An Analysis of Bike Sharing Usage: Explaining Trip Generation and Attraction from Observed Demand

Robert C. Hampshire, Ph.D
Assistant Professor of Operations Research and Public Policy
H. John Heinz III College
School of Public Policy and Management
School of Information Systems and Management
Carnegie Mellon University
4800 Forbes Ave.
Pittsburgh, PA 15213

Lavanya Marla, Ph.D
Systems Scientist
H. John Heinz III College
School of Public Policy and Management
School of Information Systems and Management
Carnegie Mellon University
4800 Forbes Ave.
Pittsburgh, PA 15213

Word Count: 5204+ 1000 (4 Figures) + 1000(4 tables) = 7204
Date: August 1, 2011

ABSTRACT

Over 100 cities around the world have deployed or have plans to deploy a bike sharing system. Bike-sharing programs enable flexibility to users by providing rentals at a variety of locations, and by facilitating one-way trips. In addition, they positively impact the environment and quality of life.

The main contribution of this paper is explaining the factors effecting bike sharing trip generation and attraction. Using usage data from bike sharing systems in Barcelona and Seville, census level demographic data, and the location of points of interest, we explain various factors effecting bike sharing usage. We employ a panel regression model estimation strategy. By using two different fixed effects models, we are able to produce consistent estimates of trip generation and attraction factors in the presence of unobserved spatial and temporal variables.

We find that the relationship between bike sharing and alternative modes of transportation can be complicated. In some settings bike sharing *competes* with alternative modes of transportation, while one can also argue that in other settings bike sharing complements. Taken together, the findings strongly support the following usage scenario: bike sharing programs in Barcelona and Seville are used mainly for commuting in the morning. In the evening a larger variety of trips purposes drive usage. These evening trips are also shorter and closer to home.

The results provide empirical foundation for cities and planners in understanding the key factors contributing to bike sharing usage.

INTRODUCTION

Bike sharing systems provide the temporary rental of publicly available bicycles. Bike sharing programs have the potential to reduce the number of cars on the roads, hence reducing congestion; they promote healthy living and are environmentally friendly. Bike sharing systems are an emerging mode of transportation. Over 100 cities have operating bike sharing programs worldwide.

However, with over 100 bike sharing programs deployed throughout the world there has yet to be a published empirical evaluation of the factors explaining trip generation or trip attraction. It is the goal of this paper to begin to lay a foundation for empirical analysis of the intensity of usage of bike sharing systems.

Many cities and planners have conducted feasibility studies of existing and proposed bike sharing systems. These studies often include a demand estimation section and methodology. For example, Philadelphia[1], New York[2], London[3], Washington D.C. and many others have published bike sharing feasibility studies in these reports many assumptions and hypothesis are presented for about demand estimation. The feasibility studies make hypotheses about various demographic, land use, economic and infrastructure factors.

The work of Krykewycz et al. (Krykewycz, Rocks, Bonnette, & Jaskiewicz, 2011) is representative of demand estimation methodology for bike sharing systems. They hypothesize about factors which contribute to trip origin and trip attraction factor.

The goal of this paper is to explain the factors influencing trip generation and trip attraction from observed by sharing usage. Using bike sharing usage data from Barcelona and Seville, Spain, census level demographic data, and points of interest data, we test the hypotheses about the various factors effecting bike sharing usage. We selected Barcelona for one cases because of its popularity and large number of trips. Additionally, Barcelona has been previously studied in the literature, see (Froehlich, Neumann and Oliver, Sensing and predicting the pulse of the city through shared bicycling) and (Froehlich, Neumann and Oliver., Measuring the pulse of the city through shared bicycle programs.) We selected Seville as a case study because of the urban redevelopment and innovative transport planning that has take place over the last ten years (The Circle of Life).

In the remainder of the paper, we first review the growing literature on bike sharing. Next, discussion the unique data sources used in the empirical estimation strategy. In Section 3, we present the general panel data model, and discuss the most salient challenges and solutions to estimation. We then provide the specific specification of the panel data models for Barcelona and Seville. Next, we provide some descriptive statistics for Barcelona and Seville before presenting the estimation results. Finally, we discuss the proper interpretation of the results and conclude.

LITERATURE REVIEW

There is an emerging literature on bike sharing systems. Several papers have studied the history of bike sharing systems (Shaheen, Guzman, & Zhang, 2010) and (deMaio, 2009). The paper by (Carballeda, Velasco and Rojo) surveys public bike systems in Spain.

[1] http://www.bicyclecoalition.org/files/PhiladelphiaBikeshareConceptStudy.pdf

[2] www.**nyc**.gov/html/dcp/pdf/transportation/**bike_share**_complete.pdf

[3] *www.tfl.gov.uk/.../**cycle-hire-scheme-feasibility**-full-**report**-nov2008.pdf*

The majority of the quantitative studies focus on state prediction using time series models. The work of Borgnat at al. (Borgnat, Abry, Flandrin, & Rouquier, 2009), (Borgnat, Fleury, Robardet & Scherrer, 2010), (Borgnat, Robardet, Rouquier, Abry, Flandrin, & Fleury, 2010) are representive of time series models for bike sharing. The papers (Kaltenbrunner, Rodrigo, Codina, & Banchs, 2010), (Vogel & Mattfeld, 2010) present time series models of bike sharing. The paper by Jensen et al. (Jensen, Rouquier, Ovtracht, & Robardet, 2010) infers the travel speeds of bikes in Lyon bike sharing program.

