Journal of Transport Geography 66 (2018) 80–90

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Community mobility MAUP-ing: A socio-spatial investigation of bikeshare demand in Chicago

Alec Biehl, Alireza Ermagun, Amanda Stathopoulos*

Department of Civil & Environmental Engineering, McCormick School of Engineering & Applied Science, Northwestern University, United States

ARTICLE INFO

Keywords:
Modifiable Areal Unit Problem
Community analysis
Chicago Divvy bikeshare
Travel demand
Model diagnostics

ABSTRACT

The expansion and evolution of bikesharing systems is a global phenomenon, which has motivated research to characterize "best practices" in both system operations and policy transferability across regions. Little is known, however, about the pros and cons of different approaches to define scale and zoning schemes in bikesharing evaluation. This research begins to address this challenge by juxtaposing station-level and community-level approaches to model and estimate the Annual Average Daily Bicyclist (AADB). We use the demand information from 459 Divvy stations in Chicago between June 1, 2015 and May 31, 2016 to assess the aggregation approaches concerning variable impacts, model specification, and prediction accuracy. Elasticity calculations, prediction error comparisons, and influence analysis reveal the importance of both built environment and sociodemographic variables in bikeshare modeling and the need to develop context-sensitive interventions. The detailed comparison of different levels of aggregation for analysis of bikeshare demand and user impact highlights that each level contributes insights to planners and policymakers. While the disaggregate approach provides the most information for planners in terms of improving bikeshare systems, there is value in adopting an aggregated approach for transport policy that accounts for potential neighborhood effects. In addition, the control for socio-demographic factors around stations highlights the variation in socio-spatial effects that planners need to account for when measuring outcomes and equity impacts.

1. Introduction

Numerous mobility services are emerging in urban areas across the world with the intent of tackling environmental issues or social disparities arising from current societal practices. Research on bikesharing schemes, particularly their diffusion within and between cities in addition to learning processes for potential users and municipalities, thus faces a critical juncture in determining their viability as a competitive travel mode alternative (Parkes et al., 2013).

In a synthesis of research findings, Ricci (2015) poses three broad questions that echo the uncertainty surrounding the future of bikesharing: (1) *Who uses bikeshare and how do these individuals use it*; (2) *What are the impacts of bikeshare on other travel modes and individual behaviors*; and (3) *How should cities implement and operate bikeshare systems*. She touches on the concerns about the risks of homogeneous user profiles that underpin the design of bikesharing schemes, resulting in wide-ranging failures to function in an equitable manner. The fact that socioeconomic characteristics are likely to hold significant explanatory power in describing cycling behaviors prompts the notion of transportation equity, or fairness, as it relates to the benefits and negative externalities of certain forms of mobility. To our knowledge, only one bikeshare study examines this issue as a primary focus, demonstrating that innovative forms of mobility are not immune to producing inequalities, particularly with respect to socioeconomic status (Goodman and Cheshire, 2014).

What this research agenda seems to also lack is a push to demarcate the trade-offs of adopting different scales of analysis, as well as defining various spatial partitioning schemes, in demand model specification. We aim to address the former challenge by juxtaposing station-level and community-level approaches to estimating the Annual Average Daily Bicyclist (AADB), defined as the mean per-day trip origin count for each station over one year of trip data for Chicago's Divvy bikeshare system. We augment the data using U.S. Census estimates, the EPA Smart Location database, and the Chicago Data Portal to additionally account for local and spatial variation in outcomes. This enables us to assess the impact of two competing aggregation approaches concerning variable impacts as well as the sensitivity of our models and findings.

Our contribution to the existing body of knowledge of bikesharing demand analysis is thus twofold. *First*, we explore the socioeconomic characteristics of station areas in conjunction with the traditional

* Corresponding author.
E-mail addresses: alecbiehl2019@u.northwestern.edu (A. Biehl), alireza.ermagun@northwestern.edu (A. Ermagun), a-stathopoulos@northwestern.edu (A. Stathopoulos).

https://doi.org/10.1016/j.jtrangeo.2017.11.008
Received 10 July 2017; Received in revised form 26 September 2017; Accepted 16 November 2017
0966-6923/

explanatory variables relating to the physical environment. This will help frame the success of a bikeshare system from a deeper social and geographic perspective, namely to increase awareness of which population groups lack access to shared mobility services. The growing popularity of adopting an equity-conscious research perspective is redefining modeling approaches to long-standing questions in travel behavior research (Lucas, 2012; Lucas et al., 2016). Expanding on this point, Médard de Chardon et al. (2017) demonstrate through a comparison of 75 bikeshare systems that their purpose and success metrics are ubiquitously elusive, which jeopardizes not only the efficiency of planning processes but also attempts to redress social equity. This latter point is imperative given recent findings that personal safety and police profiling concerns are prevalent barriers to cycling in Black and Hispanic communities (Brown, 2016).

Second, in accordance with the Modifiable Areal Unit Problem (MAUP), we demonstrate the statistical and policy implications arising from the comparison of two scales of analysis. Although the MAUP has guided research relating to several spatial problems, including land use analysis (Jacobs-Crisioni et al., 2014) and defining service areas (Wang et al., 2014), it has not been implemented in bikeshare research to our knowledge. An important result is the apparent scale invariance (e.g. distance to CBD) or lack thereof (e.g. distance to nearest transit station) of the variable elasticities. Moreover, we use cross-validation to assess the predictive capabilities of our model, in addition to measuring the influence of the predictor variables, based on several criteria, to further elucidate the implications of the observed scaling effects.

