The Journal of Transport and Land Use jtlu.org
Vol. 18 No. 1 [2025] pp. 57–75

How bikesharing changed destination distance for its users: A case study of Chicago Metropolitan Area

Shubhayan Ukil (corresponding author)
University of Michigan–Ann Arbor
ukilshub@umich.edu

Aditi Misra
University of Colorado, Denver
aditi.misra@ucdenver.edu

Abstract: Shared bike use has been growing, especially post-pandemic, because it improves personal mobility and provides an alternative to walking while increasing connectivity to transit services. Existing research has examined the impact of these services on mode share and vehicle ownership. However, these services also hold the potential to influence the distance people travel to reach destinations. In this study, we examine the impact of Divvy shared bike services in the Chicago metropolitan region on the average trip distance of its users across all trips between 2008, when the service was not operational in the city, and 2018. We use repeated cross-sectional household travel datasets from 2008 and 2018 for analysis. We perform difference-in-difference regression to calculate the change in average trip distance for the shared bike user group. As there is no way to track people in repeated cross-sectional datasets, unlike a panel dataset, we use propensity score matching to match users between the two datasets. The results indicate that the average trip distance is reduced by 0.841 km (miles) for the shared bike user group with the presence of shared bike services. Shared bike users are more likely to live in urban areas where destinations are in proximity and use multi-modal travel, which could be a reason for this group's reduced average trip distance. Given our findings, we recommend planning for shared bike services integrated with transit in urban areas and promoting mixed land use so that users can choose proximate destinations in dense urban areas.

Keywords: Bikeshare, travel distance, quasi-experimental, destination, causal effect

Article history:
Received: October 4, 2024
Accepted: November 20, 2024
Available online: February 11, 2025

1 Introduction

Micro-mobility services, which are low-speed, small, lightweight vehicles, such as bikes, e-scooters, etc., have helped bridge the last-mile connectivity issues, enabling access to destinations. Their use has grown rapidly post the COVID-19 pandemic (Fukushige et al., 2022). In addition to increasing personal mobility choices in a sustainable manner, these services hold the potential to allow people to shift their choice of retail shopping location, dining, etc., especially by improving last-mile connectivity (Martin & Xu, 2022). Micro-mobility services enable faster travel than walking (Griffin & Sener, 2016), so the reduced travel time can be used to travel farther distances. People

https://doi.org/10.5198/jtlu.2025.2597
ISSN: 1938-7849

The *Journal of Transport and Land Use* is the official journal of the World Society for Transport and Land Use (WSTLUR) and is published and sponsored by the University of Minnesota Center for Transportation Studies. This paper is also published with additional sponsorship from WSTLUR.

may also choose to travel to locations that were previously not accessible by personal vehicle due to high parking costs through improved last-mile connectivity to transit services and also as it provides an alternative to walking (Aman et al., 2021). However, existing literature has not looked into how the presence of these services may have influenced the distance shared bike users are traveling to access different destinations, which can be a proxy measure for change in destination choices with increased mobility options.

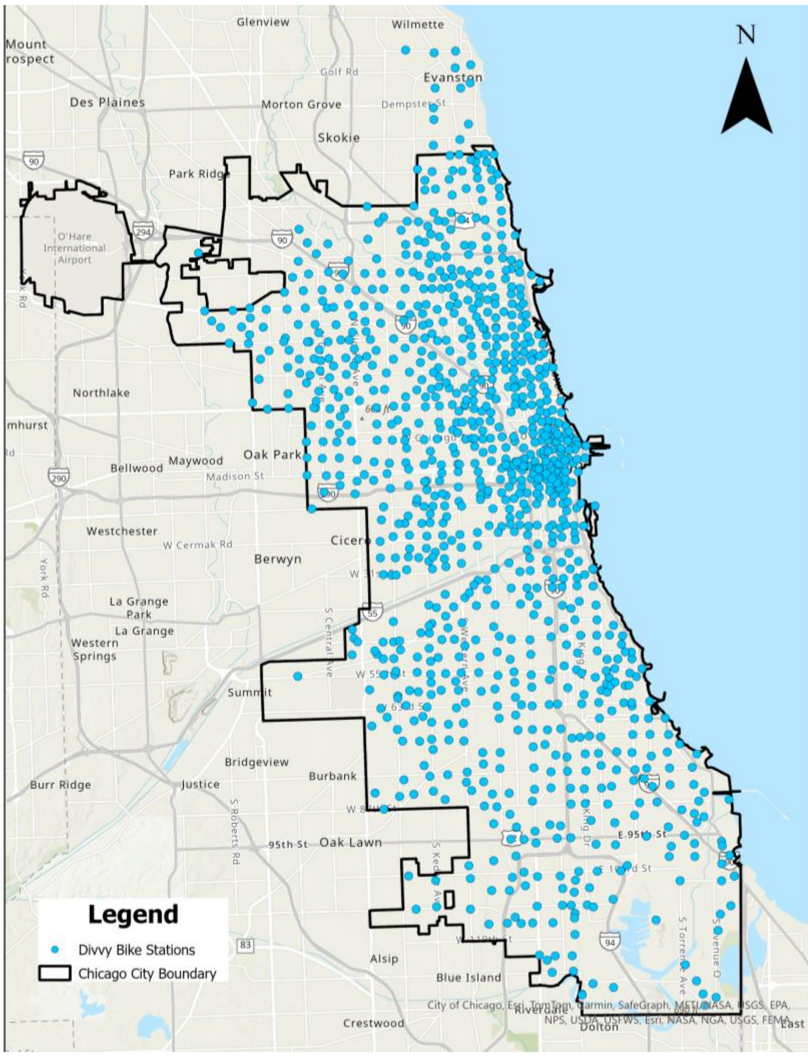

Figure 1: Divvy bike stations location (map not to scale)