Another related stream of literature focuses on the use of data mining methods such as clustering (Froehlich, Neumann, & Oliver., Measuring the pulse of the city through shared bicycle programs., 2008), (Vogel & Mattfeld., 2010), and (Froehlich, Neumann, & Oliver, Sensing and predicting the pulse of the city through shared bicycling, 2009) for short term prediction.

Further literature focuses on operational efficiency of bike sharing systems. The work of Nair, et al. (Nair, Miller-Hooks, Hampshire, & Busic, 2010) and (Lin and Yang) characterizes the spatial-temporal supply and demand asymmetries inherent in bike sharing systems. While the paper by (Raviv, Tzur and Forma) seeks to address these asymmetries by optimizing bike repositioning operations.

Trip Generation and Attraction Factors

Many cities have released bike sharing feasibility and demand forecasting studies which use similar methodologies, (see New York City[4], London[5], Philadelphia[6]). Most of these studies posit three main user groups for bike sharing: commuters, students and tourists. The feasibility studies typically use stated preference surveys and census tract level data to estimate uptake rates for each user group. Additionally, the surveys attempt to ask questions which lead to estimates of trip substitution and mode share changes. Our approach is different in that we are estimating trip generation and attraction factors from revealed preference data.

The work of Krykewycz, et al. 2011 is the most directly related to this paper. It is representative of demand estimation methodology for bike sharing systems. Krykewycz et al. (Krykewycz, et al. 2011) presents a systematic framework for estimating demand for a bicycle sharing program in Philadelphia, Pennsylvania. In addition, Krykewycz et al. (Krykewycz, et al. 2011) hypothesize about factors which contribute to trip origin and trip attraction factor. Origin factors include population density and group quarter population density. The attraction factors include job density, location of tourist attractions, and proximity to parks and recreation. They also consider network facilities and infrastructure features like rail stations, bike lanes and bus stops.

DESCRIPTION OF DATA SOURCES

Our goal is to explain select factors influencing bike sharing trip generation and attraction in Barcelona and Seville, Spain. The dependent variables are the arrival rate and departure rate of bikes in sub-

[4] www.**nyc**.gov/html/dcp/pdf/transportation/**bike_share**_complete.pdf
[5] *www.tfl.gov.uk/.../**cycle-hire-scheme-feasibility**-full-**report**-nov2008.pdf*
[6] http://www.bicyclecoalition.org/files/PhiladelphiaBikeshareConceptStudy.pdf

city districts of the city. The independent variables include census level data at the sub-city level on demographics, economic, and housing. The final dataset is comprised of points of interest (POI) data such as the location of businesses, metro stations, leisure activities, restaurants, etc

Bike Sharing Usage Data

We have developed an information systems infrastructure to capture bike sharing system state (snapshot) data in real-time via the websites of these programs. The scraped dataset spans from May – September 20, 2009. The state information was capture either every 2 minutes or 5 minutes depending on the city. Usage rate information in this study is derived from the collected state information.

Eurostat Urban Audit

The European Union's statistical agency, Eurostat, and member state national statistical agencies compile data at the intra-city level for a select number of cities in the their Urban Audit (see *ec.europa.eu/**eurostat***[7]). The Eurostat urban audit data provides variables in the categories of: demographic, economic and housing. The data is available the city level and sub-city district level. For this study we use data at the sub-city district (SCD) level. This study uses data collected in the 2006-2007 urban audit. The number of variables at the sub-city district level in 2006-2007 is limited. We were able to extract the following variables at the SCD level: **Demographic:** ResidentPopulation, ProportionFemalesToMale, PopulationDensity, **Housing:** ProportionOnePersonHouseholds, NumberHouseholds, ProportionSocialHousing[8], **Economic:** NetECActivityRateResidents[9], NumberUnemployed.

We transform this data set for our purposes in this study. We consider the ***population density*** defined by the ResidentPopulation of a SCD divided by the land area. We define the ***female population*** as the number of female residents per 1000 male residents. ***One person households*** is the number of one person households per 1000 households in a SCD. ***Social Housing*** is the number in social housing per 1000 households. ***Labor market size*** is the NetECActivityRateResidents in a SCD.

Tele Atlas Points of Interest

We use POI data from Tele Atlas [10] (see www.teleatlas.com), provider of geographic databases. The Tele Atlas data consists of the latitude and longitude of points of interest (POI) in a city. Tele Atlas places in POI into one of 68 categorizes. We reduce the number categories to 4 for the analysis. The 4 are categories: **business, transport, leisure recreation and university.** We define *job density* as the number of business POI's divided by the land area of an SCD. *Transport accessibility* is the number of transport POI's divided by the land area of an SCD.

PANEL DATA MODELS

An investigation of the usage patterns of bike sharing systems (see (Nair, Miller-Hooks and Hampshire)) shows strong time of day usage patterns. Therefore, we introduce time grouping dummy variables into the model which represent 3 time periods throughout the day: *morning*

[7] Accessed on July 1, 2011
[8] Social housing is low cost social housing provided by the municipality or a housing association - rented or bought.
[9] Activity rate is the number in work or seeking work as a proportion of the population of working age.

[10] Accessed on July 1, 2011

5am-10am, *lunch* 11am-3pm, *evening* 4pm-10pm and *night* 11pm-4am. Furthermore, it reasonable to assume that the demographic, points of interest, and network features are each relevant only for certain time periods of each day. Therefore we introduce a limited set of interaction terms between the time period dummies and either the demographic or POI variables. We also exclude the observations during the night time period. The dependent variables of interest are the average arrivals per hour and average departure per hour in SCD i at time t. Because the data set now captures trends in bike usage (arrivals and departures) for each SCD over time.

