

How land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal

Ahmadreza Faghieh 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 (Corresponding author)
Assistant Professor

Department of Civil Engineering and Applied Mechanics
McGill University
Ph: 514-398-6823, Fax: 514-398-7361
E-mail: naveen.eluru@mcgill.ca

Ahmed M. El-Geneidy
Associate Professor
School of Urban Planning
McGill University

Ph.: 514-398-8741, Fax: 514-398-8376
E-mail: ahmed.elgeneidy@mcgill.ca

Michael Rabbat
Associate Professor

Department of Electrical and Computer Engineering
McGill University
Ph.: 514-398-1847, Fax: 514-398-4470
E-mail: michael.rabbat@mcgill.ca

Usama Haq

Undergraduate Student
Department of Electrical and Computer Engineering
McGill University
Ph.: 514-398-1847, Fax: 514-398-4470
E-mail: usama.haq@mail.mcgill.ca

February 2014

For citation please use: Imani, A., Eluru, N., El-Geneidy, A., Rabbat, M. & Haq, U. (accepted). How does land-use and urban form impact bicycle flows: Evidence from the bicycle-sharing system (BIXI) in Montreal. *Transport Geography*. <http://dx.doi.org/10.1016/j.jtrangeo.2014.01.013>

ABSTRACT

Installed in 2009, BIXI is the first major public bicycle-sharing system in Montreal, Canada. The BIXI system has been a success, accounting for more than one million trips annually. This success has increased the interest in exploring the factors affecting bicycle-sharing flows and usage. Using data compiled as minute-by-minute readings of bicycle availability at all the stations of the BIXI system between April and August 2012, this study contributes to the literature on bicycle-sharing. We examine the influence of meteorological data, temporal characteristics, bicycle infrastructure, land use and built environment attributes on arrival and departure flows at the station level using a multilevel approach to statistical modeling, which could easily be applied to other regions. The findings allow us to identify factors contributing to increased usage of bicycle-sharing in Montreal and to provide recommendations pertaining to station size and location decisions. The developed methodology and findings can be of benefit to city planners and engineers who are designing or modifying bicycle-sharing systems with the goal of maximizing usage and availability.

Keywords: Bicycle-sharing systems, BIXI Montreal, BIXI arrivals and departures, linear mixed models, bicycle infrastructure, land use and built environment

1. Introduction

In recent years, there has been growing attention on bicycle-sharing systems as an alternative and complementary mode of transportation. These systems are recognized to have traffic and health benefits such as flexible mobility, physical activity, and support for multimodal transport connections (Shaheen et al., 2010). A bicycle-sharing system is intended to provide more convenience because individuals can use the service without the costs and responsibilities associated with owning a bicycle for short trips within the service area of the system. Further, a bicycle-sharing system frees individuals from the need to secure their bicycles; bicycle theft is a common problem in urban regions (van Lierop et al., 2013; Rietveld and Daniel, 2004). Another advantage associated with this system is that the decision to make a trip by bicycle can be made in a short time frame.

Currently, there are more than 4 hundred thousand public bicycles around the world and 400 cities have installed or are planning to install a bicycle-sharing system (Fishman et al., 2013). BIXI (a word formed by combining bicycle and taxi) was one of the first major public bicycle-sharing systems in North America. It was installed in 2009 in Montreal, Canada. The service began with 3000 bicycles and 300 stations. In 2012, the BIXI system had 410 stations with more than 4000 bicycles. Although bicycle-sharing systems are becoming more and more common around the world, there are relatively few studies exploring the factors affecting shared bicycle flows and usage. Fishman et al. (2013), after an extensive literature review, concluded that in order to better understand and maximize the effectiveness of bicycle-sharing programs, the evaluation of current performance of bicycle-sharing systems is crucial. Demand modeling plays an important role in determining the required capacity, and hence the success of new bicycle-sharing systems and/or the success of expanding an existing system. BIXI in Montreal is a mature system that offers a unique opportunity for understanding the factors influencing its flows and usage.

In this study, using data compiled from minute-by-minute readings of bicycle availability at all 410 stations on the BIXI website between April and August 2012, we attempt to examine the determinants of bicycle-sharing demand in Montreal. The BIXI database compiled is augmented with meteorological data, temporal characteristics, bicycle infrastructure, land use, and built environment attributes allowing us to examine the influence of these factors on bicycle-sharing system demand. Specifically, the main objective of the current paper is to quantify the influence of various factors on arrival and departure flows at the bicycle sharing station level using a general statistical modeling technique that other regions can adopt. The study employs a multilevel linear mixed modeling approach that explicitly recognizes the dependencies associated with bicycle flows originating at the same station. The model results obtained are validated using operational data compiled from 2013 (one year after the data used to fit the model). Further, we compute elasticity estimates of various attributes to illustrate the applicability of the developed model for policy analysis.

The rest of the paper is organized as follows. Section 2 provides a literature review of earlier research and positions our research. Section 3 explains the data compilation and sample formation in detail. Section 4 presents the visual representation of BIXI flows. The statistical model employed in this paper and the model estimation results are discussed in section 5. Section 6 discusses a policy exercise. Finally, Section 7 concludes the paper with recommendations for future research.

2. Literature Review

The first bicycle-sharing system was introduced in the 1960s in the Netherlands (DeMaio, 2009; Shaheen et al., 2010). Since then, there have been four generations of these systems. The first generation was “white bicycles” or free bicycles available in different locations around the city. The idea was simple: a person would pick up one of the bicycles, which were typically painted in bright colors and unlocked, ride it to his or her destination, and leave it there for the next possible user. It was free and without any time constraint. This program failed because of many stolen and vandalized bicycles. In the 1990s, a second-generation coin-deposit system was introduced as a result of the experience of the first generation of bicycle-sharing systems. Locked bicycles could be borrowed with a small deposit, which was usually refunded on return. Unfortunately, this did not eliminate the issue of bicycle theft due to user anonymity (Shaheen et al., 2010). Also, no time limit for the use of bicycles resulted in excessively long rental periods for borrowed bicycles. The third generation system added transaction kiosks to docking stations to solve the problem of user anonymity. People could rent a bicycle for only a limited amount of time. These systems became relatively successful around the world. Fourth generation systems, also called demand-responsive multimodal systems, have been built on the success of the third generation, while also improving docking stations, bicycle redistribution, and integration with other transport modes (DeMaio, 2009; Shaheen et al., 2010). BIXI belongs to the latest generation of bicycle-sharing systems. The BIXI system aggregated more than 3.4 million trips in the 2010 season (PBSC, 2013).