This study examines the impact of the presence of Divvy shared bike services on change in average trip distance for the shared bike user group in Chicago. A figure showing the presence of Divvy stations is shown in Figure 1. We hypothesize that as shared bike services become available and popular, people who choose to use these services travel to more proximate destinations because these services are primarily used in urban areas where destinations are closer, resulting in an overall lower average trip distance over time. Trip distance is not a direct travel choice but a result of decisions regarding travel mode, trip purpose, and destination choice (Maat & Timmermans, 2009).

Therefore, analyzing changes in average trip distance can help understand any shifts in destination choice when the travel mode and trip purpose are constant. We also control for built environment variables (density, entropy, etc.) at the origin and destination location to understand the changes in trip distance.

We used repeated cross-sectional Household Travel Survey datasets from 2018 and 2008 from the Chicago Metropolitan Agency for Planning (CMAP) to calculate the change in average trip distance. In the 2018 dataset, Divvy shared bike service is included as a travel mode, which is absent in the 2008 dataset. To establish a comparable group in 2008, we use Propensity Score Matching (PSM) to identify individuals who may have used the services had they been available in 2008. This user group helps us to eliminate selection and longitudinal bias when comparing the outcome variable across population groups in the two datasets. Similarly, we used PSM to identify non-users of these services in 2018, addressing selection bias. The change in the average trip distance for people using shared bike services is analyzed using Difference-in-Difference (DiD) regression between the years 2008 and 2018. DiD involves measuring the change in an outcome variable over time for a group that undergoes an intervention (people choosing to use bike-share in our case) compared to a group not subjected to the intervention (people not using bike-share in our case). Our study contributes to the existing literature on bike-sharing by longitudinally comparing the effect of bike-sharing on distance travelled, which has not been done before, using a quasi-experimental causal framework. Understanding this trend in the distance travelled is important because it can influence long-range plans, land-use planning, and traffic impact assessment studies. The results of this study can help planners and policymakers guide policies on promoting the use of bike-sharing services together with transit to increase destination options for shared bike users. Further, it can help to understand changes in the trip distribution patterns of shared bike users, which impacts demand forecasting, station placement, and rebalancing.

2 Literature review

2.1 Who are the shared bike services users?

Shared bike services are being used mostly by white, young, and middle-aged, men who are university-educated (Faghih-Imani & Eluru, 2015). They work full-time or part-time and have a middle to high household income. They live close to a bike station in urban areas (Bachand-Marleau et al., 2012), with work locations close by and they mostly do not own vehicles or have children at home. Land use and other urban form measures, such as higher job or population densities, influence the use of shared bike services. As a result, these services are primarily located in more urban locations (Faghih-Imani & Eluru, 2015; Mix et al., 2022). However, the relationship between gender and shared bike services varies across countries, depending on the mode share of biking, in general. In the United States, United Kingdom, and Australia, where mode share for bikes is lower, between 65% and 90% of cycling trips are by men. By contrast, in countries where bikes are extensively used for travel, such as the Netherlands, women cycle more than men. Hence, bike-share services are used less by females in countries with lower bicycle usage overall. Bike share service users also differ from general bike riders. The former are more racially white, younger, and usually do not own a personal bike compared to the general bike riders (Fishman, 2016). Further, these services are used more for commuting and non-work purposes. Their use is higher on weekdays than on weekends as these services are used often for daily commutes (Zhang et al., 2016). The users of these services can be classified into two groups by purpose. One group is predominantly active mode users, primarily using shared bike services for daily commutes. Whereas, the other group uses these services in combination with ride-hailing, transit, carshare, and active modes, both for commute and non-work purposes (Mohiuddin et al., 2022; Reck & Axhausen, 2021; Shaheen et al., 2012; Zhang et al., 2016).

Hence, the users of shared bike services are predominantly white men with a higher-income, full-time, or part-time job. Their demographic and socioeconomic characteristics differ from those of regular bike users. These services are used in high-density, mixed-use urban areas because of their availability, and the people living in urban areas rely on them more because of lower household vehicle ownership. When examining any change in travel or destination behavior with the growth of shared bike services, it is important to consider all these different socioeconomic variables that characterize shared bike users.

2.2 Influence on travel behavior of shared bike services

Shared bike services are changing the travel behavior of its users. They mainly substitute only walking or transit trips for work purposes and car trips to some extent for non-work purposes. It also can induce new trips to restaurants and for recreational purposes to some extent (Bieliński et al., 2021; Tatsuya Fukushige, 2021; Zhou et al., 2023). They also impact public transit ridership as they increase connectivity to transit services and lead to an increase in walking**.** However, mode substitution of bike share is influenced by weather and drops during the winter months**.** In addition, it also allows shared bike users to make quicker trips with bicycles, enabling them to reach their destinations sooner than transit service. Mode substitution of car trips with shared bikes can also reduce vehicle miles traveled (VMT) in cities with high car use for commuter trips (Fishman et al., 2014). However, the operational tasks of redistributing the shared bikes using vans or trucks can partially offset the reduction in VMT can be due to and due to some induced travel for recreational trips (Fukushige et al., 2023). Also, some studies have found that shared bikes reduce automobile use and vehicle ownership. However, the magnitude of change is mediated by the presence of transit facilities in a city (Shaheen et al., 2012).