Panel data model specifications seek to capture the change in the dependent variable along two dimensions – individuals (in our case, SCDs), and over time (in our case, over each hour in the data). Individual and time effects are captured through an estimation equation as follows: The panel estimation equation for our SCD-hour observations is:

$$y_{it} = \beta_{dem,t}x_{dem,t} + \beta_{POI,t}x_{POI,t} + \beta_{network,t}x_{network,t} + c_i + u_t + \varepsilon_{it} \qquad (1.1)$$

where **y** is the arrival (or departure) rate for the SCD-hour; $x_{dem,t}$, $x_{POI,t}$ and $x_{network,t}$ are the demographic, point-of-interest and network variables respectively, and part of the design matrix of explanatory variables; and c_t, u_i, ε_{it} are unobserved effects. The design matrix includes variables that are time invariant, which may not have any strong time-of-day effects.

Random variables c_t and u_i represent the observed variables that are dependent only on time and the SCD respectively. ε_{it} is the random idiosyncratic error term that is a function of both time and SCD. Our choice of models is related to our beliefs and assumptions about the error terms $(c_i, u_t, \varepsilon_{it})$ and their relationship to the observed variables $\mathbf{x}$.There are many scenarios where unobserved variables along the SCD and time dimensions may be correlated with the explanatory variables in the model. For example, first, the SCD level explanatory variables in the model may be correlated with the unobserved bike rebalancing scheme. The rebalancing scheme in an SCD may be related to the number of transit stops in a given SCD. Thus the unobserved rebalancing scheme is correlated with the SCD explanatory variables. Another example is that the unobserved mode share at the SCD level may be correlated with the demographic features of the resident population.

Because we believe that strong correlations can exist among the variables of the model and the unobserved effects, we adopt fixed effects panel data estimators that account for such correlations. The fixed effects estimators are consistent in the presence of unobservable variables that are correlated with the explanatory variables in a general way (see (Wooldridge), section 10.5).

Fixed Effects (between) Model

The fixed effect *between* model estimates unobserved impact of time in the model. This is critical effect to capture in bike sharing systems, as usage patterns vary temporally due to commuting behavior. The interpretation of the fixed effects between model is to explain the relative average impact of the explanatory variables between SCDs in the model. Thus the estimated coefficients using the between estimator explains why the dependent variable is larger or smaller for one SCD relative to the other SCDs in the city during a given hour.

Fixed Effects (within) Model

The fixed effects *within* model estimates a SCD specific intercept. This intercept captures the effects of unobserved variables specific to a given SCD. For example, one such unobserved effect could be the elevation of the SCD which affects usage as riding uphill may not be preferable. The interpretation of the coefficients estimated by the fixed effects *within* is the average impact of that explanatory variable on the dependent variable at time t of an SCD relative to its own time average value.

MODEL SPECIFICATIONS

We illustrate the mathematical validity of our model specifications through one example.
The following is the full model specification for the arrival rate in each SCD-Hour in the city of Barcelona:

$$ArrivalRate_{SCD,t} = \beta_{Hour}+\beta_{DayOfWeek}+\beta_{PopulationDensity,Morning}x_{SCDPopulationDensity,Morning}+ \beta_{PopulationDensity,Evening}x_{SCDPopulationDensity,Evening}+\beta_{NumberUnemployed,Morning}x_{Unemployed,Morning}+ \beta_{EconomicActivity,Morning}x_{EconomicActivity,Morning}+ \beta_{OnePersonHouseholds,Morning}x_{OnePersonHouseholds,Morning}+ \beta_{AltTransportDensity,Morning}x_{AltTransportDensity,Morning}+ \beta_{AltTransportDensity,Evening}x_{AltTransportDensity,Evening}+ \beta_{NrBusinesses,Evening}x_{NrBusinesses,Evening}+ \beta_{Restaurants,Evening}x_{Restaurants,Evening}+c_i+u_t+\varepsilon_{it},$$

where $x_{SCDPopulationDensity,Morning} = Dummy_{Morning}*x_{SCDPopulationDensity}$, etc.

Because several of the variables are created by interaction with dummy variables for time period of day (morning, evening, lunch), we ensure that multicollinearity is avoided by not including the interacted variables for all three times of day. Moreover, we also ensure that variables that are highly correlated are not all included in the model, as they might be proxies for each other and confound the standard error estimates.

RESULTS

In this section, we report and analyze the estimation results for usage rates in Barcelona and Seville. The within and between fixed effect estimation results are presented. Our dependent variables are the average arrival rate and departure rate per SCD. For both models, serial autocorrelation in the error terms is present over time. This is confirmed by serial autocorrelation tests for panel data sets (Drukker). Thus, we report all standard errors using a covariance estimator which is robust to serial correlation. In particular, we report standard errors using Arellano-White robust estimators (see (Wooldridge)). All of the results were computed by the statistical software program *R* using the *plm* package[11].

Barcelona

The bike sharing program in Barcelona, Spain named Bicing started operation on March 3, 2008. It is operated by Clear Channel. During our study period, it consisted of 402 fixed location bike stations and approximately 6000 bikes. According to the EU sponsored Optimizing Bike Sharing in Europe (OBIS) (see www.obisproject.org) working group, Barcelona has 86 train and metro stops with a bike station nearby.