Over the past few years there have been several studies devoted to examining factors affecting bicycle-sharing flows and usage. A subset of these studies conducted a feasibility analysis, proposing different bicycle-sharing programs for different cities (for example, see Gregerson et al., 2010). These studies typically aim to identify potential locations for stations and to estimate bicycle-sharing flows and usage considering socio-demographic and land-use variables (such as population and job density) as well as topological and meteorological parameters for the proposed locations. There are relatively few quantitative studies on bicycle-sharing systems employing actual bicycle usage data. Nair et al. (2013) investigated several aspects of such systems including system characteristics, utilization patterns and the connection with public transit using data from the Velib’ bicycle-sharing system in Paris, France. Buck and Buehler (2012) explored the influence of various factors — including bicycle lanes, population, number of households without a car, and retail destinations around the stations — on bicycle flows of the Capital bicycle-sharing system in Washington DC. Krykewycz et al. (2010) estimated demand for a proposed bicycle-sharing program in Philadelphia using observed bicycle flow rates in European cities. Rixey (2013) investigated the effects of demographic and built environment characteristics on average monthly bicycle usage in three different cities in the US at the station level using a regression analysis. He concluded that population density, job density, income levels, and the share of alternative commuters are all critical factors affecting bicycle-sharing ridership. The same approach has been applied by Daddio (2012) to the bicycle-sharing system in Washington DC. Wang et al. (2012), in their analysis, considered annual rates for each station and examined the effects of nearby business and job densities, socio-demographics, built environment, and transportation infrastructure variables on annual usage flows. They found that locating stations closer to jobs results in higher usage of the bicycle-

sharing system. Moreover, the presence of food-related businesses near stations has a more positive impact on arrivals and departures than non-food commercial businesses.

The objective of our research effort is similar to these previous studies. However by using aggregated monthly or yearly flow rates, these studies fail to capture the impact of variables that change in the short term; i.e., at an hourly level (such as variations in weather and time-of-day effects). Neglecting the presence of such variations usually reduces the applicability of the results obtained. Moreover, examining bicycle flows at an hourly level (or a short time frame) allows the analyst to provide the operators with bicycle demand profiles including excess and shortage information. A more recent research effort, Hampshire et al. (2013), studied the influence of bicycle infrastructure attributes and land-use characteristics on bicycle flows using aggregated hourly arrival and departure rates at the *sub-city district* (SCD) level in Barcelona and Seville, Spain. They highlighted that bicycle station density, the average capacity of stations in the SCD, and the number of points of interest in SCD are important contributors to arrival and departure rates. Contrary to the previously mentioned literature, while Hampshire et al. (2013) used a fine temporal dimension, their study fails to capture fine-grained spatial effects because the station flows studied are aggregated at the SCD level.

There have been several studies conducted using the BIXI system. These studies use survey data rather than actual bicycle flow data obtained from stations. They contribute to the literature by studying user behavior in response to bicycle-sharing systems and examine the integration of this system with public transit (Bachand-Marleau et al., 2011; Bachand-Marleau et al., 2012; Fuller et al., 2011)¹.

The current paper contributes to literature by determining the effect of meteorological data, temporal characteristics, bicycle infrastructure, land use and urban form attributes on bicycle arrival and departure flows at the station level using real data. The estimated models will allow us to predict changes to the demand profiles (arrivals and departure flows) allowing us to examine the influence of changes to the system – capacity reallocation or new station installation.

3. Data

For this study, the hourly arrival and departure rates are obtained from minute-by-minute BIXI bicycle availability data for all stations in service (410 stations) between April and August 2012. Figure 1 shows the location of BIXI stations on the Montreal Island. It is important to note that, due to severe winter conditions in Montreal, the BIXI season starts on April 15th and ends on November 15th of each year.

A sample formation exercise was necessary to obtain the arrival and departure rates from the bicycle availability data for every station. The raw data saved from the BIXI website provided information on the number of bicycles available at each station for every minute. The raw data was processed to generate minute-by-minute bicycle arrival and departure rates for every station. The arrival and departure rates obtained are not necessarily due to customer-based bicycle flows. It is important to note that bicycle-sharing system operators frequently perform rebalancing operations, removing bicycles from stations that are full and refilling the docks of

¹ A stream of studies concentrates on operational issues of bicycle-sharing systems such as maximizing efficiency of operator rebalancing program (Borgnat et al., 2011; Froehlich et al., 2009; Jensen et al., 2010; Kaltenbrunner et al., 2010; Vogel and Mattfeld, 2011). The focus of these studies is more on optimizing bicycle repositioning operations to remove bicycles from full stations and refill empty stations, and it is not particularly related to our research effort.

empty stations. Unfortunately, the occurrence of rebalancing operations is not indicated in the minute-by-minute data available, and so it is not possible to directly distinguish whether the addition (removal) of bicycles is due to customers or operators. So, we adopt a heuristic mechanism to arrive at the “true” arrivals and departures. We identify spikes of bicycle availability (or removal) in the data compiled to differentiate between customer flows and operator flows. For this purpose, we aggregate the flow rate data temporally up to a 5-minute level to capture the effect of rebalancing operations. Specifically, we assume that a rebalancing operation has occurred if the 5-minute arrival/departure rate is greater than the 99th percentile arrival/departure for that station. When such a trigger is identified, the actual bicycle flow for this 5-minute period is obtained by averaging the bicycle flow rates of the two earlier 5-minute periods and the remainder of the flow is allocated to the rebalancing operation (a slight variant of this approach is employed in Hampshire et al., 2013). After correcting for rebalancing operations, hourly arrival and departure rates for every station are obtained by aggregating this 5-minute bicycle flow data.

