to accommodate for intentional same origin and destination trips would require additional trip purpose information and is beyond the scope of this work. Therefore, we focus on trips that were destined outward.

[1] We have six models and six tables of results. If we wanted to add the results for daily users, the entire effort documented will need to be repeated for daily users. Moreover, the daily users typically account for a small share of BSS usage; for example, in New York City, only about 10% of trips are made by daily customers in 2014. To be sure, the proposed modeling framework and the systematic evaluation of the impact of sampling procedure can be applied on trips made only by daily users.

Further, we separated trips made by members and daily customers; about 86% of all the trips were made by members. Again, we consider separate sets of trips for weekday and weekend models.

CitiBike system had 330 stations in September 2013. Considering all the stations in the universal choice set of destination station choice model will result in substantial computational burden. Thus, for each trip, we randomly sample 30 stations from the universal choice set including the chosen alternative. It must be highlighted that the 30 different random stations are obtained for every individual trip. McFadden (1978) showed that the process of random sampling of alternatives does not affect the parameter estimates in multinomial logit models (see Faghih-Imani and Eluru, 2015 for a similar assumption). For the evaluation of sampling impact, we consider the sample with 50,000 trips made in weekdays and another 50,000 trips made in weekends as the base sample for our weekday and weekend models. Then from these 50,000 trips, we randomly select five sets of 20000, 10000, 5000, 3000, 2000, 1000 trips generated as our smaller samples for both weekday and weekend models. For every sample size, the information for the 30 stations is augmented with the individual trip records. A descriptive summary of base samples characteristics is presented in Table 2.

2.3. Independent Variable Generation

The independent variables considered in our analysis can be categorized into four groups: (1) weather, (2) temporal, (3) spatial variables and (4) trip attributes. It must be noted that trip attributes are only included in the destination choice models. Weather variables include hourly temperature, relative humidity, and the hourly weather condition represented as a dummy variable indicating whether or not it is raining. Considering the start time of the trips for departures and end time of the trips for arrivals, five time periods were created: AM (7:00-10:00), Midday (10:00-16:00), PM (16:00-20:00), Evening (20:00-24:00), and Night (0:00-7:00) to capture the time of the day effect on usage. For the destination choice models, the same time periods are used considering the start time of the trips.

Several variables were considered from the spatial variables group. Population density was calculated at census block level and employment density at zip code level. Other attributes were considered at a station buffer level. A 250-meter buffer around each station was found to be an appropriate walking distance considering the distances between CitiBike stations and the dense urban form of New York City; typical New York City block is about 60 meter (Kaufman et al., 2015). The length of bicycle routes and streets in the 250-meter buffer around the stations were calculated in order to examine the impact of street network and cycling facilities. The number and capacity of CitiBike stations in the 250-meter buffer were computed to capture the effect of neighbouring stations. The presence of subway and Path train stations in the 250-meter buffer were generated to examine the influence of public transit on BSS usage. The number of restaurants (including coffee shops and bars), and the area of park in the buffer were also considered as the point of interest attributes near CitiBike stations.

Trip attributes considered in destination choice model include the street network distance between the origin and destination of every trip. This distance was computed using the shortest path between origin and destination stations to investigate the travel distance influence along with other attributes. The shortest distance is computed based on the street network around the stations (excluding highways). The estimated cycling distance serves as a surrogate for the actual distance and is a reasonable reflection of the actual distance between stations. While the actual trip might involve a different route, the shortest distance would be an appropriate indicator of the distance traveled. Moreover, for the users with annual membership, the gender, and age information were available and were considered in our analysis. It must be mentioned that several exogenous variables such as gender, age or weather variables cannot be directly included within the destination choice model structure since these variables do not change across alternatives. Therefore, the interaction effects of such variables with distance variable are considered in our modelling effort.

3. MODELS

A brief description of the model structures employed in the sampling analysis is presented in this section. Specifically, we consider a linear mixed model structure for BSS demand (arrivals and departures) and a multinomial logit structure for destination choice. The same model structure is employed across the various samples chosen for analysis.

3.1. Linear mixed models

Let q = 1, 2, …, Q be an index to represent each station, d = 1, 2, …, D be an index to represent the number of days on which data was collected (sample size) and t = 1, 2, …, 24 be an index for hourly data collection period. The dependent variable (arrival or departure rate over station capacity) is modeled using a linear regression equation which, in its most general form, has the following structure:

$$y_{qdt} = \beta X + \varepsilon \qquad (1)$$

where y_{qdt} is the logarithm of normalized arrival or departure rate as the dependent variable, X is an L×1 column vector of attributes and the model coefficients, β, is an L×1 column vector. The random error term, ε, is assumed to be normally distributed across the dataset.

The error term may consist of three components of unobserved factors: a station component, a day component, and an hour-of-the-day component. We consider the station and the time-of-day to be related common unobserved effects. In this structure, the data can be visualized as 24 records for each Station-Day combination for a total of "Q stations × D days" observations. We parameterize the covariance matrix (Ω). For estimating a parsimonious specification, we assume a first-order autoregressive moving average correlation structure with three parameters σ, ρ, and φ as follows:

$$\Omega = \sigma^2 \begin{pmatrix} 1 & \varphi\rho & \varphi\rho^2 & \cdots & \varphi\rho^{23} \\ \varphi\rho & 1 & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \varphi\rho^{23} & \cdots & \cdots & \cdots & 1 \end{pmatrix} \qquad (2)$$

The parameter σ^2 represents the error variance of ε, φ represents the common correlation factor across time periods, and ρ represents the dampening parameter that reduces the correlation with time. The correlation parameters φ and ρ, if significant, highlight the impact of station specific effects on the dependent variables. The models are estimated in SPSS using the Restricted Maximum Likelihood Approach (REML).

3.2. Multinomial Logit Model

Let q = 1, 2, …, Q again be an index to represent each station, j = 1, 2, …, J be an index to represent the BSS users. Then, the random utility formulation takes the following form:

$$u_{jq} = \beta' X_{jq} + \varepsilon_{jq} \tag{3}$$

Where u_{jq} is the utility obtained by user j by selecting station q from the choice set of 30 stations. X_{jq} is the vector of attributes and β is the model coefficients to be estimated. The random error term, ε, is assumed to be independent and identically Gumbel-distributed across the dataset. The BSS user j will choose a station as the destination that offers the highest utility. With this notation, the probability expression takes the typical multinomial logit form given by:

$$P_{jq} = \frac{\exp(\beta' X_{jq})}{\sum_{q=1}^{30} \exp(\beta' X_{jq})} \tag{4}$$

The log-likelihood function can be defined as:

$$L = \sum_q \sum_j \ln(P_{jq})^{d_{jq}}$$

(5)

where d_{jq} is an indicator variable equal to 1 for the station chosen for BSS user j and 0 otherwise. By maximizing this log-likelihood function, the model parameters β are estimated. The maximum likelihood model estimation is programmed in GAUSS matrix programming language.

4. ANALYSIS AND DISCUSSION

In this study, the final model specifications for arrivals, departures, and destination choice were obtained after testing for several specifications using the corresponding base samples. The specifications were evaluated based on data fit, parameter significance and intuitiveness supported by statistical inference. The final specification based on base samples for arrivals, departures and destination choice models are presented in the first column of Tables 3 to 8.

4.1. Evaluation Measures

Considering that the base models represent the population models, we set forth estimating the same specification using the several samples prepared. As expected, the estimation on the smaller samples provides different values for the estimated parameters and different standard errors of estimates. To account for the randomness of the sampling process, we estimate the specifications on five sets of samples for each sample size. The parameters and standard errors estimated for each of five samples are compared to the base results and percentage differences at a parameter level and a model level are computed. For each sample size, we report the mean and range of percentage changes in order to show the variations not only between the different sample sizes but also between different sets of random samples for each sample size. The estimates are presented in each column of Table 3 to 8. All the base reported estimates are significant at 95% level of confidence.