Apart from the influence on mode choice, vehicle ownership, VMT, etc., shared bike services have also increased the opportunity for people to engage in different activities, primarily through improving last-mile connectivity in conjunction with a bus or train for a longer trip. The latter applies to shared bike services (Bieliński et al., 2021). Most studies have measured the impact of shared bike services in understanding mode substitution, VMT, and last/mile connectivity in existing research. However, these studies do not analyze the ability of shared bike users to engage in activities that are located farther or close by as these new services become available for travel. People may be willing to travel a different distance due to the presence of these services. Although dense, mixed-use urban structures are known to make people travel short distances (Maat & Timmermans, 2009), shared bike users, who mostly live in urban areas, may also travel farther distances. A mode that improves last-mile connectivity with a constant travel time budget may enable people to expand their activity space (Scheiner, 2010). Travel distance to destination and built environment also guide the travel mode. As shared users predominantly live in urban areas with destinations closer, with the presence of shared bike services, their average trip distance may be reduced even when they travel to a different destination.

3 Data

We utilized the most recent Household Travel Survey Datasets from 2018 and 2008, released by CMAP for the Chicago Metropolitan region (Chicago Metropolitan Agency for Planning, 2019; Chicago Metropolitan Agency for Planning, 2008). Divvy bike share has been operational in the area since 2013. These household datasets offer comprehensive trip diaries of respondents in the dataset, encompassing their trip details such as trip purpose, travel distance, travel time, travel day, trip mode, and time of day. They also provide the household and individual demographics, socioeconomic characteristics, household members information, and geographic census data of the respondents. The 2008 dataset employed a dual-frame sampling approach, combining random digit dialing with complete directory coverage and address-based samples to capture diverse demographics, including low-income, minority, renters, new residents, and cell-only households. This dataset predates the emergence of app-based mobility services and the onset of the 2008 recession. By contrast, the 2018 dataset was collected using diverse paper travel logs, web reporting, telephonic reporting, or real-time app recording methods. The data collection was done after these services had been in operation but before the onset of the COVID-19 pandemic. The travel diary does not include the cost of using the shared bike trips. Additionally, Smart Location Mapping data versions 2.0 and 3.0 (from the Environment Protection Agency (EPA) and Longitudinal Employer-Household Dynamics were used to calculate the built environment measures. The variables included population density, residential density, job density, job entropy, employment, and household entropy at a census tract level for the years 2020 (equivalent to 2018) and 2008 (with housing and population density calculated for 2010) US Census Bureau, 2023). Note that the Smart Location Mapping data version 3.0 uses 2018 Census ACS (5-Year Estimate) and 2017 Census LEHD data; both are from pre-pandemic time.

Table 1. Descriptive statistics of the different groups of users

Variables	2018 - Shared Bike Users	2018 - Shared Bike and Ride-hailing Users	2018 - Non-Shared Bike or Ride-hailing Users	2008 - Pre-Shared Bike era travelers
Count	95	24	22688	17211
Age*	36.3 (10.8)	34.7 (9.4)	43.2 (16.1)	50.7 (18.1)
Gender (M: Male, F: female)	M - 60 %, F - 40%	M - 67%, F - 33%	M - 46%, F - 54%	M - 46%, F - 54%
Driver's License Status	98%	96%	91%	90%
Hispanic Status	3%	0%	9%	3%
Race-White	85%	80%	79%	42%
Race - Black	4%	8%	10%	8%
Race-Asian	4%	8%	5%	1%
Employment Status (Full or part time)	93%	96%	73%	68%
Student Status	14%	8%	15%	10%
Education Level - bachelor’s degree	53%	38%	34%	27%
Education Level - Graduate Degree or higher	43%	58%	27%	25%

Disability Status	1%	0%	5%	8%
HH Vehicle Ownership - 0	46%	42%	9%	7%
HH Vehicle Ownership - 1	39%	54%	27%	29%
HH Vehicle Ownership - 2	11%	0%	44%	44%
HH Vehicle Ownership - 3	3%	0%	14%	14%
HH Vehicle Ownership - 4 or more	1%	4%	6%	6%
HH Size*	2.14 (1.08)	1.79 (0.9)	3.09(1.4)	2.67 (1.3)
HH Income - < 35k	12%	4%	10%	16%
HH Income - 35k - 50k	4%	4%	7%	12%
HH Income - 50k-59k	2%	4%	7%	8%
HH Income - 60k-75k	10%	30%	10%	12%
HH Income - 75-100k	9%	4%	16%	19%
HH Income - >100k	63%	54%	50%	33%
Residence in City of Chicago Boundary	93%	93%	32%	32%
Employment Density at residential location**	25.15 (57.1)	39.60 (78.9)	9.13 (55.7)	6.41 (29.7)
Residential Density at residential location**	27.82 (24.2)	26.49 (23.3)	8.55 (15.0)	7.68 (18.3)
Population Density at residential location**	45.80 (31.9)	44.83 (30.3)	17.23 (22.9)	15.83 (42.8)
Employment and Household Entropy at residential location**	0.55 (0.2)	0.60 (0.1)	0.56 (0.1)	0.62 (0.2)