Although the BIXI season starts April 15th every year, only a subset of the stations begin functioning within the first ten days of the season. Hence, from 2012 BIXI data, we removed the month of April and restricted our sample to the four months of May, June, July and August. Subsequently, to obtain a reasonable sample size, we randomly select two days for every station in our database. The arrival/departure rates in overnight hours (1 AM to 6 AM) are very low. Thus, we aggregate the bicycle flow rates in the overnight time period as one record, generating 20 records for every day (one for the period 1 AM to 6 AM, and one for each remaining hour of the day). Further, to account for the influence of station capacity on bicycle flows, we normalized our dependent variable (arrivals or departures at a station) with station capacity. The final sample consisted of 16400 records (20 hours \times 2 days \times 410 stations) of normalized arrival and departure rates at a station level. The data sample compiled is well distributed across the four months (percentages of April, May, June and July range between 22.4 and 26 percent) and across all 7 days in a week (daily shares range from 12.8 to 15.6 percent). To be sure, the data sample employed in our analysis forms a small share of the entire data compiled. If the objective is to estimate a linear regression model, large sample size would not be an issue. However, in our paper, we estimate a linear mixed model (described in Section 5) whose structure results in longer model run times for larger samples. Further, employing very large samples for model estimation might result in data over-fit and inflated parameter significance. Two separate models are developed to examine the arrival rates and departure rates at every station.

3.1. Independent variable generation

The independent variables considered in our analysis can be categorized into three groups: (1) weather, (2) temporal and (3) spatial variables. Weather variables include hourly temperature, relative humidity, and the hourly weather condition represented as a dummy variable indicating whether or not it is raining. The temporal variables considered aim to capture time-of-day and day-of-the-week effects. Specifically, the day is divided into four periods: morning (6AM-10AM), mid-day (10AM-3PM), PM (3PM-7PM) evening (7PM- 12AM). The influence of weekend vs. weekday was also taken into account. Further, to account for young individual users in the downtown core of Montreal, we included a Friday and Saturday night dummy variable to test for possible increase in BIXI usage during these periods compared to other times.

To examine the spatial determinants influencing bicycle usage at each station, two classes of spatial variables were used: a) Bicycle infrastructure and b) Land-use and built environment variables. The bicycle infrastructure variables included are at both the *traffic analysis zone* (TAZ²) level and the buffer level. A 250 meter buffer around each station was found to be an appropriate walking distance considering the distances between BIXI stations. Bicycle infrastructure variables were used to examine the effect of cycling facilities on the bicycle demand and usage of the bicycle-sharing system. The length of bicycle facilities (bicycle lanes and bicycle paths) in the buffer was calculated to capture the impact of placing BIXI stations near bicycle facilities on the usage of the bicycle-sharing system. Moreover, the length of minor roads (local streets and collectors) and major roads (arterials and highways) in the buffer were calculated to identify cyclist preference of routes. The number and capacity of BIXI stations in the 250 meter buffer were computed to capture the effect of neighbouring stations.

Land-use and built environment characteristics are the other group of variables considered in our analysis. To study the influence of the *central business district* (CBD), the distance from each station to the CBD was computed. The walkscore corresponding to every station is also generated³. The presence of metro and bus stations near a BIXI station and the length of bus lines in the 250 meter buffer were generated to examine the influence of public transit on bicycle arrival and departure rates. We also considered three types of points of interest near each station: (1) the number of restaurants (including coffee shops and bars), (2) the number of other commercial enterprises and (3) a categorical variable indicating whether or not the BIXI station is near a university. The TAZ level variables considered in our analysis include population density and job density of the TAZ associated with each BIXI station. To provide an illustration of the data compiled, we provide a descriptive summary of the sample in Table 1.

4. Visual Representation of BIXI flows

In order to better understand the spatial and temporal variation of bicycle usage in the BIXI system, we represent the bicycle arrival and departure rates of every station visually using a geographic information system. For this purpose, the bicycle flows of every station in every day of June were considered. To conserve space, we mainly focus on the AM and PM time periods in our visualization exercise. We compute the average hourly arrival and departure flows at every station for the AM and PM time periods. The patterns are presented in Figure 2. Several interesting observations can be made from the results. First, we can see that flows are much higher for the BIXI system during the PM period. One plausible explanation for the trend is that employed individuals might find it easier to bicycle home since they are presumably not in as much of a rush as when going to work in the morning. These individuals might decide to arrive at work using less strenuous modes (such as bus or metro). Furthermore, people might also consider riding the BIXI as a useful exercise after work or might make short trips within the CBD — for instance, going from work to a restaurant. It is also possible that during the evening peak hour the population using BIXI includes students and other individuals without the typical schedule (e.g., workers in restaurants and coffee shops, and non-workers). Second, the higher concentration of arrival rates in CBD in the morning peak hour confirms the use of bicycle-sharing system for daily commute purposes. Third, the results indicate that bicycle flows are

² Traffic Analysis Zone represents the unit of demarcation for urban metropolitan planning purposes.

³ The walkscore is a walkability index based on the distance to amenities such as grocery stores, restaurants, etc (see Carr et al., 2011 and <http://www.walkscore.com/> for more information).

more spatially widespread in the evening peak compared to morning peak. Overall, the visualization provides a brief overview of bicycle flows in Montreal using the BIXI system.

5. Analysis and Discussion

5.1. Linear mixed models

The most common methodology employed to study continuous dependent variables such as arrival and departure flows is the linear regression model. However, the traditional linear regression model is not appropriate to study data with multiple repeated observations. In our empirical analysis, we observe the arrivals and departures at the same station at an hourly level for each station. Hence to recognize this, we employ a multilevel linear model that explicitly recognizes the dependencies associated with the bicycle flow variable originating from the same BIXI station. Specifically, 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. The linear mixed model collapses to a simple linear regression model in the absence of any station specific effects. A brief description of the linear mixed model is provided below.