The impact of sample size on the estimation results are examined by the following measures: 1) the capability to produce the same parameter estimates of the base sample, 2) the significance of the parameter represented by the standard error and 3) the prediction capability as a measure of goodness of fit to predict the same results for data hold-out sample. For each variable in each sample size model, we calculate the percentage error with respect to base estimate of the variable coefficient and standard error for every set of a random sample of that size. Then, for each variable, we present the mean percentage change and the range of percentage change within the five random samples of smaller samples. For usage models, in order to evaluate the parameters estimated, mean absolute percentage error (MAPE) and root mean square error (RMSE) of estimated parameters with respect to the base sample' estimates are calculated. In addition, the MAPE for the change in standard error of estimates with respect to base values are also generated. In order to better demonstrate the models' performance, we also indicate the number of parameters which become statistically insignificant at 95% level of confidence when we use smaller samples. In order to show the prediction capability of models, we used the data from the first week of October 2013 (i.e. the next week after our base sample for model estimation) to validate the estimated models by each sample. We used the first five weekdays of October for our weekday's models and the first Saturday and Sunday of October for our weekend's models as validation data. The same data procedure described in sample preparation for models estimation was repeated in order to compute hourly arrivals and departures. For each sample, the model developed was used to generate predictions of hourly arrival and departure rates and the predictions were compared with the observed rates in the validation sample. Again, to compare the prediction performance, we calculate two error metrics of mean absolute error (MAE) and RMSE. All the measures for evaluating arrival and departure models performance are presented in the bottom rows of Table 3 and Table 4 for weekday models and Table 5 and Table 6 for weekend models.

For the destination choice model, the same measures including the mean and range of percentage change for each variable, and MAPE and RMSE as aggregated measure are used in order to evaluate the performance of models to produce the estimated parameters of the base sample. Also, the mean and range of percentage change for standard errors of estimates and the number of parameters which become statistically insignificant at 95% level of confidence are calculated. However, for prediction capability measure, different procedure compared to the arrivals and departures models was needed. For this purpose, we employ a hold-out sample of 5000 trips in weekdays and 5000 trips in weekends as our validation sample. The same data preparation and choice set generation for estimation samples is exercised for the validation sample. The parameters estimated by each sample size models were used to compute the probability of choosing a station for 30 stations of choice set for each of the 5000 trips. In order to evaluate the performance of models in prediction, two metrics are used: a) the predictive log-likelihood: the sum of the log of the probability of chosen station across the validation sample, and b) the percentage of correct prediction (correct prediction is defined as assigning the highest probability to the chosen station). Again, all the measures for evaluation of sample size impact on the performance of destination choice models are presented in the bottom rows of Table 7 for weekday models and Table 8 for weekend models.

4.2. Evaluation results

In this section, we discuss the impact of sample size on the performance of models estimated. It is important to note that the weekday and weekend base models are slightly different. Some variable effects were insignificant in weekend models compared to the variable effects in weekday models. These include: (1) for arrivals model, the length of rails within buffer and the interaction of job density variable with AM; (2) for departures model, the interaction of population density and job density variables with AM; (3) for destination choice model, the presence of subway station in the buffer and the interaction of distance variable with temperature . The only variable that was significant in weekend models and insignificant in weekday models is the area of the park in the buffer variable for destination choice model. For the usage models, the weekend models have more stable performance; i.e. effect of sample size in arrivals models and departures models are less in weekend models (lower RMSE for both estimated parameters and prediction) which are expected since the usage of the system is lower in weekends. However, for destination choice models, the weekday and weekend models provide almost similar performances with the slightly better performance of weekday models in producing the base parameters and slightly better performance of weekend models in prediction.

For weekday usage models, the performance of models on smaller sample size to produce the estimated parameters of the base sample varies by MAPE of 2.54 to 9.85 and 2.28 to 9.56 for arrivals and departures, respectively. The results show that as we choose smaller sample size until two days, we have almost similar results (less than four variables -out of 20- become insignificant) as the base case. We observe a huge jump in percentage change of standard errors when we use only one day to estimate arrivals or departure models. The results for the standard error changes indicate that as sample size decreases, standard error of estimates increases substantially thus altering inference i.e. the variable might be considered insignificant. The population density variable and time period dummy variables (AM, Midday, PM, and Evening) present the lowest variation in the predicted estimates even in one-day sample in both arrivals and departures. On the other hand, the maximum variation of estimated parameters is associated with the interaction of built environment attributes with temporal variables such as the interaction of population and job density with AM and PM, and also the length of bicycle lanes in the buffer and temperature. The level of variation of variables can be recognized as a measure of variables’ importance in examining the arrivals and departures. In terms of prediction capability, we do not observe any significant difference (loss) due to the use of smaller samples.

For weekend usage models, similar trends are observed. The MAPE varies from 2.9 to 8.4 for arrivals and from 2.3 to 9.9 for departures. Again, we observe a huge drop in performance of models when we use only one day to estimate the usage models. The results show that as sample size decreases, the MAPE and the standard error of estimates increases. The lowest variation in the predicted estimates in both arrivals and departures is associated with the time period dummy variables, population density variable and the number of restaurants in the buffer variable while the largest variation of estimated parameters is observed for to the job density variable. The sample size does not have a significant impact on the prediction performance of both arrival and departure models. Based on the evaluation measures discussed above and specifically the escalation of error relative to standard errors, we suggest a minimum sample of three days for weekday models and two days for weekend models in order to sufficiently analyze BSS hourly usage.

For destination choice models, as sample size decreases, the error measures increase. The MAPE (RMSE) of model performance to generate the estimated parameters of base model increases from 7.5 (0.096) for sample size of 20,000 trips to 43.7 (0.575) of sample size of 1,000 trips for weekday models while the corresponding values for weekend models are an increase from 10.1 (0.142) to 46.6 (0.738). Again, as expected, the standard error of estimate increases when sample size reduces which is also demonstrated through the number of insignificant parameters. In addition, in general, the station capacity variable and the distance variable show the lowest variation in the predicted estimates for both weekday and weekend models, clearly indicating the importance of these two attributes in destination choice process of CitiBike users. Further, we can clearly see that as sample size decreases, the prediction capability of estimated models marginally reduces as highlighted by the predictive LL and percentage of correct prediction measures. However, the impact of sample size on prediction performance is not as substantial as that of the impact on parameters estimated. In total, based on the results and considering the increase in standard error, we can recommend a minimum sample size of 5000 trips for both weekday and weekend models for examining users' destination choice and more generally users' behaviour towards BSS.

Overall, from our research exercise, we observe that employing different sample sizes have a stronger impact on the parameters estimated relative to the prediction capability. There is a clear trade-off between the use of smaller sample size and the closeness of estimated results with the base results. Models estimated with smaller samples are likely to increase the error in parameter inference as well as prediction performance. It is important to note that in none of the estimated models (arrivals, departures, destination choice) for smaller samples, the sign of estimated coefficients changed from the base sample results. We also want to highlight that when we have considered a sample size of about half of the base sample (10 days out of 21 days of weekday models, 5 days out of 9 days of weekend models, 20000 trips out of 50000 trips), we obtain the results that are very similar to the results obtained for the base sample. Specifically, a) for usage models, the MAPE is less than 3% and 1 or 2 variables out of about 20 variables become insignificant and b) for destination choice models, the MAPE is less than 10% and three or fewer variables become insignificant. The decision on the exact size of the sample to be employed will need to be specific to each individual dataset and to be examined based on the system knowledge of the analyst. In the absence of any information, our recommended sample size for arrivals, departures, and destination choices for weekdays and weekends can be employed as minimum requirements (as far as possible).