.* Mean with Standard Deviation in Parenthesis
**Calculated at census tract level
HH refers to household

The 2018 and 2008 raw datasets have 23,708 and 23,819 person observations, respectively. Only the people who live within the seven-county CMAP region are used for analysis. Further, we dropped the observations with missing demographic, socioeconomic, and household information, the key variables for analysis. In the 2018 dataset, 95 persons used shared bike (Divvy) services. Also, 24 persons used shared bike services along with other app-based ride-hailing services. However, as they are low in the count to be analyzed independently and are demographically similar to the shared bike users, they were considered shared bike users for analysis. Shared bike users who used the services to travel for either non-work purposes or both work and non-work purposes were considered in the study. Descriptive statistics of the key demographic and socioeconomic characteristics of the different categories of users used in the PSM are given in Table 1. Age and household size are continuous variables, whereas race, driving license status, employment status, education status, and disability status are binary variables. We recategorized household vehicle ownership into a categorical variable. The household vehicle ownership category of 4 or above was recategorized into a single category of "4 or more." Household Income is a categorical variable with six different levels of less than $35k, $35-50k, $50k- 59k, $60k-75k, $%75-100k, and more than $100k.

There has been an overall change in some of the key demographic and socioeconomic variables of the people in the 2018 household travel survey dataset compared to the 2008 dataset for the CMAP area. The mean age reduced, the percentage of Hispanic respondents increased, and the percentage of White and Asian populations increased in the region. Also, the percentage of people employed increased, the percentage of people with a bachelor's degree, and those with a master's degree increased. Further, people reduced their household vehicle ownership, and household income increased in the region. The demographic, socioeconomic, and household characteristics of shared-bike users are similar to the known characteristics of users of these services (as discussed in Section 2). Most shared bike users in the 2018 Household Travel Survey dataset lived within the Chicago City limits. Their residential census tract has a higher residential population and job density than the overall samples in the dataset, which matches the known residential built environment of shared and micro-mobility users in general.

4 Methodology

To effectively understand the impact of a treatment on the outcome for the treatment group, it is important to compare it to the change in outcome variable of the control group. The counterfactual comparison separates the impact of other confounding temporal events over time to calculate the effect of the treatment on the treatment group. Hence, in this study, we examined the effect of the presence of shared bike services (treatment) on the average trip distance of the shared bike group (treatment group) between the pre- and post-shared bike era through a simple linear DiD regression using Pooled Ordinary Least Square estimator (Wooldridge, 2021).

We use CMAP's repeated cross-sectional household travel survey datasets to perform the analysis. However, the 2008 dataset does not have a known group of potential shared bike users that can be readily compared to the shared bike users from the 2018 dataset. Also, the counterfactual group of non-shared bike users in the pre- and post- shared bike era are not known in the two datasets from 2008 and 2018, respectively. Hence, we use Propensity Score Matching (PSM to identify the potential shared bike user group (pre-treatment group), which is similar to the 2018 shared bike users (post- treatment group) that helps to eliminate any cross-sectional bias and longitudinal incomparability between the groups in 2008 and 2018. PSM involves pairing statistically similar individuals both cross-sectionally and longitudinally. In the absence of a panel dataset, which is cost-intensive and often has a high attrition rate, quasi-experimental methods like PSM help to create a balanced treatment and control group and provide reliable results from any subsequent causal analysis (Stuart et al., 2014). In this study, the treatment corresponds to the presence of shared bike services in 2018 compared to the base year of 2008. The propensity score is defined as the conditional probability of assigning a person to a specific treatment group based on observed demographic, socioeconomic, and other covariates (Rosenbaum & Rubin, 1983). In a simple PSM, binary logistic regression estimates the propensity score for observations that receive treatment and those that do not. It is a balancing score for the distribution of measured covariates between the treatment and control groups. For each observation in the treatment group, an observation with a similar propensity score, based on the covariates, from the group that did not receive treatment serves as its control. In this analysis, the calculated score for each shared bike user represents the probability of being a shared bike services user group member. Hence, any outcome variable between these groups is best analyzed at a group level.

In this study, we employed two-dimensional PSM to address both cross-sectional bias and longitudinal incomparability bias between treatment and control groups cross-sectionally and over time, improving the results of any subsequent causal analysis. Zhong et al. (2021) discussed using two-dimensional PSM to eliminate selection bias and longitudinal incomparability bias when matching users from repeated cross-sectional datasets over two time periods. This method identifies four groups: pre-treatment, post-treatment, pre-control, and post-control based on a known group (pre-treatment) that are balanced in the covariates, reduce bias, and mimic a randomized controlled experiment. It helps to control for the selection bias between the two groups, i.e., the unequal probability of a subject being assigned to treatment or a control group from the same dataset at a cross-sectional level. It simultaneously controls for longitudinal incomparability between treatment and control groups over time. Also, matching and regression methods perform best in combination, as PSM helps to hold the parallel trend assumption. It minimizes the bias that may occur if the treatment and control groups are different (Becker & Hvide, 2013).