Let $q = 1, 2, \dots, Q$ be an index to represent each station, $d = 1, 2, \dots, D$ be an index to represent the various days on which data was collected and $t = 1, 2, \dots, 20$ 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:

$$y_{qdt} = \beta X + \varepsilon$$

where y_{qdt} is the normalized arrival or departure rate as dependent variable, X is an $L \times 1$ column vector of attributes and the model coefficients, β , is an $L \times 1$ 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. Due to the substantial size of the data and the number of independent variables considered in our study, it is prohibitively burdensome, in terms of run time, to estimate the combined influence of the three components simultaneously. Thus, we consider the station and the time-of-day to be related common unobserved effects. In this structure, the data can be visualized as 20 records for each Station-Day combination for a total of 820 observations. Estimating a full covariance matrix (20×20) is computationally intensive while providing very little intuition. Hence, 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:

$$\Omega = \sigma^2 \begin{pmatrix} 1 & \phi\rho & \phi\rho^2 & \dots & \phi\rho^{19} \\ \phi\rho & 1 & \dots & \dots & \dots \\ \vdots & \vdots & & \vdots & \vdots \\ \phi\rho^{19} & \dots & \dots & \dots & 1 \end{pmatrix}$$

The parameter σ 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) that is slightly different from maximum likelihood

(ML) approach. The REML approach estimates the parameters by computing the likelihood function on a transformed dataset. The approach is commonly used for linear mixed models (Harville, 1977).

5.2. Model fit measures

In our study, two model frameworks were estimated for arrivals and departures: (1) a linear regression model and (2) a linear mixed model. The final model selection was based on the restricted log-likelihood and Bayesian Information Criterion metrics. Our model estimation process was guided by considerations of parsimony and intuitiveness. The two model frameworks were compared using the *log-likelihood ratio* (LLR) test. For the arrivals model, the LLR test statistic was significant at any reasonable level of significance (the LLR test-statistic value was 3632, significantly higher than the corresponding chi-square value for two additional degrees of freedom (ϕ and ρ)). Similarly, for the departures model, the LLR test statistic was significant at any reasonable level of significance (the LLR test-statistic value was 3491). The LLR test comparisons clearly highlight the suitability of the mixed modeling approach employed in our analysis for examining the determinants of BIXI usage in Montreal.

5.3. Results

In this section, we discuss the results of linear mixed model estimation to understand the different effects of meteorological, spatial and temporal elements on the bicycle usage in the BIXI bicycle-sharing system. It must be noted that we considered several specifications but only the statistically significant results for arrival and departure rates are presented in Table 2.

5.3.1. Weather variables

As expected, there is a positive correlation between temperature and the arrival and departure rates. On the other hand, humidity has a negative impact on the arrival and departure rates. People are less likely to ride a bicycle in rainy or very humid time periods. However, the rainy weather variable is not significant for the arrival rate model. This might be explained by the idea that the weather has a stronger effect on the decision of taking out a bicycle than on returning it.

5.3.2. Temporal Variables

People tend to bicycle more on weekdays than weekends, as highlighted by the negative coefficient of the weekend variable. The interpretation of the time-of-day variables needs to be judiciously undertaken due to the presence of interaction effects with population density and university variables. Nevertheless, we clearly observe that the BIXI system is more predominantly used during the PM period relative to other times of the day. The likelihood of using bicycle-sharing systems increases on Friday and Saturday nights, indicating a propensity of young individual users in the downtown core of Montreal during these periods compared to other days.

5.3.3. Bicycle Infrastructure Variables

In this section, the results for parameters related to bicycling infrastructure variables are explained. The bicycle flows and usage of the bicycle-sharing system increase when there are more bicycle facilities (bicycle lanes, bicycle paths, etc.) nearby a BIXI station (in agreement with the findings of Buck and Buehler, 2012). While the length of minor roads in a 250 meter buffer of each station is associated with a positive impact on arrival and departure rates, the

length of major roads has a negative effect. The results indicate that BIXI usage is more likely to occur in densely populated neighborhoods. The impact of the number of BIXI stations and the BIXI capacity in a 250 meter buffer need to be examined as a combination. At first glance, it might seem unintuitive that the impact of capacity is negative on BIXI usage. However, the result recognizes that as the number of stations increases we simultaneously increase the capacity. Hence, the estimates obtained are the overall effect of adding stations as well as capacity. In fact, the capacity variable is almost 25 times smaller than the positive impact associated with the number of BIXI stations, highlighting that adding more stations with capacity of 10-15 (the typical size in Montreal) is likely to increase BIXI usage more than adding a few large stations. The result provides an indication that adding stations with very large capacity is not as productive for arrivals and departures as adding smaller stations.

5.3.4. Land Use and Built Environment Variables

It is expected that the arrival and departure rates decrease when a BIXI station is located farther from the CBD. This is supported by the negative coefficient of the distance-to-the-CBD variable. BIXI users often combine their trip mode with the metro more than other modes of transport (Bachand-Marleau et al., 2012); this is also recognized by the positive impact of the presence of metro stations near BIXI stations in the results (similar results can be seen in Nair et al., 2013). In general, the number of restaurants in the vicinity of a BIXI station increases the usage of that station (similar to the findings of Wang et al., 2012, Hampshire et al., 2013). While the presence of this type of business has a negative impact on the departure rate of a BIXI station in the AM period, it intuitively has a positive influence in both arrival and departure rates in the PM period, reinforcing the attraction of bicycle-sharing systems for restaurant customers. The number of all other commercial enterprises in the 250 meter buffer of each station during PM and evening time periods is associated with negative impact. The coefficient associated with the presence of a university campus on a BIXI station's arrival rate has, interestingly, the opposite sign in the AM and PM periods. BIXI stations near universities are more likely to experience a higher volume of bicycles arriving in the AM than in the PM. While for the departure rates model, the negative coefficient for the AM period has the similar explanation, the university variable does not have a significant influence in PM period. This is plausible since students and teachers tend to have more flexible schedules and usually do not have a fixed time for the end of a work day. The effect of population and job density are incorporated in both models at the TAZ level. BIXI stations in TAZs with higher population density tend to have higher arrival and departure rates (see Rixey, 2013; Wang et al, 2012, similar results). The opposite sign of job density in the AM and PM in the arrival rate models highlights the likely use of bicycle-sharing systems for daily work commute trips.