5. CONCLUSION

This paper examined the impact of sample size on hourly usage and users' destination choice preferences employing data from New York City's CitiBike. Towards this end, we evaluated the BSS data from two perspectives: 1) system usage – what contributing factors influence hourly arrival and departure rates at a station level, 2) user destination choice – what factors contribute to users' preference of destination station choice. For system models, we estimated linear mixed models for hourly arrivals and departures on one month of data as our base model and compared it with the estimation on a set of smaller samples. We recognized the distinct behaviour of BSS in weekends by estimating separate models for weekdays and weekends. We considered 21 days of weekdays as our base sample and 10 days, 5 days, 3 days, 2 days and 1 day (five sets each) as our smaller samples while for weekends we considered 9 days of weekends as our base sample and 5 days, 3 days, 2 days and 1 day (five sets each) as our smaller samples. For destination choice models, we estimated multinomial logit model on 50,000 trips made in weekdays and 50,000 trips made in weekends as our base model and again on five sets of smaller samples (20000, 10000, 5000, 3000, 2000, 1000 trips) for weekday and weekend models. We examined the impact of sample size on the estimation results based on three measures: the capability to produce the same parameters estimate of the base sample, the comparison of standard errors and the prediction performance.

As expected, the estimation with the smaller samples provided different values for the estimated parameters and standard error of estimates. For usage, the performance of models on smaller sample size to produce the "true" parameters were within 10% of the base case. However, the increase in the standard errors were substantial (i.e. the confidence interval for the parameters was large). We observed a huge drop in performance of models when we used only one day to estimate arrivals or departure models. The results clearly indicated that when sample size decreases, standard error of estimates gradually increases and the confidence in estimated parameters reduces and more number of variables became insignificant. The mean number of insignificant parameters varied from 1.2 to 6.4 variables. In terms of prediction capability, we were not able to observe any significant difference due to the use of smaller samples. For destination choice models, as sample size decreased, the error measures increased. The MAPE of model performance to generate the estimated parameters of the base model for weekday (weekend) models increased from 7.5 (10.1) for the sample size of 20,000 trips to 43.7 (46.6) for the sample size of 1,000 trips. Again, as expected, the standard error of estimate increased when sample size reduced. The impact on the confidence of estimated parameters in destination choice models is slightly higher than the usage models; the mean number of parameters that become insignificant varies between 2.6 to 11 variables. Further, we observed that as sample size decreases, the prediction capability of estimated models marginally reduced. However, the impact of sample size on prediction performance is not as substantial as that impact on parameters estimated. The results suggested a minimum sample size of three days data for weekday analysis and two days data for weekend analysis to examine BSS demand and 5000 trips for weekday and weekend models for examining users' destination choice. While these results cannot be generalized across all urban regions, these guidelines are intended to serve as minimum requirements for sample sizes analyzing BSS in the absence of any system level guidelines.

To be sure, the study is not without limitations. While the two model structures have been extensively tested, these specifications might not be applicable for other regions. Hence, the transferability of sample sizes cannot be generalized to other urban regions. However, considering the guidelines from our research as minimum requirements will ensure that the sample sizes employed for analyzing BSS are reasonable even for other urban regions. Moreover, the data employed in our analysis is actual trip data from the BSS operator. We do not have information on potentially latent demand. Thus, developing improved models that consider the truncated nature of the arrival and departure rates might be useful. Furthermore, it is possible to consider advanced models for usage and destination choice. For example, in the usage models, it is possible that potential unobserved spatial correlations exist between demand across stations close to one another (see Faghih-Imani and Eluru, 2016). Accommodating for such potential spatial correlations in a sampling exercise would be a future direction of research.

ACKNOWLEDGMENT

The authors are grateful for the insightful feedback from four anonymous reviewers on an earlier version of the manuscript.

References

Bachand-Marleau, J., B. Lee, and A. El-Geneidy, 2012. Better Understanding of Factors Influencing Likelihood of Using Shared Bicycle Systems and Frequency of Use. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2314, pp. 66-71.

Buck, D., Buehler, R., 2012. Bike lanes and other determinants of capital bikeshare trips. Paper presented at the 91st Transportation Research Board Annual Meeting 2012, Washington, DC.

Buck, D., D. Buehler, P. Happ, B. Rawls, P. Chung, and N. Borecki, 2013. Are Bikeshare Users Different from Regular Cyclists? A First Look at Short-Term Users, Annual Members, and Area Cyclists in the Washington, D.C., Region. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2387, pp. 112–119.

Faghih-Imani A. and N. Eluru, Role of Bicycle Sharing System Infrastructure on Usage: Evidence from Montreal, Presented at the 5th Innovations in Travel Modeling Conference, Baltimore, MD, 2014.

Faghih-Imani, A., and N. Eluru (2016), "Incorporating the Impact of Spatio-Temporal Interactions on Bicycle Sharing System Demand: A Case Study of New York CitiBike System," Journal of Transport Geography, Vol 54, June 2016, pp. 218-227

Faghih-Imani A., and N. Eluru, 2015. Analyzing Bicycle Sharing System User Destination Choice Preferences: An Investigation of Chicago's Divvy System, *Journal of Transport Geography*, Vol. 44, pp. 53-64.

Faghih-Imani A., N. Eluru, A. El-Geneidy, M. Rabbat, and U. Haq, 2014. How does land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal, *Journal of Transport Geography*, Vol. 41, pp. 306-314.

Fishman, E., 2016. Bikeshare: A review of recent literature, *Transport Reviews*, Vol. 36, pp. 92-113.

Fishman, E., S. Washington and N. Haworth, 2015. Bikeshare's impact on active travel: Evidence from the United States, Great Britain, and Australia. *Journal of Transport & Health*, Vol. 2, pp. 135-142.

Fuller, D., L. Gauvin, Y. Kestens, M. Daniel, M. Fournier, P. Morency, and L. Drouin, 2011. Use of a new public bicycle share program in Montreal, Canada. *American journal of preventive medicine*, Vol. 41, No. 1, pp. 80-83.

Gebhart, K. and Noland, R., 2014. The Impact of Weather Conditions on Bikeshare Trips in Washington, DC. *Transportation*, Vol. 41, pp. 1205-1225.

González, F., C. Melo-Riquelme, and L. de Grange. 2016. A combined destination and route choice model for a bicycle sharing system. *Transportation*, Vol. 43, pp.407-423.

Goodman, A., J., Green, and J. Woodcock. The role of bicycle sharing systems in normalising the image of cycling: An observational study of London cyclists. *Journal of Transport & Health*, Vol. 1, 2014, pp. 5-8.

Jäppinen, S., T. Toivonen, and M. Salonen, 2013. Modelling the potential effect of shared bicycles on public transport travel times in Greater Helsinki: An open data approach. *Applied Geography*, Vol. 43, pp. 13-24.

Kaufman, S. M., L. Gordon-Koven, N. Levenson and M. L. Moss, 2015. Citi Bike: The First Two Years. The Rudin Center for Transportation Policy and Management.

Lathia, N., S., Ahmed, and L., Capra. Measuring the impact of opening the London shared bicycle scheme to casual users. *Transportation Research Part C: Emerging Technologies*, Vol. 22, 2012, pp. 88-102.

Mahmoud, M. S., W. El-Assi, and K. Nurul Habib. 2015. Effects of Built Environment and Weather on Bike Sharing Demand: Station Level Analysis of Commercial Bike Sharing in Toronto. Paper presented at the 94th Transportation Research Board Annual Meeting, Washington, DC.