[11] http://cran.r-project.org/web/packages/plm/index.html

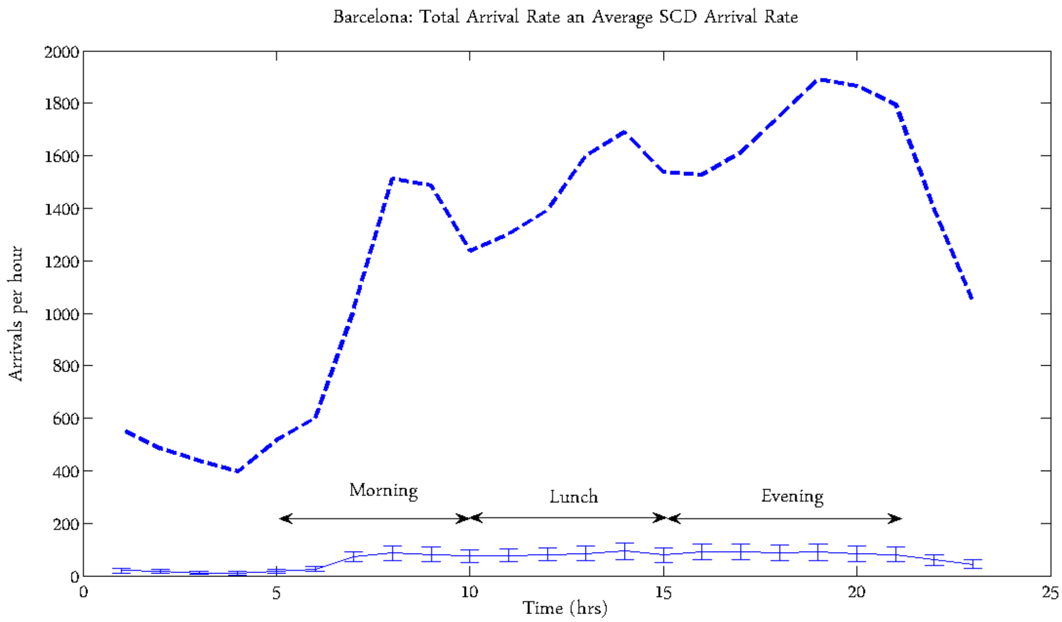


Figure 1 Barcelona Total Arrival Rate (---) and Average and Standard Deviation of the SCD Arrival Rate for 24 hours

Barcelona has 56 sub-city districts in the Eurostat urban audit. The Tele Atlas data contains 6,893 points of interest for Barcelona. During the observation period, there are 402 bike stations in Barcelona. Barcelona has 168 points of interest that are categorized as transport, 2,809 as businesses, 401 as leisure and 60 as universities. Figure 3 is a GIS map of the sub-city district boundaries, location of the bike stations and the points of interest.

We use data from 34 days of observations. From these days, we construct 782 hourly observations of the arrival rate and departure rate averaged at the hourly level. Only weekdays are selected for analysis in this study. In our dataset there are 28,632 average trips per day.

The dotted line in Figure 1 shows the city wide total arrival rate for Barcelona for each hour in a 24 hour period. It shows the morning, lunch and evening behavior of the arrival rate. The total city arrival rate has an upward trend during the morning, lunch and evening periods. The total arrival rate attains a maximum during the evening period around 7pm. Also, the lunch period arrival rate peak is larger than the morning commuting peak. The solid line in Figure 1 displays the hourly arrival rate averaged over all 56 SCDs. In Figure 1 the bars on the solid line represent the standard deviation of the arrival rate amongst the 56 SCDs.

Table 1 Barcelona: Departure Rate

	Between	*Within*
Number of Stations	4.4537e+00 ***	
Population Density (Morning)	4.0006e-01 ***	3.5754e-01 ***
Population Density (Evening)	2.8805e-01 ***	1.7913e-01 ***
NumberUnemployed (Morning)	-1.9172e-03 ***	-2.7882e-03 ***
Labor Market Size (Morning)	5.7916e-04 ***	6.9966e-04 ***
OnePersonHouseholds (Morning)	1.1586e-04 **	-4.6674e-04 ***
Transport Accessibility (Morning)	-6.0730e-02 ***	-2.2330e-01 ***
Transport Accessibility (Evening)	-3.0540e-02	-1.0094e-01**
Job Density (Evening)	-6.2733e-02 ***	-1.0094e-01 ***
Restaurants (Evening)	2.6217e-01 ***	2.0862e-01 ***
N	56	56
T	1141	1141
R^2	64.148	27.244

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 2 Barcelona: Arrival Rate

	Between	*Within*
Number of Stations	4.4951e+00 ***	
Population Density (Morning)	1.6797e-01 ***	
Population Density (Evening)	1.1755e-01 ***	1.0121e-01 ***
FemalePopulation (Evening)	-7.5228e-03 ***	-5.8206e-03 ***
OnePersonHouseholds (Evening)	1.1795e-03 ***	3.3976e-04 ***
Labor Market Size (Evening)	6.1926e-04 ***	5.5880e-04 ***
Transport Accessibility (Morning)	-2.6536e-02	-1.8543e-02
Transport Accessibility (Evening)	1.1185e-01 ***	9.8786e-02 **
Job Density (Morning)	3.5987e-02 ***	
Universities (Morning)	-9.2421e-02	4.3600e-01 ***
Restaurants (Lunch)	1.0345e-01 ***	1.3790e-01 ***
N	56	56
T	1141	1141
R^2	65.077	21.572

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Tables 1 and 2 present the estimates for the fixed effects (within) and fixed effects (between) models for Barcelona, for departure and arrival rate respectively. Our discussion of results considers the both models together. We emphasize the sign and significance of the coefficient effects in the discussion. The point estimates generated by the two models may differ due to model specification. The explanatory variables are to be interpreted as trip attraction factors when the arrival rate is the dependent variable, and trip generation factors when the departure rate is the dependent variable.