The rest of the paper is organized as follows. Section 2 reviews the literature on bikeshare demand while Section 3 introduces the MAUP, along with some applications in the transportation literature. Section 4 details the data preparation process. We subsequently provide the analytical methods and results in Section 5. The final section summarizes our findings and offers suggestions for future work. We believe that a more systematic implementation of this framework would add transparent value for urban and transportation planners in terms of "diagnosing" their methodological approaches, guiding their selection of a preferred level of analysis for a given analytical goal, and evaluating appropriate "success" factors for bikesharing schemes.

2. Bikesharing: what do we know?

We synthesize the literature on bikesharing demand analysis from two distinct standpoints. First, we examine the relationship between the level of data aggregation and research aims, as well as that between the unit of analysis and methodological approach utilized. Second, we overview the factors relevant to modeling and predicting demand, stressing the importance of both physical and social environment characteristics.

2.1. Demand estimation

Table 1 summarizes 13 studies we identified as representative of the current literature on bikesharing demand. We do not intend this table to be exhaustive, but rather a fair coverage of analytical methods and geographical scales. We find that most of the research examines bikesharing at a disaggregate level.

Two primary approaches characterize these papers: classification or prediction. The classification approach typically relies on some form of profiling algorithm to characterize demand or mobility patterns. The units of analysis are either users (Vogel et al., 2014) or stations (Borgnat et al., 2011), both of which produce clusters or typologies based on temporal characteristics and primary trip purposes. The prediction approach attempts to forecast bikeshare demand using a set of explanatory variables that could have a significant impact on station usage. One example is the use of Poisson regression to predict the log of average daily trips based on weather conditions and calendar events (Corcoran et al., 2014).

Six of the studies in Table 1 aggregate bikeshare data to present some overall phenomena in the system that are not obvious using disaggregated data. Faghih-Imani and Eluru (2016) develop a single measure based on number of stations, station capacity, and areal size to represent bikeshare infrastructure in Montreal's Traffic Analysis Zones (TAZs). They estimate three spatiotemporal econometric models to determine the impact of various land use and infrastructure features on arrival and departure rates. Even when only conducting simple exploratory analyses, it is possible to examine the performance of a bikeshare system in different neighborhoods (Ahillen et al., 2016) or for comparison between traditionally deprived and wealthy areas of a city (Goodman and Cheshire, 2014). An important research gap is identified as none of the existing studies attempt to compare how the choice of zoning scheme or geographic scale could influence model results.

2.2. Impact of physical and social environment variables

Table 2 synthesizes the findings related to various explanatory variables that might affect bikeshare demand: the plus sign represents a positive correlation, the minus sign represents a negative correlation, and zero indicates statistical insignificance. Four studies relied solely on trip data and sought to either classify usage patterns or develop a new methodology for predicting usage. Consequently, Table 2 does not include these studies.

As stated by Wang and Zhou (2017), "the built environment impact on travel behavior is perhaps the most studied topic in travel behavior research." This should not be surprising: if the physical built environment is not conducive to increase use of any particular travel mode, such as cycling, then policymakers' attempts to enhance both individual (e.g. increased physical activity) and environmental (e.g. reduction in fossil fuel emissions) well-being through behavior change would be futile. In an extensive review of active travel, Wang et al. (2016) point to land use form, safety measures, accessibility, and street connectivity as major influences on individuals' physical activity within the urban landscape. In the specific context of bikeshare demand, we see that most of the explanatory variables pertain to the (built) environment. For instance, mixed land use, a higher number of restaurants, and stronger network connectivity typically have a positive impact, while distance to the central business district and high levels of precipitation/humidity/wind typically exhibit negative effects. Some of the variable effects are more nuanced due to the temporal analyses within the studies. As an example, the importance of commercial and recreational land use types for inducing trips depends on the day of the week (Zhang et al., 2017; Zhou, 2015). Additionally, the impacts of station capacity and number of public transit and other bikeshare stations in proximity will differ depending on whether system members or temporary customers are analyzed (Faghih-Imani and Eluru, 2015), as well as across studies.

Table 2 also shows that personal or social/economic factors are underrepresented in existing bikeshare demand studies. Researchers should be cautious, however, when using solely objective data in constructing models of bikeshare usage for policy formulation. From a more general perspective, Moudon et al. (2005) determine that, while a favorable built environment is a necessary condition for promoting cycling, it is not sufficient: personal factors, such as sociodemographic characteristics and perceptions, are more important determinants. Similar conclusions are evident in more recent literature on bikeshare use, especially in terms of combining traditional survey data with novel operational "big data" sets (Mateo-Babiano et al., 2016). In addition, there is an increased awareness of the benefits of qualitative research in the transportation literature, typically as part of a mixed-method design. These types of studies have reinforced the findings of quantitative work regarding the significance of both actual and perceived levels of accessibility and safety (Fishman et al., 2012), as well as the legibility of the transportation network and its impact on navigational and wayfinding exercises for aspiring cyclists (Ó Tuama, 2015).