McFadden, D. Modeling the Choice of Residential Location. In *Spatial Interaction Theory and Planning Models* (A. Karlqvist et al., eds.), North Holland Publishers, Amsterdam, Netherlands,1978.

Meddin, R. and P. DeMaio, 2015. The Bike-Sharing World Map. Retrieved 20th July, 2015 from http://www.bikesharingworld.com

Murphy, E., and J., Usher, 2015. The Role of Bicycle-sharing in the City: Analysis of the Irish Experience, *International Journal of Sustainable Transportation*, Vol. 9, pp. 116-125

Nair, R., Miller-Hooks, E., Hampshire, R., and A., Busic, 2013. Large-Scale Vehicle Sharing Systems: Analysis of Velib. *International Journal of Sustainable Transportation*, Vol. 7, pp. 85-106.

NHTS, 2009. U.S. Department of Transportation, Federal Highway Administration, National Household Travel Survey 2009. URL: http://nhts.ornl.gov.

O'Brien, O., Cheshire, J., and M. Batty, 2014. Mining bicycle sharing data for generating insights into sustainable transport systems, *Journal of Transport Geography*, Vol. 34, pp. 262-273.

Rixey, R., 2013. Station-Level Forecasting of Bikesharing Ridership: Station Network Effects in Three U.S. Systems. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2387, pp. 46-55.

Rudloff, C. and Lackner, B., 2014. Modeling Demand for BikeSharing Systems: Neighboring Stations as Source for Demand and Reason for Structural Breaks. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2430, pp. 1-11.

Shaheen, S., S. Guzman and H. Zhang, 2010. Bikesharing in Europe, the Americas, and Asia Past, Present, and Future. *Transportation Research Record: Journal of the Transportation Research Board*, No. 2143, pp. 159-167.

Wang, X., G. Lindsey, J. Schoner and A. Harrison, 2015. Modeling Bike Share Station Activity: Effects of Nearby Businesses and Jobs on Trips to and from Stations. Forthcoming *Journal of Urban Planning and Development*.

Zhao, J., W. Deng and Y. Song, 2014. Ridership and effectiveness of bikesharing: The effects of urban features and system characteristics on daily use and turnover rate of public bikes in China. *Transport Policy*, Vol. 35, pp. 253-264.

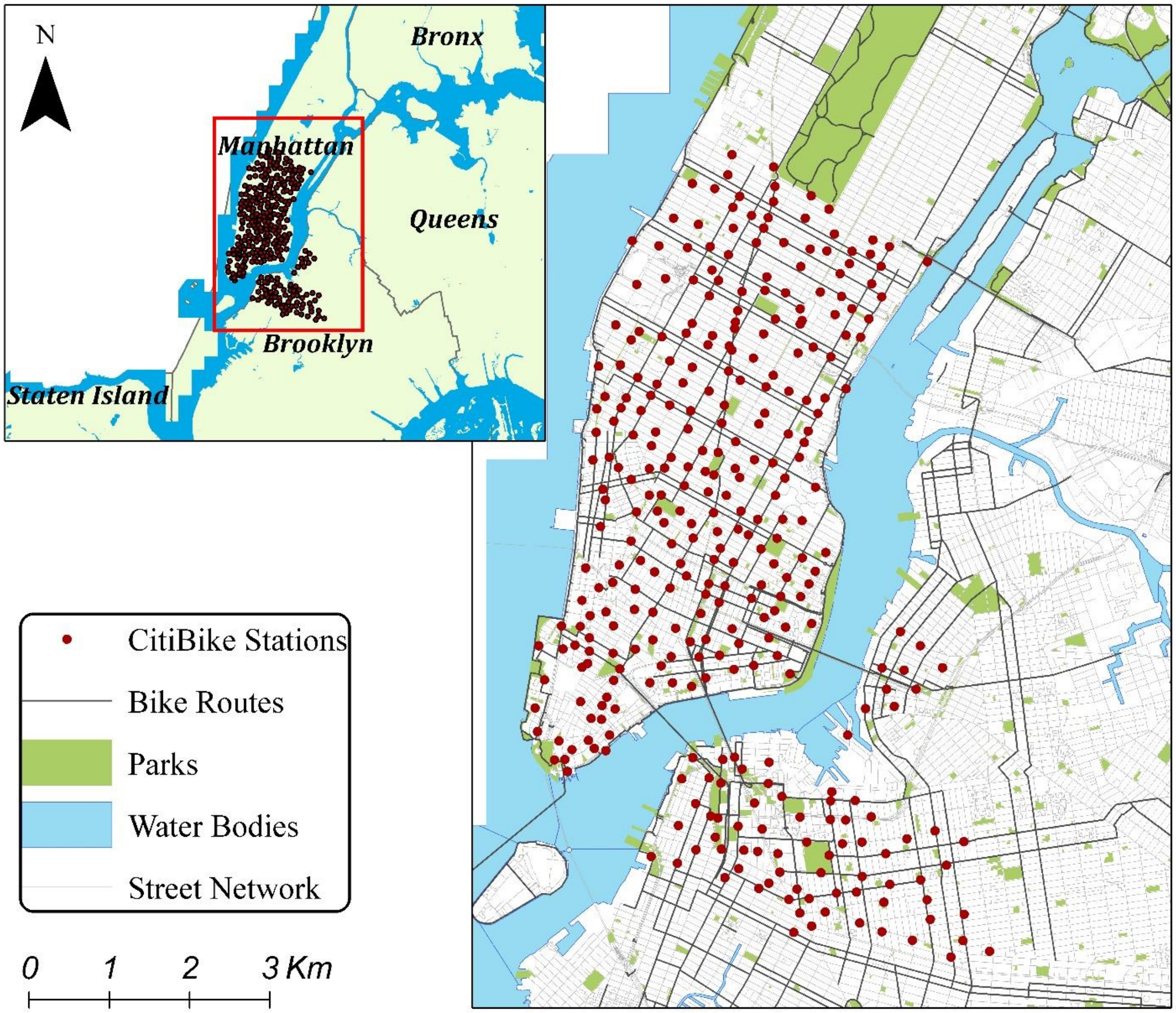

Figure 1. Map of CitiBike System in New York City

Table 1. Summary of Recent Literature on BSS

Study	Temporal Level	Spatial Level	BSS	Sample size	Analysis Level	Analysis Framework Employed
Buck & Buehler 2012	Daily	Station	Capital Bike, Washington	91	System	Linear Regression
Rixey 2013	Monthly	Station	3 Cities in US	265	System	Multiple Regression
Faghih-Imani et al. 2014	Hourly	Station	BIXI, Montreal	16400	System	Linear Mixed Model
Faghih-Imani & Eluru 2014	5 periods in a day	TAZ	BIXI, Montreal	8225	System	Panel Ordered Logit Model
Gebhart & Noland 2014	Hourly	System	Capital Bike, Washington	10968	System	Negative Binomial for Trip Rates & OLS regression for Trip Duration
Rudloff & Lackner 2014	Hourly	Station	CityBike Wien, Vienna	16489	System	Various Count Models
Zhao et al. 2014	Daily	City	Various Cities in China	69	System	Linear regression and its variants such as Partial Linear regression
Mahmoud et al. 2015	Monthly	Station	Bike Share Toronto	Station analysis: 960 OD-Pair analysis: 6316	System and Users	OLS regression
Wang et al. 2015	Daily	Station	Nice Ride, Minneapolis	116	System	OLS regression and Negative Binomial
Faghih-Imani & Eluru 2015	Trips	Station	DIVVY, Chicago	6000	Users	Multinomial Logit Model