We find that the *network* variable representing the number of stations per SCD has a positive and significant impact on both trip generation and attraction, indicating that SCDs with more number of stations attract as well as generate more trips. This is likely because an SCD with a higher number of stations offers more options to the user to borrow or deposit bikes, leading to the SCD being a larger source/destination of bike trips.

The *demographic* variable representing population density is positive and significant in the morning and evening. A positive significant population density in the morning for departures and a positive significant population density in the evening for arrivals support a usage scenario in which departures in the morning are likely to originate near home and arrivals in the evening are likely to terminate near home. Additionally, positive significance for population density in the evening for departures and arrivals also indicates the possibility of other trips (leisure or restaurants) following work trips. Also, SCDs with a higher proportion of female residents seem to be associated with lower trip attraction rates in the evening.

The results for *economic* variables on trip generation support the narrative in which SCDs with smaller labor markets have lower trip generation levels. That is, higher unemployment seems to affect trip generation negatively. This estimate in the *within* model also indicates that if the unemployment in a given SCD were to increase, fewer trips would originate from there. Additionally, the positive coefficient for labor market size on trip generation strongly suggests that bike sharing is used for work commutes; and more trips are generated in SCDs or time periods with larger labor markets. This is supported by the fact that in Barcelona, day passes are not offered for tourists, and most activity results from regular subscribers.

The POI variables provide a rich set of interpretations for trip generation and attraction. The *restaurant* variable has a positive and significant impact on both trip generation and attraction in the evening, indicating trips to and from restaurants in the evening period. This supports the hypothesis of shorter leisure- or food-oriented trips in the evening (as was indicated by the population density parameters). Additionally, *job density* in the evening is negatively correlated with trip generation and positively correlated with trip attraction in the morning. These results together support the usage scenario in which usage is driven by commuting to work in the morning, and dominated by leisure trips in the evening.

For the *transport accessibility* variable we find a significant negative effect in the morning on trip generation. Thus SCDs with more transport alternatives have lower levels of bike sharing departures in the morning relative to other SCDs. This also indicates that if other modes of transport were to increase in an SCD, bike sharing trips will decrease. However, the impact of transport accessibility has an insignificant impact on trip attraction in the morning. This result is consistent with asymmetric morning commuting patterns found in Nair et al. (Nair, Miller-Hooks and Hampshire). During the evening, transport accessibility plays a different role where it has a significant positive impact on trip attraction and a weakly significant negative impact on trip generation.

These results suggest that bike sharing is a *competitor* to other modes of transportation. That is, more accessibility to other modes leads to fewer bike sharing trips. Another phenomenon argued for in the literature is that of *complementarity* – occurring when a user uses multiple modes - another mode of transportation to a bike station, followed by a bike sharing trip. Under this mechanism the presence of alternative modes of transportation are complementary to bike sharing and is positively associated with the intensity of bike sharing usage. In our models, though transport accessibility is negatively associated with trip generation and attraction, we cannot rule out the presence of complementary multi-modal travel behavior. This is because we cannot explicitly observe trip origins and destinations in our dataset.

Seville

The bike sharing program in Seville, named Sevici, and started operation on July 24, 2007. It is operated by JCDecaux. During our study period, it consisted of 271 fixed location bike stations and approximately 2000 bikes. According to the EU sponsored Optimizing Bike Sharing in Europe (OBIS) (see www.obisproject.org) working group, Seville has 19 train and metro stops with a bike station nearby. In contrast to Barcelona, Seville has a flat elevation profile.

We use data from 34 sub-city districts in Seville. The Tele Atlas data contains 1,700 points of interest for Seville. Seville has 168 points of interest that are categorized as transport, 808 business, 278 leisure and 30 universities. Figure 4 presents a GIS map of the sub-city district boundaries, location of the bike stations and the points of interest.

We use data from 24 days of observations. Only weekdays are selected for the analysis in this study. From these days, we construct 220 hourly observations of the arrival rate and departure rate averaged at the hourly level. In the dataset, there is an average of 8,173 trips per day in Seville.

The dotted line in Figure 2 shows the city wide total arrival rate for Seville for each hour in a 24-hour period. It shows the morning, lunch and evening behavior of the arrival rate. One key observation is that the activity in the lunch period is more prominent relevant to the morning periods as compared to Barcelona. Additionally, the peak during the lunch period is later in the day compared to Barcelona. The solid line in Figure 2 displays the hourly arrival rate averaged

