Transportation Research Part A 97 (2017) 177–191

Contents lists available at ScienceDirect

Transportation Research Part A

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

An empirical analysis of bike sharing usage and rebalancing: Evidence from Barcelona and Seville

Ahmadreza Faghih-Imani [a], Robert Hampshire [b], Lavanya Marla [c,*], Naveen Eluru [d]

[a] Department of Civil Engineering and Applied Mechanics, McGill University, Canada
[b] Transportation Research Institute, University of Michigan, United States
[c] Department of Industrial and Systems Engineering, University of Illinois at Urbana-Champaign, United States
[d] Department of Civil, Environmental and Construction Engineering, University of Central Florida, United States

ARTICLE INFO

Article history:
Received 4 September 2013
Received in revised form 18 October 2016
Accepted 19 December 2016

Keywords:
Bike sharing
Rebalancing
Linear mixed model
Points of interest

ABSTRACT

Over 400 cities around the world have deployed or have plans to deploy a bike sharing system. However, the factors that drive their usage and the amount of rebalancing they require are not known precisely. A knowledge of these factors would allow cities to design or modify their systems to increase usage while lowering rebalancing costs. We collect station-level occupancy data from two cities and transform station occupancy snapshot data into station level customer arrivals and departures to perform our analysis. Specifically, we postulate that arrivals and departures from stations can be separated into: (i) arrivals (and departures) due to consumers, and (ii) arrivals (and departures) due to the system operators for rebalancing the system. We then develop a mixed linear model to estimate the influence of bicycle infrastructure, socio-demographic characteristics and land-use characteristics on customer arrivals and departures. Further, we develop a binary logit model to identify rebalancing time periods and a regression model framework to estimate the amount of rebalancing. The research is conducted using bike sharing data from Barcelona and Seville, Spain. The resulting modeling framework provides a template for examining bicycle rebalancing in different contexts, and a tool to improve system management of bicycle sharing systems.

1. Introduction

Bike sharing systems are an emerging mode of transportation that provide the temporary rental of publicly available bicycles. These programs have the potential to reduce car usage in dense neighborhoods, hence reducing congestion; additionally they promote healthy living and are environmentally friendly. Over 400 cities have operating bike sharing programs worldwide, including North and South America, Europe and Asia (see http://bike-sharing.blogspot.com). In the process, many cities and planners have conducted feasibility studies of existing and proposed bike sharing systems. These studies often include demand estimation and corresponding methodology. For example, Philadelphia (DVRP, 2011), New York (NYC, 2011), London (TFL, 2011), and many others have published bike sharing feasibility studies. In these reports, many assumptions and hypotheses are presented for demand estimation. The feasibility studies make hypotheses about various socio-demographic, land use, economic and infrastructure factors. The work of Krykewycz et al. (2011) is representative of demand

* Corresponding author.
E-mail addresses: seyed.faghihimani@mail.mcgill.ca (A. Faghih-Imani), hamp@umich.edu (R. Hampshire), lavanyam@illinois.edu (L. Marla), naveen.eluru@ucf.edu (N. Eluru).

http://dx.doi.org/10.1016/j.tra.2016.12.007
0965-8564/

estimation methodology for bike sharing systems. They hypothesize about factors which contribute to trip generation and trip attraction. These factors are grouped into three main categories: socio-demographic, land-use and infrastructure spatial attributes. Our approach also uses these three categories while employing system usage data.

The goal of this paper is to explain the factors influencing trip generation and trip attraction from station occupancy snapshot data. Earlier studies have shown that snapshot data can provide a useful representation of bike share usage (de Chardon and Caruso, 2015). We transform station occupancy snapshot data into two categories: (i) customer arrivals and departures, and (ii) system operator rebalancing (removal and refill), in order to perform this analysis. Augmenting the system data from Barcelona and Seville in Spain with census level socio-demographic data, and points of interest data, we test the hypotheses about the various factors affecting bike sharing usage. We select Barcelona as the primary city for our study because of its popularity and large number of bike trips. Additionally, Barcelona has been previously studied in the literature, see Froehlich et al. (2008, 2009). The goal of these two papers is to predict the availability of bikes at each station. In contrast, our goal is not prediction, but explaining the factors that contribute to trip generation and attraction. We also select Seville as a case study because of the urban redevelopment and innovative transport planning that has taken place over the last ten years (Cycling Mobility, 2011).