Table 2. Descriptive Summary of base samples characteristics

Continuous Variables	Min	Max	Mean	Std. Deviation
Hourly Arrivals (Annual Members)	0	83.00	3.74	5.26
Hourly Arrivals (Daily Customers)	0	39.00	0.61	1.50
Hourly Departures (Annual Members)	0	102.00	3.74	5.36
Hourly Departures (Daily Customers)	0	34.00	0.61	1.49
Temperature (°C)	8.3	34.4	19.64	4.82
Relative Humidity (%)	27.0	94.2	60.96	16.01
Length of Bicycle Facility in 250m Buffer (m)	0	1022.7	314.95	178.82
Area of Parks in 250m Buffer (m^2)	0	95209.9	10181.87	15169.65
Number of Restaurants in 250m Buffer	0	545	54.35	92.21
Number of CitiBike stations in 250m Buffer	0	4.00	1.24	1.01
Capacity of CitiBike stations in 250m Buffer	0	169.00	43.93	38.93
Station Capacity	3.00	67.00	34.35	10.76
Pop Density (people per $m^2 \times 1000$)	0.01	67.20	24.87	14.68
Job Density (jobs per $m^2 \times 1000$)	0	432.52	55.83	53.83
Trip Distance (km)	0.05	12.68	3.97	2.30
Trip Duration (min)	1.02	89.57	12.79	8.87
Members Age	16.00	96.00	37.33	10.95

Categorical Variables	Percentage
Rainy Weather	2.6
Weekends	30.0
Subway Station in 250m Buffer	49.7
Path Train Station in 250m Buffer	4.2
Female Members	24.7

Table 3. Arrivals Models Estimation Results - Weekdays

	Base Sample	% Change in β relative to Base for Smaller Sample with the size of									
	β	10 days		5 days		3 days		2 days		1days	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Constant	-4.142	0.2	[-0.5 0.8]	0.5	[-2.7 7.9]	0.0	[-4.6 3.3]	1.1	[-0.5 3.9]	-0.9	[-6.5 5]
Built Environment Variables											
Length of Bicycle Facilities in buffer	0.041	8.0	[-2.5 12.3]	-17.8	[-23 -13.6]	6.4	[-14.8 35.5]	-11.9	[-59.2 19.8]	69.4	[40.2 119.2]
Presence of Subway Station in Buffer	0.118	2.0	[-3.8 7.2]	-2.0	[-11.3 3.6]	1.0	[-12.3 15.9]	-1.0	[-24.4 27]	5.0	[-13.1 29.6]
Presence of Path Train Station in Buffer	0.362	1.5	[-0.2 4.6]	-4.5	[-10.1 -1.2]	-4.7	[-9.3 0.8]	-2.9	[-7.3 6.7]	-2.6	[-35.3 16]
Length of Rails in Buffer	-0.024	-1.4	[-28.2 32.7]	-17.7	[-40 -3.9]	-12.9	[-79.6 87]	-9.3	[-61.8 45.6]	-17.7	[-88.1 70.6]
Area of Parks in Buffer	-2.949	-1.5	[-9.5 6.1]	0.0	[-24.6 10.9]	2.3	[-12.1 20.7]	0.2	[-38.5 24.7]	-6.3	[-39.8 12.2]
Number of Restaurants in Buffer	0.745	1.3	[-3.2 4.7]	6.4	[3.6 11.2]	-2.2	[-27.2 14.8]	-0.7	[-4.7 3]	-1.0	[-21.3 17.3]
Population Density	12.931	-0.2	[-1.9 1.2]	-0.1	[-0.7 1]	0.8	[-2.1 4.3]	-1.9	[-8.3 3.6]	-1.7	[-3.7 1.7]
Population Density*AM	-4.331	1.3	[-7.2 15.5]	7.8	[-18.8 32.5]	-0.1	[-29.1 25.7]	13.8	[-9.7 27.1]	-13.1	[-29.3 17.7]
Population Density*PM	3.103	5.1	[-5.9 20.2]	1.6	[-13.4 28.9]	-6.7	[-30.8 12.2]	1.7	[-39.1 57.1]	29.8	[1 71.6]
Job Density	1.064	-2.7	[-10.8 4.4]	-4.2	[-9.8 6.5]	-7.7	[-20.4 9.8]	-3.1	[-23.8 15.1]	6.7	[-40.8 53.5]
Job Density*AM	4.945	-1.4	[-3 3.4]	-2.2	[-11.3 2.7]	0.4	[-6.2 9.4]	2.0	[-4.2 8.5]	3.5	[-5.6 12]
Job Density*PM	0.231	15.4	[-47.3 72.4]	-22.3	[-99.5 40]	60.0	[22.9 153.3]	20.5	[-170 137.3]	-2.6	[-132.2 125.4]
Weather & Temporal Attributes											
AM	1.276	0.1	[-2.6 1.9]	0.3	[-3.8 3.6]	-0.2	[-2.5 3.2]	1.0	[-4.6 6.3]	-1.7	[-4.8 1.3]
Midday	1.103	0.3	[0 0.8]	-1.1	[-3.5 0]	-0.2	[-0.9 0.7]	0.3	[-1.3 3.5]	-0.9	[-7.5 2.3]
PM	1.470	-0.4	[-1.8 0.8]	0.0	[-1.5 2]	-0.2	[-2.8 1.1]	-1.6	[-6.4 0.8]	-2.0	[-4.8 0.8]
Even	0.790	0.1	[-0.8 0.8]	-0.1	[-0.9 1]	0.2	[-3.3 2.9]	-2.1	[-5 -0.2]	-1.8	[-4.6 -0.5]
Temperature	0.0110	2.7	[-5.5 7.7]	-2.9	[-19.2 18.8]	-1.3	[-75.5 38.6]	20.7	[-13.9 62]	-14.3	[-94 49.2]
Relative Humidity	-0.005	-0.2	[-4 3.5]	-8.4	[-78.9 18.2]	-0.2	[-22.5 23.1]	-5.1	[-12.4 2.6]	11.8	[-31.2 31.2]
Rainy Weather	-0.221	5.2	[-3.8 11.6]	9.2	[-5.5 51.4]	-6.4	[-28.1 7.6]	-4.1	[-23.3 19.8]	4.2	[-42.4 36.7]
% Change in Std.Err	-	44.9	[44 45]	105.0	[99 110]	165.7	[164 168]	226.1	[220 231]	360.2	[359 369]
Number of Insignificant Parameters	-	1.6	[1 2]	3	[3 3]	3.2	[3 4]	3.8	[3 5]	6.4	[5 8]
Parameters MAPE	-	2.54		5.47		5.68		5.25		9.85	
Parameters RMSE	-	0.044		0.085		0.141		0.082		0.183	
Prediction MAE	3.09	3.09		3.10		3.09		3.09		3.09	
Prediction RMSE	5.49	5.49		5.52		5.50		5.50		5.50	