5.4. Model Validation

The model estimation results for arrival and departure rates were validated using data from May 2013 (one year after the data used to fit the model). The bicycle availability data was compiled from minute-by-minute readings from the BIXI system for all the stations in May 2013. The same data compilation process described in sample preparation for model estimation (see section 3) was repeated to compute bicycle arrival and departure rates. The model developed in section 5.3 was used to generate predictions of bicycle arrivals and departures, and the predictions were compared with the observed values in the validation dataset. Specifically, we calculated two error metrics to evaluate model prediction performance: a) Root Mean Square Error (RMSE) and

b) Mean Absolute Error (MAE). Furthermore, we computed the absolute error as a percentage of station capacity and examined the number of stations with less than 5% error, between 5 and 10% error, between 10 and 15% error, between 15 and 20% error, between 20 and 25% error, greater than 25% error. These measures were computed for the entire sample as well as for specific time periods of the day. The validation exercise results are presented in Table 3. Overall, the predicted arrival and departure rates are reasonably close to the observed rates with absolute error of around 1.8 bicycles per hour. The results indicate that for about 90% of the records the error in prediction is within 20%. The fit for the arrival model is slightly better than the fit for the departure model. In terms of time of day, we can see that the performance of the model in the PM period is relatively inferior to the performance of the model for other time periods. However the results are satisfactory considering the larger rates of arrival and departure in the PM period. The validation highlight the predictive ability of the proposed framework to examine BIXI system bicycle flows (arrivals and departures).

6. Policy Exercise

To better illustrate the magnitude of effects of variables on the use of BIXI system we computed the elasticity effects for both arrival and departure models by computing the percentage change of arrival/departure rate due to changes to the exogenous variables.

In this part, we focus on the following variables: 1) increasing the length of bicycle facilities by 10% in the 250 meter buffer; 2) increasing the number of stations in the buffer without increasing the capacity in the buffer, i.e., we reallocate capacity to add a new station; 3) increasing the station capacity by the average station size (19); and 4) increasing the number of restaurants by 50% of average number in the 250 meter buffer. The elasticity effects are computed as a percentage difference in arrivals and departures relative to the base case. The measures generated are presented in Table 4.

The following observations can be made from the results presented. First, an increase in the bicycle infrastructure variables (length of bicycle facilities, stations and/or capacity) leads to an increase in usage of the bicycle-sharing system, as expected, since the presence of infrastructure plays a great role in cyclists' decision to use such a system. These effects are marginally higher for departures than for arrivals. Second, and more strikingly, we see that increasing the number of stations without increasing capacity in the buffer has a greater impact than increasing capacity by as much as an average station. We believe that this result is quite useful for future BIXI system planning purposes and for other bicycle-sharing operators. The result clearly underscores the need to reallocate very large stations as smaller stations with lower capacity in multiple locations to increase BIXI system usage. Third, it is interesting to see that the bicycle infrastructure variables in the buffer have the most significant impact on arrival rates in the AM period and on departure rates in the Night period, while the effect of station capacity has a similar trend during the day for both arrivals and departures. Finally, we see that increasing the number of restaurants results in an increase of bicycle usage especially in the PM and Evening periods. This finding can be helpful when trying to understand the best areas to allocate new stations to ensure high usage.

7. Conclusion

This study examined the factors influencing the usage flows of a bicycle-sharing system in Montreal, Canada. It contributes to the literature by capturing the effect of meteorological data, temporal characteristics, bicycle infrastructure, land use, and built environment attributes on bicycle arrival and departure flows at the station level, using data obtained from the BIXI system. The multilevel model estimation approach provides intuitive results for both arrival and departure rates. It is observed that people are more likely to use a bicycle-sharing system under good weather conditions. While during the weekends the bicycle usage reduces, Friday and Saturday nights are positively related to arrival and departure rates. The bicycle flows are expected to decrease when we go farther from CBD. The accessibility measures are plausibly correlated to bicycle usage for every station. Restaurants, other commercial enterprises, and universities in the vicinity of a station significantly influence the arrival and departure rates of the BIXI station. The BIXI system variables, number of stations and capacity, have an intricate relationship with arrivals and departures. Specifically, we observe that adding a BIXI station has a predominantly stronger impact on bicycle flows compared to increasing station capacity. Population density of a station's TAZ positively affects the bicycle flows while the effect of the TAZ job density variable has an opposite sign in the AM and the PM.

The model estimation results for arrival and departure rates were validated using the data from May 2013. Overall, the predicted arrival and departure rates are reasonably close to the observed rates with absolute error of around 1.8 bicycles per hour. Further, in almost 90% of the validation records, the error in prediction was less than 20% of the capacity of that station. Further, to examine the impact of exogenous variables on BIXI arrivals and departures, the impact of changes to exogenous variables is considered. The results provide interesting insights. The most prominent result from the exercise highlights the intricate relationship between the number of stations and bicycle capacity. The relationship suggests that adding additional stations (either by relocating existing capacity from large stations or adding new bicycle slots) is more beneficial in terms of arrival and departure flows compared to adding capacity to existing stations. The finding is very important for decision makers planning to install new BIXI stations in Montreal or for decision makers planning new bicycle sharing schemes in other cities.

ACKNOWLEDGEMENTS

The authors would like to acknowledge financial support from Natural Sciences and Engineering Research Council (NSERC) of Canada under the Discovery Grants program. The authors would also like to acknowledge the critical input of two anonymous reviewers and editors of the special issue.