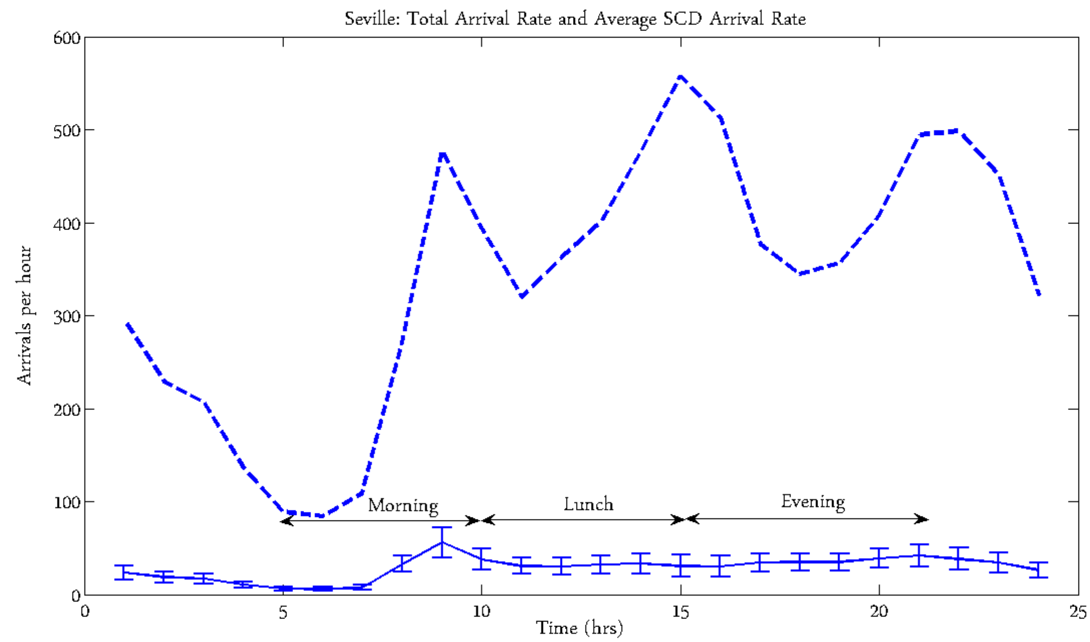


Figure 2 Seville Total Arrival Rate (---) and Average and Standard Deviation of the SCD Arrival Rate for 24 hours

over all 34 SCDs. In Figure 2 the bars on the solid line represents the standard deviation of the arrival rate amongst the 34 SCDs.

Tables 3 and 4 present the results for both the departure and arrival rate for Seville. The estimates for both the fixed effects (within) and fixed effects (between) model are presented.

In Seville (similar to Barcelona), the *network* variable representing the number of stations per SCD has a positive and significant impact on both trip generation and attraction, indicating higher demand in such SCDs. The sign and significance of the *demographic* variable representing population density - positive for trip generation and attraction – indicates strong usage for work trips during the day and for other trips in the evening, as seen in Barcelona. This also supports a usage scenario in which departures in the morning are likely to originate near home and arrivals in the evening are likely to terminate near home.

The parameters of *economic* variables for the *between* model indicate that *labor market size* has a significant positive impact on the departure rate and arrival rate. This strongly suggests that in the morning users commute to work, and depart from work in the evening using bike sharing, as in the case of Barcelona. Also, we find other trip generators and attractors in the evening. The *between* model implies SCDs with larger labor market size both generate and attract more trips, likely those related to work. For the POI variables, we find that *restaurants* at lunch have a positive significant effect on trip generation and attraction. Further, at lunch *job density* has a positive significant impact on trip attraction. These results are consistent with the macro-level trend in Seville (see Figure 2), in which there is relatively more activity (both generation and attraction) during the lunch period than in Barcelona. We hypothesize that this could be driven by either siesta behavior (The Circle of Life) or the presence of tourists in Seville.

The impact of *transport accessibility* in Seville suggests that lunch activity is more influential than in Barcelona.. During the lunch and evening periods in Seville, the between model estimates a significant negative effect on bike sharing trip generation and attraction (indicating competition like in Barcelona); while the within model provides insignificant estimates for the impact of transport accessibility on trip generation in the evening and trip attraction both at lunch and in the evening. This result might be driven by the factor that Seville has a relatively small train network.

The density of *leisure* POI has a positive and significant impact on both trip generation and attraction in the evening. Additionally, *job density* in the evening is positively correlated with trip attraction. Because the job density category includes shopping centers and retail outlets, this suggests strong usage in the evening, and that trip purposes in the evening are diverse. These results together support the scenario in which usage is driven by commuting to work during the day, and by leisure trips in the evening. Also, since we have an increase in both arrivals and departures with the density of leisure POI's in the evening, this implies shorter trips in the evening. This further strongly suggests that the trip purposes are more varied in the evening and of shorter duration.

Table 3 Seville: Departure Rate

	Between	*Within*
Number of Stations	9.9961e-01 ***	
Population Density (Morning)	1.0982e-01 ***	0.080937 **
Population Density (Evening)	2.3963e-02	0.109870 ***
Labor Market Size (Morning)	5.3971e-04 ***	
Restaurants (Lunch)	5.0606e-01 ***	0.214660 *
Job Density (Evening)	5.1331e-02	-0.058912
Leisure (Evening)	1.3318e+00 ***	0.288760 **
Transport Accessibility (Lunch)	-8.7647e-01 ***	-0.617826 ***
Transport Accessibility (Evening)	-6.3242e-01 ***	-0.542751
Restaurants (Evening)	-6.1587e-02	0.145320 *
N	38	38
T	456	456
R^2	22.4	24.59

Table 4 Seville: Arrival Rate

	Between	*Within*
Number of Stations	1.020522 ***	
Population Density (Evening)	0.068846 **	1.8017e-01 ***
Labor Market Size (Evening)	2.5490e-04 ***	5.2372e-01 ***
Restaurants (Lunch)	0.627942 ***	1.7231e-01 ***
Job Density (Lunch)	0.074013 *	1.7231e-01 ***
Job Density (Evening)	0.124780 ***	1.8044e-01 ***
Leisure (Evening)	2.0411e-01 **	1.331195 ***
Transport Accessibility (Lunch)	-0.559277 ***	4.9102e-02
Transport Accessibility (Evening)	-0.592984 ***	-1.1197e-01
Restaurants (Evening)	0.089248 .	3.8342e-01 ***
N	38	38
T	456	456
R^2	20.51	25.31

The results however, have limited utility in *predicting* the level of bike sharing usage in an SCD with a given set of observed explanatory variables. This is due to unobserved variables in the model. Though the estimation technique used is robust to correlation between unobserved variables and the explanatory variables in the model, the omitted variables impact the predictive power of the results.