Table 4. Departures Models Estimation Results - Weekdays

	Base Sample	% Change in β relative to Base for Smaller Sample with the size of									
	β	10 days		5 days		3 days		2 days		1days	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Constant	-4.141	0.4	[-0.6 1]	0.7	[-3.2 6.7]	0.0	[-4.3 4.4]	1.7	[-0.2 4]	0.8	[-4.5 7.2]
Built Environment Variables											
Length of Bicycle Facilities in buffer	0.063	4.8	[-1.2 15.7]	-7.5	[-21.3 6.1]	11.6	[-16.7 35.6]	-2.3	[-28.6 24]	22.5	[-17.2 66.1]
Presence of Subway Station in Buffer	0.175	2.0	[-1.2 6.8]	-0.2	[-12.2 7.7]	3.8	[-11.8 10.2]	1.1	[-15.1 12]	6.3	[-6 18.7]
Presence of Path Train Station in Buffer	0.370	-0.5	[-2.8 1.4]	-2.4	[-9.2 6.8]	-4.2	[-12.6 4.1]	-3.1	[-14.1 19]	-1.1	[-36.5 15.5]
Length of Rails in Buffer	-0.040	-5.8	[-24.8 5.9]	-6.2	[-30.7 33.5]	-2.6	[-20.6 27.7]	3.3	[-13.3 30.7]	-30.7	[-75.6 40.9]
Area of Parks in Buffer	-3.062	-3.4	[-10.5 3.3]	-2.4	[-17.9 5.2]	6.1	[-3.6 13.8]	-4.7	[-21.5 4.2]	-16.1	[-41.7 51.3]
Number of Restaurants in Buffer	0.811	1.5	[-3.5 5.5]	1.9	[-4 8.1]	-0.1	[-12 11.5]	-5.1	[-15.3 4.1]	3.7	[-22.4 32.8]
Population Density	11.951	0.6	[-2.6 3.9]	-0.3	[-2.2 1.6]	1.5	[-1.7 5.3]	-1.9	[-9.2 7.1]	2.4	[-1.8 11.8]
Population Density*AM	5.666	-3.1	[-12 10.8]	-3.3	[-33 13.8]	-12.7	[-22.9 1.6]	-21.2	[-46.1 5.1]	12.3	[-55.9 79]
Population Density*PM	-1.187	0.1	[-51.3 67.2]	-10.2	[-66.2 40.9]	-42.4	[-89.7 -5.1]	-24.3	[-104.6 34.9]	-13.3	[-87.9 79.4]
Job Density	1.348	-3.1	[-4.9 -1.2]	0.4	[-2.8 7.6]	3.9	[-0.5 9.6]	3.3	[-13.8 25.1]	3.1	[-19.6 30.9]
Job Density*AM	-0.360	5.3	[-34.2 64.2]	-10.3	[-72.9 114.6]	54.7	[-46.5 113.4]	17.6	[-105.9 175.1]	60.3	[-18.5 144.1]
Job Density*PM	3.072	1.8	[-4.8 5.5]	-1.8	[-14.1 4.5]	-2.9	[-11.8 6.8]	3.4	[-6.4 13]	3.8	[-17.4 20.3]
Weather & Temporal Attributes											
AM	1.485	0.1	[-1.8 1.7]	-0.7	[-4.2 4]	0.5	[-2.7 3.2]	2.8	[-0.9 7.3]	1.3	[-4.5 8.1]
Midday	1.125	-0.1	[-0.9 0.4]	-0.4	[-2.7 0.7]	-1.8	[-3.7 -0.8]	0.1	[-1.8 3.7]	-0.3	[-4.8 3.2]
PM	1.433	-0.5	[-2.4 0.9]	0.2	[-1.8 1.9]	-1.0	[-5 1.5]	-1.9	[-2.7 -1]	0.0	[-3.9 4.6]
Even	0.756	-0.4	[-1 0.6]	-0.2	[-1.9 1.7]	-1.4	[-4.8 1]	-3.5	[-5.6 0.5]	-0.7	[-6.3 5.9]
Temperature	0.006	8.4	[-4.4 19.9]	8.5	[-47.2 82]	-14.9	[-107.8 64]	50.3	[-7.1 132.3]	-7.6	[-89.8 95.3]
Relative Humidity	-0.005	-0.8	[-5.9 3.6]	-4.8	[-64.6 17.6]	-3.5	[-34.3 25.7]	-6.5	[-20.7 7.6]	-1.9	[-35.9 14.2]
Rainy Weather	-0.269	3.0	[-3.2 8.2]	1.9	[-18.7 48.9]	-5.5	[-13.3 3.5]	-4.5	[-36.2 17.4]	-2.9	[-61.1 50.4]
% Change in Std.Err	-	45.0	[44 45]	104.8	[99 110]	166.4	[165 168]	226.7	[218 233]	359.6	[357 370]
Number of Insignificant Parameters	-	1.2	[1 2]	2.4	[1 3]	2.8	[2 4]	3.8	[3 5]	5.8	[5 6]
Parameters MAPE	-	2.28		3.21		8.75		8.13		9.56	
Parameters RMSE	-	0.032		0.047		0.165		0.143		0.171	
Prediction MAE	3.06	3.06		3.07		3.07		3.06		3.06	
Prediction RMSE	5.59	5.60		5.61		5.61		5.59		5.58	

Table 5. Arrivals Models Estimation Results - Weekends

	Base Sample	% Change in β relative to Base for Smaller Sample with the size of							
	β	5 days		3 days		2 days		1days	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range
Constant	-4.035	-0.2	[-1.1 1]	0.4	[-0.3 1.5]	-0.3	[-2.8 5.5]	2.2	[-2.2 5.5]
Built Environment Variables									
Length of Bicycle Facilities in buffer	0.065	5.0	[-4.2 16]	-10.3	[-26.6 12.3]	-20.2	[-30.7 -1.6]	12.8	[-8.9 58.9]
Presence of Subway Station in Buffer	0.040	-7.7	[-15.9 1.5]	-8.8	[-41.8 20.7]	-23.5	[-64.5 5]	0.3	[-48.4 41]
Presence of Path Train Station in Buffer	0.327	1.3	[-6.4 6.3]	-2.1	[-13.5 16.5]	13.3	[-0.3 32.3]	-10.4	[-53.2 19.8]
Area of Parks in Buffer	-1.730	3.0	[-32.2 27.5]	-9.9	[-32.9 8.2]	3.2	[-11.6 24.1]	-20.4	[-79.7 37.5]
Number of Restaurants in Buffer	0.804	3.2	[-2 8.9]	-3.0	[-10.4 14.1]	0.4	[-18.7 10.1]	-1.0	[-9.4 14.3]
Population Density	15.211	-0.8	[-2.2 0.7]	0.0	[-2.4 3]	-0.6	[-10.3 6]	-4.5	[-9.4 -0.6]
Population Density*AM	-2.266	10.1	[-0.1 18]	-0.6	[-67.8 23.5]	-13.5	[-49.7 22.3]	11.7	[-59.5 56.1]
Population Density*PM	3.212	7.9	[-12.9 31.4]	-2.8	[-17.8 25.5]	5.5	[-24.2 37.1]	15.5	[-30.8 57.5]
Job Density	-0.498	2.5	[-5.2 13.4]	0.9	[-16.7 30.2]	-1.5	[-28.9 29.1]	-39.4	[-72.8 7.8]
Job Density*AM	0.656	-1.8	[-11.3 15.9]	8.6	[-15 46.6]	-17.2	[-84.7 80.9]	7.5	[-31.2 55.2]
Weather & Temporal Attributes									
AM	0.656	0.8	[-1.6 3.6]	-1.2	[-9.5 2.4]	-0.4	[-8 4.5]	-0.3	[-6.7 9.3]
Midday	1.458	-0.1	[-1.3 0.8]	-0.4	[-0.8 0]	0.3	[-0.8 1.2]	0.2	[-2.7 2.8]
PM	1.199	-1.2	[-2.2 -0.1]	-0.5	[-2.7 1.7]	0.6	[-2.8 3.9]	-1.2	[-9.3 3.6]
Even	0.639	-0.7	[-1.9 0.6]	0.8	[-2.4 4.8]	0.7	[-1.2 3.9]	0.0	[-2.5 6.2]
Temperature	0.022	2.0	[-4.8 8.3]	4.6	[-3.1 12.3]	-1.6	[-19.1 41.6]	9.4	[-38.1 46.6]
Relative Humidity	-0.010	1.9	[-2.4 5.5]	-1.0	[-3 2.5]	-1.8	[-13.1 4.3]	-6.6	[-9.9 -2.2]
Rainy Weather	-0.204	-1.2	[-13.6 17]	9.1	[1.2 20.3]	4.3	[-19.3 20.6]	-7.5	[-22.1 26.5]
% Change in Std.Err	-	33.9	[33 35]	73.7	[73 75]	112.4	[112 115]	203.3	[200 206]
Number of Insignificant Parameters	-	1.2	[1 2]	3.4	[3 4]	5.2	[4 6]	6.2	[6 7]
Parameters MAPE	-	2.9		3.6		6.0		8.4	
Parameters RMSE	-	0.040		0.052		0.097		0.127	
Prediction MAE	2.21	2.22		2.22		2.21		2.21	
Prediction RMSE	3.71	3.71		3.72		3.70		3.70	