References

- Bachand-Marleau, J., Larsen, J., El-Geneidy, A., 2011. Much-Anticipated Marriage of Cycling and Transit How Will It Work? *Transportation Research Record* 2247, 109-117.
- Bachand-Marleau, J., Lee, B., El-Geneidy, A., 2012. Better Understanding of Factors Influencing Likelihood of Using Shared Bicycle Systems and Frequency of Use. *Transportation Research Record* 2314, 66-71.
- Borgnat, P., Abry, P., Flandrin, P., Robardet, C., Rouquier, J.B., Fleury, E., 2011. Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective. *Advances in Complex Systems* 14, 415-438.

- 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.
- Carr, L.J., Dunsiger, S.I., Marcus, B.H., 2011. Validation of walk score for estimating access to walkable amenities. *British Journal of Sports Medicine* 45 (14), 1144-1148.
- Daddio, D., 2012. Maximizing Bicycle Sharing: An Empirical Analysis of Capital Bikeshare Usage. University of North Carolina at Chapel Hill.
- DeMaio, P., 2009. Bike-sharing: history, impacts, models of provision, and future. *Journal of Public Transportation* 12, 41-56.
- Fishman, E., Washington, S., Haworth, N., 2013. Bike Share: A Synthesis of the Literature. *Transport Review* 33, 148-165.
- Froehlich, J., Neumann, J., Oliver, N., 2009. Sensing and Predicting the Pulse of the City through Shared Bicycling. 21st International Joint Conference on Artificial Intelligence (Ijcai-09), Proceedings, 1420-1426.
- Fuller, D., Gauvin, L., Kestens, Y., Daniel, M., Fournier, M., Morency, P., Drouin, L., 2011. Use of a New Public Bicycle Share Program in Montreal, Canada. *American Journal of Preventive Medicine* 41, 80-83.
- Gregerson, J., Hepp-Buchanan, M., Rowe, D., Vander Sluis, J., Wygonik, E., Xenakis, M., McCormack, E., 2010. Seattle Bicycle Share Feasibility Study. Seattle: University of Washington.
- Hampshire, R., M. Lavanya and N. Eluru, 2013. "An Empirical Analysis of Bike Sharing Usage and Rebalancing: Explaining Trip Generation and Attraction from Revealed Preference Data" Technical Paper, Heinz College, Carnegie Mellon University.
- Harville, D.A., 1977. Maximum Likelihood Approaches to Variance Component Estimation and to Related Problems. *Journal of the American Statistical Association* 72, 320-338.
- Jensen, P., Rouquier, J.B., Ovtracht, N., Robardet, C., 2010. Characterizing the speed and paths of shared bicycle use in Lyon. *Transportation Research Part D: Transport and Environment* 15, 522-524.
- Kaltenbrunner, A., Meza, R., Grivolla, J., Codina, J., Banchs, R., 2010. Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system. *Pervasive and Mobile Computing* 6, 455-466.
- Krykewycz, G., Puchalsky, C., Rocks, J., Bonnette, B., Jaskiewicz, F., 2010. Defining a Primary Market and Estimating Demand for Major Bicycle-Sharing Program in Philadelphia, Pennsylvania. *Transportation Research Record*, 117-124.
- Nair, R., Miller-Hooks, E., Hampshire, R., Busic, A., 2013. Large-Scale Vehicle Sharing Systems: Analysis of Velib. *International Journal of Sustainable Transportation* 7, 85-106.
- PBSC 2013, PBSC Urban Solutions. <http://www.publicbikesystem.com/what-we-achived/case-studies-info/?id=1>.
- Rietveld, P., Daniel, V., 2004. Determinants of bicycle use: do municipal policies matter? *Transportation Research Part A: Policy and Practice* 38, 531-550.

Rixey, R., 2013. Station-Level Forecasting of Bike Sharing Ridership: Station Network Effects in Three U.S. Systems. Paper presented at the 92nd Transportation Research Board Annual Meeting 2013, Washington, DC.

Shaheen, S., Guzman, S., Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia Past, Present, and Future. *Transportation Research Record* 2143, 159-167.

van Lierop, D., Grimsrud, M., and El-Geneidy, A., 2013. Breaking into bicycle theft: Insights from Montreal, Canada. *Forthcoming International Journal of Sustainable Transportation*.

Vogel, P., Mattfeld, D., 2011. Strategic and Operational Planning of Bike-Sharing Systems by Data Mining - A Case Study. *Lecture Notes in Computer Science* 6971, 127-141.

Wang, X., Lindsey, G., Schoner, J., Harrison, A., 2012. Modeling bike share station activity: the effects of nearby businesses and jobs on trips to and from stations. Paper presented at the 92nd Transportation Research Board Annual Meeting 2012, Washington, DC.