CONCLUSION

This study presents the first results on explaining the intensity of bike sharing from actual system usage data. The work represents an empirical test of the many hypothesized trip generation and attraction factors made by cities with existing bike sharing system and those cities planning bike sharing programs. We used panel data based models to understand the factors influencing regional and temporal behavior of bike sharing systems in Barcelona and Seville.

We find that number of bike stations, population density and labor market size are strong indicators of trip generation and attraction in both Barcelona and Seville. In Barcelona, indicators such as female population and number of unemployed seemed to explain part of the usage, whereas usage seemed more uniform across socio-demographic factors in Seville. In Seville trips focused around lunch activities seemed significant and contributed to a larger portion of usage than in Barcelona. However, with the current data set, we cannot make direct comparisons based on the relative values of estimated parameters between Barcelona and Seville. This is because we cannot say if the coefficients in the Barcelona model are significantly different from the coefficients in the Seville model. This would require a different estimation strategy and hypothesis testing not carried out in this study.

The usage patterns observed are consistent with commuting patterns during the day which begin at home in the morning and end at home in the evening; with shorter but more diverse leisure- or food-oriented trips that begin and end at home in the evening (following work). In both cities, alternate modes of transportation are seen to have competitive impacts on bike usage, though this does not rule out inter-modal complementarity with bikes.

References

Borgnat, Pierre, et al. "Shared Bicycles in a City: A Signal Processing and Data Analysis Perspective." 2010. Scientific Commons. 1 Febuary 2010 <http://www.scientificcommons.org/58104633>.

—. "Spatial analysis of dynamic movements of Vélo'v, Lyon's shared bicycle program." 2010. Scientific Commons. 1 Febuary 2010 <http://www.scientificcommons.org/57719278>.

—. "Studying lyon's vlo'v: A statistical cyclic model." European Conference on Complex Systems 2009 - ECCS'09 . 2009.

Carballeda, Lisístrata, et al. "PUBLIC BIKE SYSTEMS IN SPAIN." Association for European Transport and contributors (2010).

deMaio, Paul. "Bike-sharing: History, impacts, models of provision, and future." Journal of Public Transportation 12 (2009): 41-56.

Drukker, D. M. "Testing for serial correlation in linear panel-data models." Stata Journal 3 (2003): 168-177.

Froehlich, J., J. Neumann and N. Oliver. "Sensing and predicting the pulse of the city through shared bicycling." Twenty-First International Joint Conference on Artificial Intelligence (IJCAI-09). 2009.

Froehlich, J., J. Neumann and N. Oliver. "Measuring the pulse of the city through shared bicycle programs." International Workshop on Urban, Community, and Social Applications of Networked Sensing Systems- UrbanSense08. 2008.

Jensen, Pablo, et al. "Characterizing the speed and paths of shared bicycles in Lyon." Transportation Research Part D: Transport and Environment 15.8 (2010).

Kaltenbrunner, Andreas, et al. "Urban cycles and mobility patterns: Exploring and predicting trends in a bicycle-based public transport system." Pervasive and Mobile Computing 6. 2010.

Krykewycz, Gregory R., et al. "Defining a primary market and estimating demand for major bicycle-sharing program in Philadelphia, Pennsylvania." Transportation Research Record: Journal of the Transportation Research Board 2143 (2011): 117-124.

Lin, Jenn-Rong and Ta-Hui Yang. "Strategic design of public bicycle sharing systems with service level constraints." Transportation Research Part E 47 (2011): 284-294.

Nair, Rahul, et al. "Large-scale bicycle sharing systems: Analysis of velib." International Journal of Sustainable Transportation (2011).

Raviv, Tal, Michal Tzur and Iris A. Forma. "Static repositioning in a bike-sharing system: Models and solution approaches." 2010.