Table 6. Departures Models Estimation Results - Weekends

	Base Sample	% Change in β relative to Base for Smaller Sample with the size of							
	β	5 days		3 days		2 days		1days	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range
Constant	-4.122	-0.5	[-1.2 0.1]	0.1	[-1.7 2.6]	-0.1	[-3.2 6.2]	1.5	[-1 5.2]
Built Environment Variables									
Length of Bicycle Facilities in buffer	0.090	-0.2	[-3.2 3.9]	-6.1	[-20.1 3.5]	-5.5	[-17.9 7]	19.1	[-2.3 43.1]
Presence of Subway Station in Buffer	0.100	-2.0	[-10.3 9.4]	-3.0	[-23.1 14.3]	-2.6	[-31.4 19.4]	5.3	[-5.4 10.3]
Presence of Path Train Station in Buffer	0.324	2.8	[-7.9 13]	-6.0	[-16.5 2.1]	8.3	[-1.9 23.2]	-24.0	[-60.8 4.9]
Length of Rails in Buffer	-0.036	-7.0	[-35.4 5]	-11.4	[-31.2 3.9]	-31.2	[-47.9 -6.4]	5.1	[-22.8 38.9]
Area of Parks in Buffer	-1.796	-3.1	[-33.4 14.1]	-10.0	[-39.6 21.2]	-1.5	[-23.9 38.8]	-5.3	[-78.3 50.3]
Number of Restaurants in Buffer	0.894	2.4	[-2 6.8]	-2.6	[-10.3 6.4]	-5.2	[-19.8 11.9]	-3.4	[-21.5 7.9]
Population Density	13.100	-0.2	[-1.3 0.9]	-0.1	[-4.7 7.1]	2.0	[-6.4 11.5]	-4.3	[-7.7 -0.8]
Population Density*AM	5.017	-4.7	[-14.7 8.7]	10.5	[1.1 21.3]	-16.0	[-28.1 -1.8]	1.7	[-57.9 44.4]
Job Density	-0.379	8.1	[-4.3 30]	-0.4	[-25.7 37.4]	-4.5	[-46.1 8.5]	-59.1	[-138.2 -12.1]
Job Density*PM	0.738	-0.5	[-13.4 32.2]	-1.1	[-37.7 41.1]	18.2	[-17.6 83.2]	-28.2	[-79.4 8.2]
Weather & Temporal Attributes									
AM	0.610	0.3	[-0.9 1.5]	-3.0	[-5.4 -0.1]	3.4	[1.2 6.5]	1.4	[-7.6 13.2]
Midday	1.546	0.2	[-0.1 0.5]	-0.4	[-1.5 0.7]	-0.5	[-2 1.2]	0.2	[-3.4 3.4]
PM	1.298	-0.6	[-1.9 0.5]	-0.2	[-1.6 1.4]	0.4	[-3.5 1.9]	2.6	[-1.3 5.5]
Even	0.650	0.4	[-0.7 1.6]	0.1	[-3.6 1.6]	0.4	[-1.8 1.6]	1.8	[-1.2 5.8]
Temperature	0.017	-0.6	[-7.3 4.7]	0.7	[-9.5 14.6]	-7.1	[-37.7 43.6]	1.8	[-19.7 34]
Relative Humidity	-0.009	2.6	[0.5 9.2]	-1.8	[-11.4 4.9]	-3.1	[-20.7 6.5]	-6.9	[-32 2.8]
Rainy Weather	-0.224	-5.0	[-12.6 -1.3]	3.8	[-9.3 21.5]	-8.0	[-20.6 -0.6]	-7.0	[-44.2 22.1]
% Change in Std.Err	-	34.0	[33 34]	74.0	[73 75]	113.4	[112 117]	204.6	[202 207]
Number of Insignificant Parameters	-	1.8	[1 3]	3	[2 4]	3.4	[2 4]	5.4	[5 6]
Parameters MAPE	-	2.3		3.4		6.5		9.9	
Parameters RMSE	-	0.033		0.050		0.101		0.174	
Prediction MAE	2.19	2.20		2.20		2.19		2.19	
Prediction RMSE	3.65	3.65		3.65		3.65		3.64	