The matching process used for analysis is adapted based on the needs of the study and is represented in Figure 2. We included all measured baseline characteristics of shared bike users in the PSM model that can affect their use of shared bike services. Four groups of users were identified, with three pairs of matchings. The groups were after-change users of shared bike services (Group A), after-change non-users of shared bike services (Group C), before-change people who may have used shared bike services if they were available (Group B), and before-change people who may not have used shared bike services even if they were available (Group D).

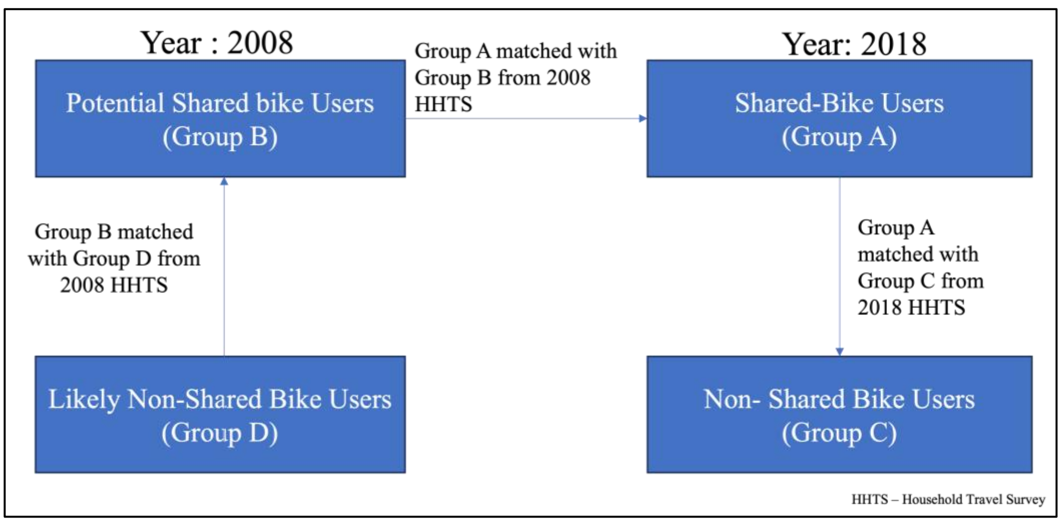

Figure 2. Two-dimensional propensity score matching diagram

Firstly, longitudinal matching was performed by matching shared bike users in the 2018 Household Travel Survey dataset (Group A) to people from the 2008 Household Travel Survey dataset based on the demographic, socioeconomic, and residential built environment covariates. Group B, as identified from this matching, represents the people who might have potentially used the shared bike services if they had been available in 2008. Next, a subset of non-users of shared bike services (Group C) was identified from the 2018 Household Travel Survey dataset by cross-sectional matching with the shared bike users in 2018 (Group A) based on the same covariates. This group represents the non-users of shared bike services in 2018. The final group is those who might not have used shared bike services even if they had been available in 2008 (Group D). This group was identified from the 2008 Household Travel Survey dataset based on cross-sectional matching with the group (Group B) who might have potentially used these services if they were available in 2008. It needs to be noted that all four groups of users are mutually exclusive, such that the same user from the 2008 Household Travel Survey dataset cannot be both in Group B and D.

The imbalance in matching between the treatment and control groups is measured by the Standardized Mean Difference (SMD), a balance statistic used to measure the balance of covariate distribution between the treatment and control groups. The SMD (δ) is calculated with the formula (Oakes & Johnson, 2006) shown in Equation 1:

$$\delta = \frac{|\bar{X}t - \bar{X}c\,|}{\sqrt{\frac{s_t^2 + s_c^2}{2}}} \qquad (1)$$

where, $\bar{X}t$ is the mean of the covariate for the treatment group
$\bar{X}c$ is the mean of the covariate for the control group
s_t is the standard deviation for the treatment group
s_c is the standard deviation for the control group

SMD less than 0.1 (10%) between treatment and control groups across covariates indicates good balance.

After PSM, DiD regression was performed based on the four identified groups of current and potential users and non-users of shared bike services to estimate the change in average trip distance for the shared bike user group. Figure 3 shows how the change in outcome is measured for the treatment group (intervention effect) due to the intervention. The change in outcome due to all other factors is teased out by accounting for the change in outcome for the control group. The DiD effect is assumed to be homogenous across time and observations in the treatment group, which is an important assumption for measuring the impact. Covariates that were not perfectly balanced between the treatment and control group in PSM were controlled in DiD regression. The regression adjustment helps to clean up the small residual covariate imbalance between the groups (Stuart, 2010). Covariates that do not vary over time were added as control variables to improve the precision of estimates in DiD regression (Angrist & Pischke, 2009). Controlling for covariates in DiD regression helps to hold the parallel trend assumption (Huntington-Klein, 2022).