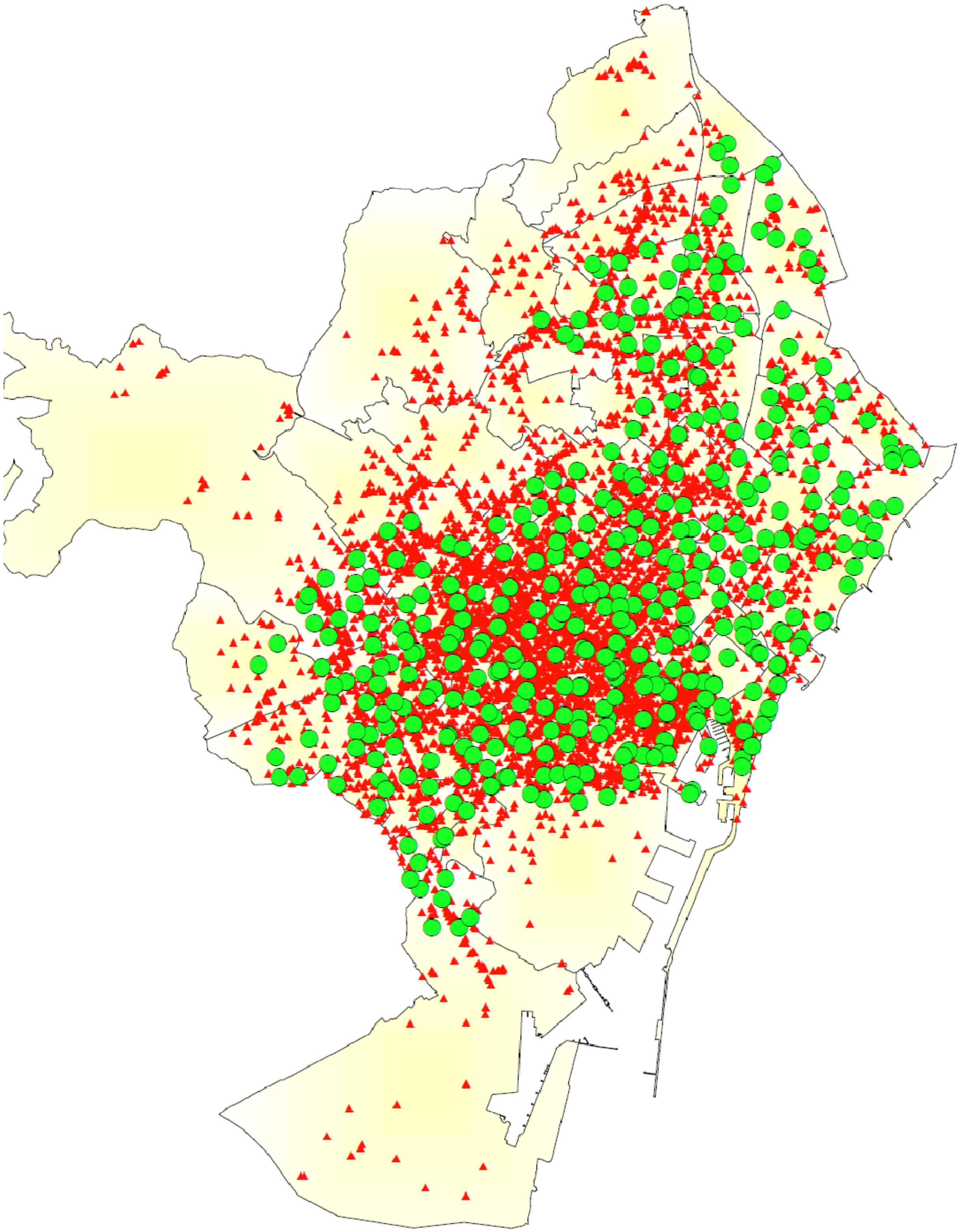
Roberto Diez Pisonero. "Actuation and Promotion Mechanisms of Urban Tourism: The Case of Seville (Spain)." TURIZAM 15.Issue 1 (2011): 26-39 .

Shaheen, Susan A., Stacey Guzman and Hua Zhang. "Bikesharing in europe, the americas, and asia past,present, and future." Journal Transportation Research Record: Journal of the Transportation Research Board 2143 (2010): 159-167.

"The Circle of Life." Cycling Mobility 1 2011: 72-79.

Vogel, Patrick and l Dirk Christian Mattfeld. "Sensing and predicting the pulse of the city through shared bicycling." 12th World Conference on Transport Research (WCTR). 2010.

Wooldridge, Jefferey M. Econometric Analysis of Cross Section and Panel Data. Cambridge: The MIT Press, 2002.



1

2 **Figure 3 Barcelona: Bike Station Locations (circles) and Points of Interest (triangles)**

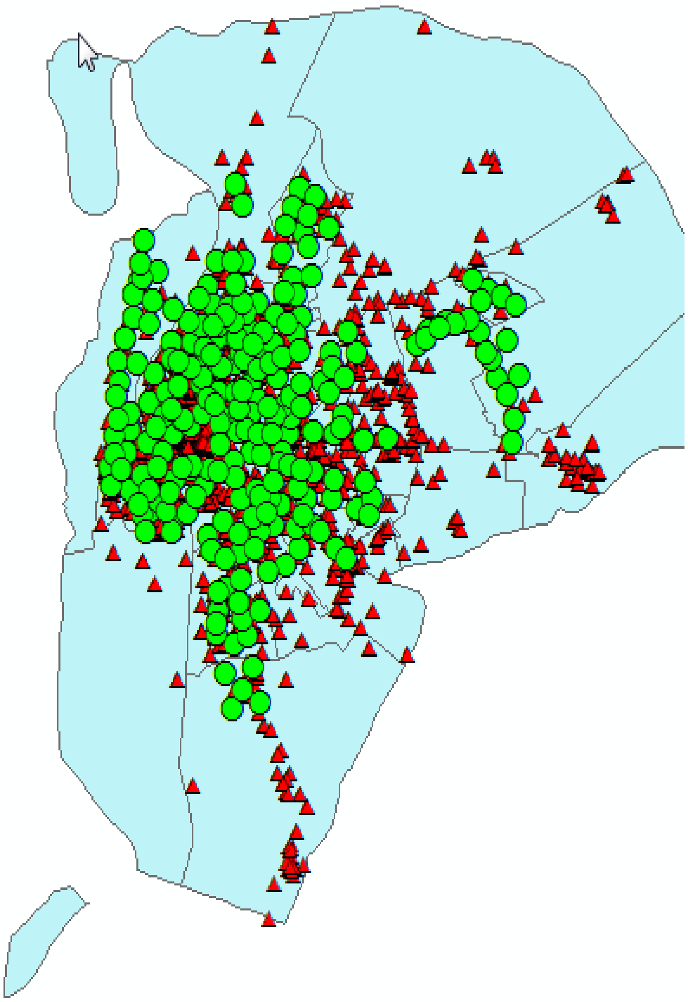


Figure 4 Seville: Bike Station Locations (circles) and Points of Interest (triangles)