Table 7 Destination Choice Models Estimation Results - Weekdays

	Base Sample	% Change in β relative to Base for Smaller Sample with the size of											
Parameters	**β**	**20000**		**10000**		**5000**		**3000**		**2000**		**1000**	
		Mean	**Range**	**Mean**	**Range**	**Mean**	**Range**	**Mean**	**Range**	**Mean**	**Range**	**Mean**	**Range**
Built Environment Variables													
Station Capacity	0.017	-0.8	[-5.8 2.9]	-0.7	[-5.2 7.5]	-1.2	[-6.4 6.4]	-3.5	[-22 24.3]	-12.6	[-30.6 11]	24.9	[7.5 38.2]
Presence of Subway Station in Buffer	0.057	15.4	[-10.1 34.6]	-22.0	[-50.8 3.7]	-2.6	[-62.1 69.3]	28.4	[-57.2 106.3]	4.4	[-70.9 150.8]	-7.6	[-49.6 25.8]
Presence of Path Train Station in Buffer	0.109	4.4	[-21.9 20.7]	-17.2	[-46 25.8]	5.6	[-49.1 113.9]	4.9	[-108.3 118.3]	-2.8	[-71.2 77.4]	11.0	[-115.8 178.9]
Length of Rails in Buffer	-0.086	-12.0	[-28.5 2.7]	13.1	[-23.9 45.2]	-4.0	[-24.2 29.9]	-34.6	[-63.9 15.8]	-16.5	[-48.5 7.9]	-44.5	[-72.8 4.8]
Number of Restaurants in Buffer	0.465	-1.8	[-27.1 14.3]	7.2	[-7.9 30.5]	16.9	[-16.2 41.2]	-3.5	[-76.1 59.6]	33.9	[-8.9 99.7]	-13.2	[-160.6 131.4]
Population Density	3.513	-9.3	[-35.8 10.4]	6.0	[-11.2 35.5]	-12.3	[-67 28.3]	-7.7	[-63.8 26.3]	-10.8	[-99.7 58.9]	82.9	[-14.2 134.7]
Population Density*AM	-11.899	-7.3	[-13.1 1.7]	13.7	[-0.2 33.4]	-0.8	[-22.2 34.9]	-8.4	[-11.9 -2.2]	-2.7	[-63.5 33.7]	15.7	[-43.9 56.5]
Population Density*PM	2.903	-0.2	[-38.7 40.6]	-19.4	[-63.8 37.2]	24.8	[-61.7 127.5]	-14.9	[-74.4 111.2]	57.6	[-34.6 136]	-48.0	[-153.1 56.1]
Job Density	-0.360	-16.6	[-78.9 55.2]	4.2	[-21.3 47.1]	1.1	[-107.4 73.1]	-101.4	[-443.7 92.2]	75.6	[-134.1 259]	104.5	[-103.6 299.1]
Job Density*AM	4.374	1.0	[-6.9 8.1]	-1.5	[-6.3 3.2]	0.9	[-9.9 24.4]	-8.0	[-36.7 24]	18.2	[3.6 25.9]	7.7	[-22.2 55.3]
Job Density*PM	-1.941	3.6	[-10.6 12.6]	2.3	[-11 21]	0.5	[-32.4 24.6]	18.2	[-55.1 159.8]	-40.5	[-64.4 15.1]	-2.8	[-112.5 107.2]
Trip Attributes													
Distance	-0.484	-1.9	[-13.4 10]	-0.3	[-4.4 3.3]	-13.2	[-35.6 4.9]	-0.2	[-21 19.5]	-4.3	[-39.5 24.3]	-16.3	[-74 31.2]
Distance*Female	0.025	-5.0	[-57.5 34]	-15.9	[-75.7 73.7]	29.4	[-96.8 116.6]	-55.9	[-125.1 2.8]	64.1	[-140.1 318.2]	-43.6	[-167.2 127.9]
Distance*Temperature	0.119	-8.3	[-134.5 112.2]	-24.0	[-71.2 66.2]	-126.0	[-447.7 88.2]	60.9	[-317.1 519.6]	-36.8	[-588.6 322.8]	-119.4	[-439.8 121.4]
Distance*Humidity	-0.105	12.1	[-35.2 49.1]	-6.9	[-46.8 38.3]	64.2	[13.4 133.8]	29.8	[-6.9 85.8]	21.3	[-84.9 102.2]	97.4	[-282.5 448.1]
Distance*Rainy	-0.136	-20.0	[-47.7 13.3]	42.2	[-11.8 131.8]	39.5	[-106.7 377.3]	53.4	[-10.7 130]	-39.4	[-142.6 84.4]	59.7	[-124.8 679.7]
% Change in Std.Err	-	57.8	[54 60]	124.7	[120 135]	217.3	[200 251]	309.4	[276 355]	394.2	[380 404]	615.2	[525 677]
Number of Insignificant Parameters	-	2.6	[2 3]	5.6	[3 7]	6.2	[4 9]	8.2	[7 9]	10	[9 12]	11	[9 12]
Parameters MAPE	-	7.5		12.3		21.4		27.1		27.6		43.7	
Parameters RMSE	-	0.096		0.164		0.384		0.385		0.355		0.575	
Predictive LL	-14799.5	-14801.3		-14801.4		-14808.8		-14814.7		-14819.9		-14850.4	
% of Correct Prediction	11.88	11.80		11.82		11.86		11.84		11.88		11.86	

Table 8 Destination Choice Models Estimation Results - Weekends

Parameters	Base Sample β	% Change in β relative to Base for Smaller Sample with the size of 20000		10000		5000		3000		2000		1000	
		Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range	Mean	Range
Built Environment Variables													
Station Capacity	0.015	1.8	[0 3.4]	-3.9	[-10.1 4.7]	-0.7	[-8.8 12.8]	0.7	[-18.2 13.5]	-8.6	[-25.7 1.4]	-2.2	[-58.1 28.4]
Presence of Path Train Station in Buffer	0.072	4.0	[-16.9 37.6]	-45.4	[-81.3 -10.4]	-12.8	[-100 105.6]	13.9	[-76.7 113.6]	-3.3	[-144.7 128.7]	-138.9	[-303.8 -52.8]
Area of Parks in Buffer	1.744	-27.0	[-64.8 3.3]	-8.5	[-44.1 26.9]	22.2	[-26.9 76.3]	-45.1	[-121.7 53.8]	-13.3	[-156.1 222.8]	-88.2	[-194.7 -31.9]
Length of Rails in Buffer	-0.060	-17.9	[-36.5 -2.5]	-15.6	[-43.8 12.2]	-5.5	[-90.3 51.2]	-30.8	[-105.7 -1.8]	30.2	[-98.2 88.7]	98.5	[-79.8 345.7]
Number of Restaurants in Buffer	0.663	-0.2	[-21.7 7.8]	-2.6	[-15.1 8.7]	-1.1	[-27.1 18.1]	-29.2	[-51.3 14.8]	-4.1	[-23 19.4]	14.2	[-96.8 95.2]
Population Density	6.208	2.8	[-8.3 12.5]	0.1	[-18.4 29.4]	-7.5	[-37.8 11.9]	15.7	[-22.7 62.3]	-9.3	[-60.7 47.8]	-9.3	[-212.9 80.4]
Population Density*AM	-8.534	3.6	[-19.6 21.1]	2.4	[-35.9 25]	-4.3	[-49.2 29.9]	4.6	[-126.7 124.4]	-13.3	[-102.5 46.3]	8.7	[-100.6 108.7]
Population Density*PM	3.210	-14.9	[-42.7 26.9]	-13.5	[-103 13.3]	-10.6	[-111.9 44.5]	-11.3	[-115 87.6]	17.5	[-163.3 244.9]	-20.5	[-198.7 317.9]
Job Density	-2.147	-5.7	[-19.7 4]	-2.2	[-23.5 17.2]	-14.8	[-36.5 0.5]	-10.8	[-44 15.9]	-16.8	[-35.4 1.2]	2.5	[-37.4 58.2]
Job Density*AM	1.261	1.4	[-38.5 39.8]	0.3	[-21 59.7]	-18.3	[-111.1 52.5]	-1.5	[-199.7 193.5]	-84.8	[-287.4 110.7]	45.2	[-322.4 271.6]
Job Density*PM	-1.006	25.0	[-10.9 62.9]	-25.0	[-80.9 50.6]	11.9	[-59 100.1]	-38.1	[-186.5 145.8]	-50.5	[-228.3 92.8]	-19.0	[-215.8 152.2]
Trip Attributes													
Distance	-0.540	-2.2	[-6.3 2.5]	1.2	[-4.9 5.4]	-6.6	[-17.4 4.5]	-0.9	[-7.2 6.3]	-12.8	[-32.3 -4.1]	1.8	[-24.6 44.1]
Distance*Female	0.031	-18.8	[-23.5 -12.9]	-28.6	[-100 39.4]	18.2	[-48.1 94.5]	12.9	[-162.9 263.5]	-66.1	[-179 51.3]	86.5	[-111.3 302.3]
Distance*Humidity	-0.043	22.5	[-73.1 107.7]	-55.9	[-151.3 29.3]	118.8	[-78.2 305.4]	25.1	[-206.6 139.1]	247.2	[-42.2 664.4]	27.6	[-679.4 448.9]
Distance*Rainy	-0.124	-14.0	[-113.9 31.8]	25.1	[-82.9 225.6]	-2.7	[-107.1 155]	70.2	[-159.7 479.1]	52.1	[-264.9 479.6]	183.2	[-330.9 569.8]
% Change in Std.Err	-	57.1	[50 60]	123.4	[117 133]	213.8	[200 222]	309.0	[284 350]	402.5	[380 434]	624.9	[572 731]
Number of Insignificant Parameters	-	2.8	[1 4]	5.8	[5 6]	7.6	[5 8]	8.8	[8 10]	9.4	[9 10]	10.4	[9 12]
Parameters MAPE	-	10.1		14.4		16.0		19.4		39.4		46.6	
Parameters RMSE	-	0.142		0.228		0.327		0.280		0.731		0.738	
Predictive LL	-14596.0	-14597.0		14599.5		14602.1		14608.0		14608.8		14635.1	
% of Correct Prediction	12.98	13.02		13.08		12.83		12.84		12.83		12.79	