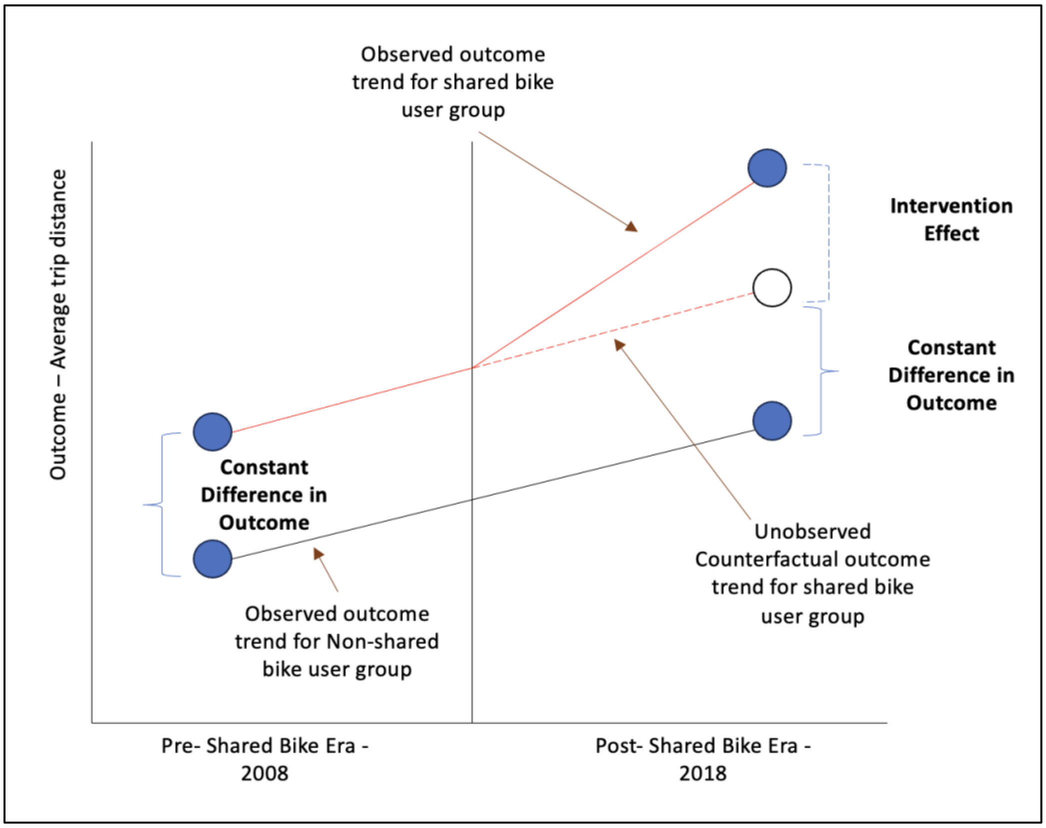

Figure 3. Schematic for difference-in-difference regression

The DiD regression model is represented in Equation 2:

$$y = \beta_0 + \beta_1 * t + \beta_2 * c + \beta_3 * t * C + \beta_4 * Cov + \varepsilon \qquad (2)$$

where, y is the outcome of interest

t is the time variable

C is the change or intervention variable

β_0 is the baseline average

β_1 is the time trend in the control groups

β_2 is the difference between the treatment and control group pre-treatment/intervention

β_3 is the difference in changes over time for the treatment group because of the treatment/intervention

β_4 is a vector of the coefficient for the covariates matrix of control variables used to improve the precision of β_3

Cov is the c matrix of the control variables in the model

ε is the error term, which is normally distributed

5 Results

5.1 Propensity score matching

PSM was implemented in Rstudio Version 2022.12.0+353 with the MatchIt package. We matched people between the four user groups based on the methodology described in

Section 3. We included all potential confounding covariates that influence treatment and outcome. Also, the PSM model did not include any post-baseline covariates that could have been influenced or modified by the treatment (presence of shared bike services) to avoid “post-treatment” bias (Austin, 2011; Stuart et al., 2014). The matched sample was created with the nearest neighbor matching algorithm, and matching was undertaken without replacement to keep the independence-of-case assumption valid (Beal & Kupzyk, 2014). The final balance between samples on the covariates for all three pairs of matchings where the SMD is lowest across covariates is presented in Table 2 for users of shared bike services. The key covariates used in the matching process when treatment is the presence of shared bike services were the demographic and socioeconomic parameters that describe the users of these services, as discussed in section 59. Also, the urban built environment variables were included as covariates as they are associated with the presence of shared bike services as well as the built environment (density, diversity, etc.) in the residential location of a person influences their destination choice, and indirectly their trip distance (as discussed in section 59). Hence, the built environment parameters of the residential location of users were also included as covariates in the matching process. The groups in the matching process are named alphabetically for ease of reading, as described in section 3.