The proposed research effort evaluates two hypotheses and their corresponding questions. First, we hypothesize that bike-sharing demand is influenced by bicycle infrastructure (bicycle station numbers and capacity), land use (population density, employment density and points of interest) and temporal variables (such as temperature and humidity). We answer the question of quantifying the impact of these various factors on bike-sharing demand, in a multivariate setting. Second, we hypothesize that rebalancing requirements at each station can be partitioned into the quantity of rebalancing at a station and the frequency of rebalancing. While these metrics are likely to be substantially affected by the same set of variables that influence bike-sharing demand, we expect the influence of these variables to be different for rebalancing. We answer the question of how the influences of these factors differ for rebalancing, as compared to demand, in magnitude and sign. For instance, while mixed land use areas increase bicycle sharing demand they might also reduce demand for rebalancing as bicycle flows occur across all parts of a dense neighborhood.

The outline of this paper is as follows. In Section 2, we review relevant literature. In Section 3, we discuss the unique data sources used in the empirical estimation strategy. We present the linear mixed model and discuss the most salient challenges and solutions to its estimation in Section 4. In Section 5, we discuss the estimation results for Barcelona and Seville, including the explanations for user trips, rebalancing operations caused by imbalance of bikes in the system, as well as policy implications for system design. We conclude in Section 6.

1.1. Contributions

The major contributions of this paper are as follows. First, we generate a unique dataset through a combination of three data sets – (1) station occupancy snapshots collected from the operators' websites; (2) socio-demographic, economic and housing data from census data (for this paper, from Eurostat); and (3) Points of Interest data that describe land-use from TeleAtlas. Second, we separate user arrivals and departures from operator rebalancing (removal or refill of bikes) using a heuristic approach. Through empirical analysis we test assumptions about the factors that influence customer arrivals and departures and rebalancing refill and removal. We apply a methodology for analyzing such systems using behavioural models, specifically, linear mixed models. Finally, we present the first empirical analysis of system rebalancing by the operator focused on understanding the factors creating such imbalances, using an approach consisting of a binary logit model (for identifying stations that need rebalancing) and a linear regression model for the amount of rebalancing. This analysis can help in creating plans for rebalancing well in advance, as well as in creating incentive mechanisms for customers to rebalance bikes.

2. Literature review

Demand estimation for non-motorized travel modes have been fairly well-studied (see Rietveld et al., 2001; Cao et al., 2006; Chatman, 2005; Handy et al., 2006; Kitamura et al., 1997; Schwanen and Mokhtarian, 2005). The FHWA report (FHWA, 1999) describes key factors for trip generation for non-motorized modes. Other studies have attempted the demand analysis of bicycle usage and trip rates (Chatman, 2005), usage and mode choice (Baltes, 1996; Beck and Immers, 1996; Cervero and Duncan, 2003; Hunt and Abraham, 2007) and usage and travel mileage (Ewing et al., 2005). There are relatively fewer studies about bicycle ownership and its relation to bicycle use. Also, measures of bicycle usage are subject to inaccuracy in travel surveys and are therefore often poorly documented (BTS, 2000).

Bike sharing systems, however, are expected to have some similar and some differing characteristics compared to other non-motorized modes. For example, a characteristic differentiating bike sharing systems from other non-motorized systems is that they do not necessitate ownership of bikes and therefore facilitate increased complementarity between biking and transit. A characteristic common to bike sharing and other non-motorized systems is the age group population that is expected to favor their use.

There is emerging literature on bike sharing systems. Studies such as Shaheen et al. (2010) and deMaio (2009) have described the history of bike sharing systems while Carballeda et al. (2010) survey public bike systems in Spain. The majority