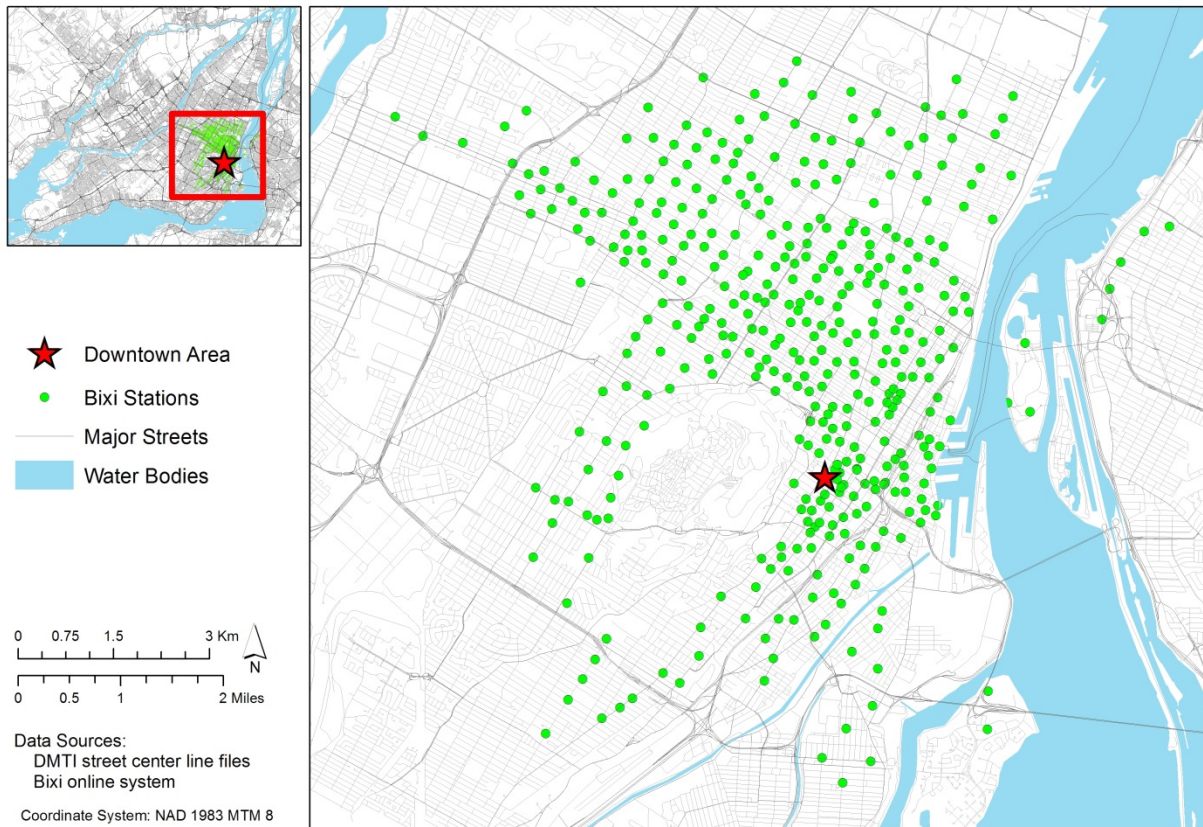


Figure 1 BIXI stations in Montreal Island

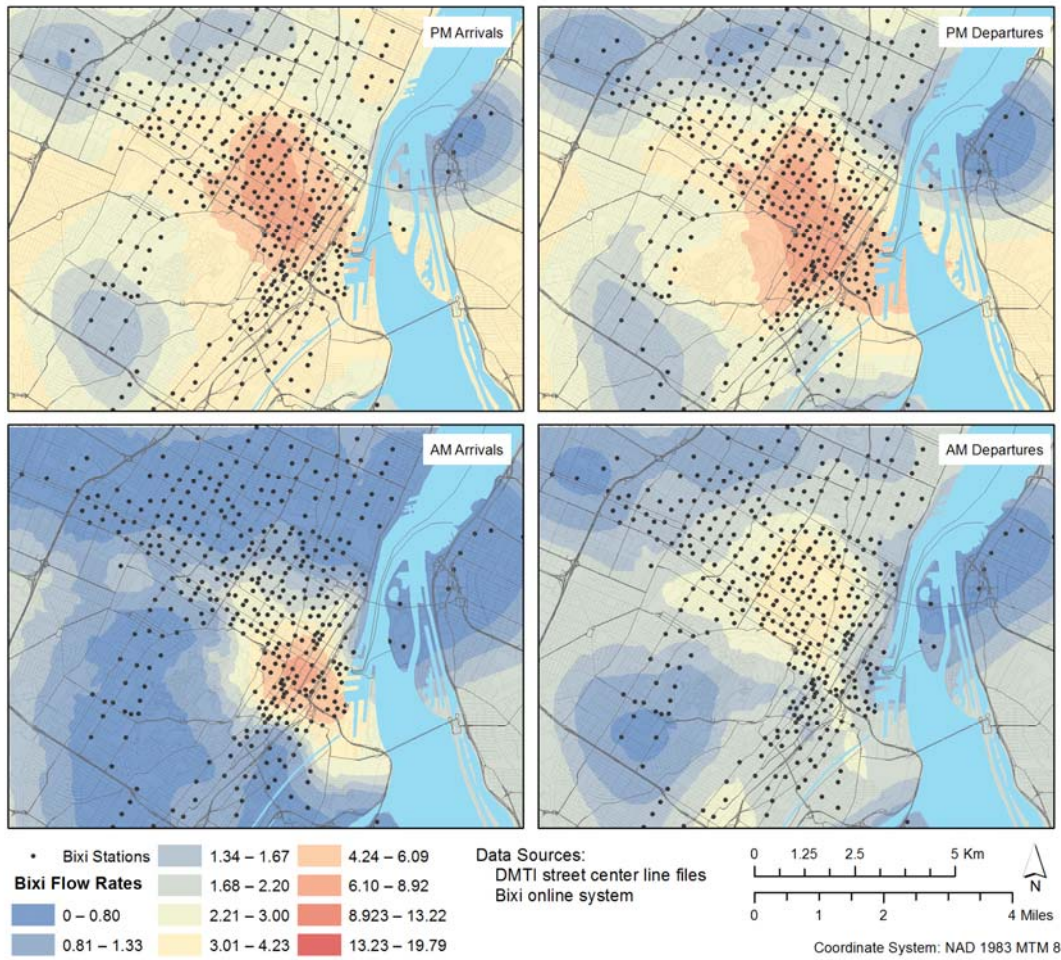


Figure 2 Spatial Distribution of Average Arrival and Departure Rates in Peak hours