Table 2. Matching results for shared bike services users

	Group B matched based on Group A			Group C matched based on Group A			Group D matched based on Group B		
Covariates	**Means Treatment Group - A**	**Means Control Group - B**	**Std. Mean Diff.**	**Means Treatment Group - A**	**Means Control Group - C**	**Std. Mean Diff.**	**Means Treatment Group - B**	**Means Control Group - D**	**Std. Mean Diff.**
Age	36.03	36.53	-0.05	36.03	36.88	-0.08	36.5	36.09	0.04
Gender - Male	0.61	0.58	0.07	0.61	0.50	0.22	0.58	0.61	-0.05
Gender - Female	0.39	0.42	-0.07	0.39	0.50	-0.22	0.42	0.40	0.05
Vehicle License Ownership	0.97	0.96	0.11	0.97	0.99	-0.11	0.96	0.98	-0.13
Hispanic status	0.03	0.03	0.00	0.03	0.03	0.00	0.03	0.04	-0.11
Race- White	0.84	0.84	0.00	0.84	0.87	-0.09	0.84	0.82	0.07
Race- Black	0.05	0.04	0.04	0.05	0.04	0.04	0.04	0.05	-0.04
Race- Asian	0.05	0.04	0.04	0.05	0.03	0.08	0.04	0.05	-0.04
Employed	0.93	0.91	0.10	0.93	0.97	-0.13	0.91	0.90	0.03
Student Status	0.13	0.16	-0.10	0.13	0.09	0.10	0.16	0.17	-0.02
Education - bachelor’s degree	0.50	0.55	-0.10	0.50	0.51	-0.02	0.55	0.50	0.12
Education - Graduate or higher	0.54	0.49	0.10	0.54	0.50	0.08	0.49	0.57	-0.17
HH Income - < 35k	0.10	0.17	-0.22	0.10	0.14	-0.14	0.17	0.17	0.00
HH Income - 35k - 50k	0.04	0.03	0.04	0.04	0.04	0.00	0.03	0.03	0.00
HH Income - 50k- 59k	0.03	0.03	-0.05	0.03	0.03	-0.05	0.03	0.03	0.00
HH Income - 60k-75k	0.13	0.14	-0.02	0.13	0.09	0.12	0.14	0.15	-0.02

HH Income - 75-100k	0.08	0.07	0.06	0.08	0.07	0.06	0.07	0.07	0.00
HH Income - >100k	0.61	0.55	0.12	0.61	0.62	-0.02	0.55	0.55	0.02
Household Vehicle Ownership - 0	0.45	0.42	0.07	0.45	0.44	0.03	0.42	0.29	0.27
Household Vehicle Ownership - 1	0.42	0.45	-0.07	0.42	0.45	-0.07	0.45	0.56	-0.22
Household Vehicle Ownership - 2	0.08	0.09	-0.03	0.08	0.08	0.03	0.09	0.11	-0.06
Household Vehicle Ownership - 3	0.03	0.03	0.00	0.03	0.02	0.05	0.03	0.03	-0.05
Household Vehicle Ownership - 4 or more	0.02	0.01	0.07	0.02	0.02	0.00	0.01	0.01	0.00
Gross residential density (HU/acre) at residential location*	27.56	26.05	0.06	27.56	29.11	-0.06	26.05	23.26	0.10
Gross Population density (people/acre) at residential location*	45.61	40.75	0.15	45.61	47.23	-0.05	40.75	36.76	0.12
Employment and Household Entropy at residential location*	0.57	0.54	0.13	0.57	0.57	0.00	0.54	0.54	0.03
Employment density at residential location*	28.07	25.63	0.04	28.07	23.13	0.08	25.63	17.98	0.12

DK/RF - Don't Know/ Refuse to answer.
Calculated at census tract level
HU refers to housing units.

In the matching process, with the treatment being the presence of shared bike services in 2018 compared to the base year of 2008, the covariates are mostly well balanced between the treated and control subjects when matching is done to identify Group B based on Group A. The SMD for the vehicle license ownership, household income category of more than 100k, employment, density gross population density, and household entropy at residential location variables are marginally higher or lower than the acceptable range of 0.1 (10%). However, the household income category of less than 35k is -0.22, considerably higher than the acceptable range of SMD. In the case of matching done to identify Group C based on Group B, both from the 2018 Household Travel Survey dataset, some covariates have SMD greater than 0.1(10%). The covariates are gender, driving license ownership, employment status, and household income categories of less than 35k and 60-75k. The SMD value is marginally higher or lower than the acceptable range of 0.1 (10%) for all covariates except for gender (-0.22 for the female category). In the case of matching to identify users from Group D based on Group B, both from the 2008 Household Travel Survey dataset, some of the covariates have SMD greater than 0.1. The driving license ownership, Hispanic status, education level of bachelor's degree, gross population density, and employment density have SMD marginally above 0.1. However, the covariates of graduate or higher education level and household vehicle ownership categories of 0 and 1 vehicle have SMD of -0.17, 0.27, and -0.22, respectively. The covariates that have SMD above the acceptable range are tested as control variables in DiD regression, explained in Section 5.1.

5.2 Difference-in-difference regression

With the matched PSM results, the change in average trip distance for the 2018 shared bike user group due to the presence of shared bike services was analyzed with linear DiD regression. As matching has not been done at the trip level to calculate the change in trip distance, it is difficult to identify a trip of a non-shared bike user in 2018 or a potential shared bike user in 2008 that is comparable to a shared bike trip from the shared bike user group in 2008. Also, shared bikes are often used together with other modes. Hence, all trips performed by users in all four groups from the household travel survey datasets were included to perform DID regression to best estimate the average treatment effect for shared bike users on trip distance. Further, though Divvy shared bike services did not begin operating uniformly across Chicago area in 2013, the post-impact change being measured after five years in 2018 helps to hold the homogeneity assumption. It can be assumed that all people in the city experienced the presence of shared bike services after five years, and the average shared distance would be homogenous among people. The trips performed by the users in the four groups were processed to remove any trip that was the first for a user as it does not have an origin location person in the dataset. Also, any trip where the distance was unusually high or low for the time spent in travel by a specific mode, travel mode was missing, or those with origin or destination outside CMAP's seven-county region were removed. All other trips across all modes used by the users (walking, bike, automobile (both driving and as a passenger), CTA (bus and train), PACE, METRA, other transit, private shuttle, taxi, carpool/ vanpool, Uber/Lyft) were included in the analysis as PSM is not done at a trip level.