Table 1 Descriptive Summary of sample characteristics

Continuous Variables	Min	Max	Mean	Std. Deviation
Temperature (°C)	5.9	33	20.90	5.19
Relative Humidity (%)	24	99	61.40	16.70
Elevation (m)	14.3	154.8	49.22	24.33
Station Distance to CBD (km)	0.11	9.26	3.45	1.91
Length of Bicycle Facility in 250m Buffer (km)	0	2.49	0.70	0.51
Length of Minor Roads in 250m Buffer (km)	1.14	6.48	3.56	0.83
Length of Major Roads in 250m Buffer (km)	0	5.73	1.14	1.02
Length of Bus Lines in 250m Buffer (km)	0	12.33	2.81	1.94
Area of Parks in 250m Buffer (m ²)	0	194907	14551	26962
Number of Restaurants in 250m Buffer	0	194	24.00	35.31
Number of other Commercial Enterprises in 250m Buffer	0	1989	121.59	206.85
Walkscore	14	97	62.3	15.7
Number of BIXI stations in 250m Buffer	1	8	2.23	1.46
Capacity of BIXI stations in 250m Buffer	7	223	46.89	40.49
Station Capacity	7	65	19.53	7.95
TAZ Pop Density (people per m ² × 1000)	1.01	187.79	59.38	31.62
TAZ Job Density (jobs per m ² × 1000)	0.07	4078.13	141.19	528.96
Categorical Variables	Percentage			
Rainy Weather	9.7			
Weekends	26.5			
Friday & Saturday Nights	8.0			
Metro Station in 250m Buffer	21.7			
Station in Downtown area	17.1			
Station in Oldport area	4.9			
University in 250m buffer	17.1			
School in 250m buffer	40.7			

Table 2 Model Estimation Results

Parameter	Arrival Rate		Departure Rate	
	Coefficient	t-statistic	Coefficient	t-statistic
Intercept	0.0784	3.066	0.0584	2.271
Weather Variables				
Temperature	0.0048	8.829	0.0047	8.576
Relative Humidity	-0.0013	-8.556	-0.0012	-7.765
Rainy Weather	-0.0035	-0.697	-0.0124	-2.457
Time Variables				
Weekend	-0.0451	-7.031	-0.0506	-7.838
AM	-0.0259	-5.982	0.0548	11.768
Midday	-0.0186	-4.078	0.0065	1.418
PM	0.0734	15.042	0.0526	10.824
Friday & Saturday Nights	0.0608	10.218	0.0735	12.215
Bicycle Infrastructure Variables				
Length of Bicycle Facility in 250m Buffer	0.0342	5.911	0.0361	6.200
Length of Minor Roads in 250m Buffer	0.0110	2.645	0.0112	2.668
Length of Major Roads in 250m Buffer	-0.0173	-5.224	-0.0189	-5.659
Number of BIXI stations in 250m Buffer	0.0254	4.923	0.0241	4.662
Capacity of BIXI stations in 250m Buffer	-0.0011	-5.581	-0.0010	-5.206
Land use and Built Environment Variables				
Station Distance to CBD	-0.0101	-4.974	-0.0110	-5.408
Metro Station in 250m Buffer	0.0202	2.762	0.0181	2.465
Number of Restaurants in 250m Buffer*PM	0.0005	3.459	0.0006	5.844
Number of Restaurants in 250m Buffer	0.0004	3.691	0.0005	4.276
Number of Restaurants in 250m Buffer*AM	--	--	-0.0007	-6.504
Number of other Commercial Enterprises in 250m Buffer*PM	-0.0001	-4.343	--	--
Number of other Commercial Enterprises in 250m Buffer*Evening	-0.0001	-5.246	-0.0001	-3.201
University in 250m buffer * AM	0.0228	2.780	-0.0352	-4.052
University in 250m buffer * PM	-0.0367	-4.253	-	-
TAZ Pop Density	0.1603	1.804	0.1613	1.805
TAZ Job Density * AM	0.0607	10.354	0.0142	2.036
TAZ Job Density * PM	-0.0230	-3.338	0.0197	2.875
ARMA Correlation Parameters				
σ	0.0256	66.613	0.0262	67.282
ρ	0.8928	114.741	0.8942	105.994
ϕ	0.3546	35.216	0.3459	33.982

Table 3 Validation Results

		Overall	Night	AM	Midday	PM	Evening
Arrival	Mean Absolute Error	1.840	1.649	1.971	1.828	2.426	1.696
	Root Mean Square Error	2.843	2.441	3.554	2.664	3.321	2.359
	Percentage of results with Absolute Error less than 5% Station Capacity	37.6	39.6	48.9	37.1	25.3	39.0
	10% Station Capacity	65.2	71.9	75.2	65.8	49.1	67.5
	15% Station Capacity	82.0	88.4	86.5	83.9	68.8	85.2
	20% Station Capacity	90.4	94.3	91.0	92.1	82.2	93.6
	25% Station Capacity	94.1	96.1	93.4	95.0	90	96.5
Departure	Mean Absolute Error	1.888	1.631	1.937	1.873	2.555	1.720
	Root Mean Square Error	2.884	2.529	2.981	2.680	3.782	2.452
	Percentage of results with Absolute Error less than 5% Station Capacity	36.4	45.4	38.2	35.6	27.3	41.0
	10% Station Capacity	64.0	75.6	66.6	63.8	51.2	69.2
	15% Station Capacity	81.2	89.1	82.9	82.3	69.7	85.7
	20% Station Capacity	89.8	93.8	89.7	91.5	82.1	93.4
	25% Station Capacity	93.6	95.7	92.7	95.1	88.8	96.3

Table 4 Elasticity Effects for Arrival and Departure Rates*

Variable change considered		Overall	Night	AM	Midday	PM	Evening
Arrival	Bicycle facility length in 250m buffer increased by 10%	1.78	0.67	3.01	2.06	1.19	1.27
	Number of BIXI stations in 250m buffer increased by 1 (<i>Neighbouring capacity remains same</i>)	19.51	-6.76	26.46	25.08	13.43	23.43
	Station capacity increased by 4	14.17	19.55	9.22	12.90	17.09	15.40
	Number of restaurants increased by 50%	4.68	-1.39	5.42	5.14	6.18	4.80
Departure	Bicycle facility length in 250m buffer increased by 10%	2.09	3.94	2.61	1.83	1.41	1.73
	Number of BIXI stations in 250m buffer increased by 1 (<i>Neighbouring capacity remains same</i>)	26.83	46.74	28.65	22.50	16.06	32.34
	Station capacity increased by 4	13.34	8.11	12.04	14.70	17.20	11.34
	Number of restaurants increased by 50%	5.77	11.87	-3.03	5.72	9.17	8.22

* *The percentage change of arrival/departure rate due to changes to the exogenous variables*