The DiD regression model was implemented in Rstudio Version 2022.12.0+353. We provided the 95% or 90% confidence interval for each estimate reported, along with standard error and p-values. The DiD model was adjusted for the demographic, socioeconomic, and built environment covariates that had SMD above 0.1 in the PSM model and were statistically significant in the DiD regression (as explained in Section 5). We controlled for the travel mode trip purpose and built environment characteristics at the origin/destination trip in the DiD model of average trip distance as the trip distance is not a travel choice but a result of these decisions. It helps to hold the parallel trend assumption. Also, we controlled for the built environment at the origin and destination of trips as they influence the outcome variable of trip distance. The average trip distance of the shared bike user group in 2018 was 2.77 km (miles), a non-user of the services in 2018 was 5.42 km (miles), a potential shared bike user in 2008 was 6.04 km (miles), and a potential non-user of the services in 2008 was 6.44 km (miles).

Table 3. DiD results of change in average trip distance of shared bike users

Variables	Estimate	Standard Error	p-value
(Intercept)	9.101*	0.688	< 2e-16
Trip Purpose: Work	0.974*	0.282	5.61e-04
Trip Purpose: Non-work activities with flexible destinations (shopping, dining, etc.)	-1.088*	0.266	4.44e-05
Trip Mode: Transit	4.361*	0.341	< 2e-16
Trip Mode: NMT	-4.047*	0.267	< 2e-16
Trip Mode: Bike share	-2.213*	0.452	1.05e-06
Gender: Female	-1.497*	0.222	1.95e-11
Education level: Graduate degree or higher	-0.438*	0.218	0.0451
Employment Status: Yes	1.575*	0.457	0.0006
Household vehicles - 1	-0.093	0.242	0.7002
Household vehicles - 2	1.361*	0.427	0.0015
Household vehicles - 3	0.199	0.685	0.7709
Household vehicles - 4 or more	-1.301**	0.732	0.0757
Residential density at home location	0.067*	0.022	0.0032
Population density at home location	-0.052*	0.017	0.0029
Employment and Population Entropy at home location	-1.336*	0.61	2.86e-02
Population density at origin#	-0.018*	0.004	8.22e-06
Employment density at destination #	-0.001*	0	6.89e-05
Residential density at destination#	0.07*	0.02	7.40e-04
Population density at destination#	-0.069*	0.016	2.18e-05
Year	-1.238*	0.353	0.0005
Trip by Shared Bike User	-0.089	0.336	0.7897
Year: Trip by Shared Bike User	-0.841**	0.464	0.0701
Adjusted R-squared: 0.2973			

* Significant at 95% Confidence interval, ** Significant at 90% Confidence interval
# calculated at census tract level
NMT: Non-Motorized Travel

For the shared bike user group, the linear DiD regression results for the change in average trip distance are presented in Table 3. The average trip distance of the shared bike user group in 2018, compared to potential shared bike users in 2008, was reduced by 0.841 km (miles), with the presence of shared bike services as a travel mode in 2018, significant at a 90% confidence interval (p-value: 0.07). The overall change in trip distance between 2008 and 2018 for the shared bike user group was -2.08 km (miles), with the presence of shared bike services and other factors that may have changed over time. The average trip distance for the non-user group of shared bike services in 2018, compared to the potential non-user group of the same services in 2008, also decreased by 1.23 km (miles), significant at a 95% confidence interval (p-value: 0.0004). The difference in average trip distance between the potential user group of shared bike services and the comparable group who would not have used the services, even if the services were available in 2008, is -0.089. However, it is not significant at a 95% or 90% confidence interval (p-value: 0.78). Although the overall average trip distance decreased for both user and non-non-user groups between 2008 and 2018, with the presence of shared bike services, the average trip distance decreased further for shared bike users.

6 Discussion and study limitations

For the shared bike user group, the DiD model of average trip distance shows that the trip distance was reduced by almost a mile between 2008 and 2018 with the presence of shared bike services. The change in trip distance is likely because shared bike users are more multi-modal travelers. They performed 57% of their trips using an active mode (walking trips: 38% and shared bike trips:19%) and 19% of trips with a transit mode, much less than non-users of shared bike services in 2018 (56%). The eco-friendly multi-modal travel of shared bike users can also be a result of the residential location of the shared bike users, as 92% of them live within the Chicago city limits, where transit and shared bikes are more easily accessible. By contrast, 70% of the matched potential users of shared bike services in 2008 lived within the Chicago city limits. A change in average trip distance with the presence of shared bike services could influence the mode share of shared bike users and the distribution of trips, which can impact travel demand forecasts. Any policy actions intended to improve sustainable travel with bike-share use need to plan for bike-share infrastructure, considering the new trip patterns. As diverse forms of micro-mobility services grow it can alter the economic viability of destinations to some extent. Further, environmental impact analysis performed to analyze the effect of bike share services should consider both changes in mode share and distribution of trips, as including the latter phenomenon improves the accuracy of impact.

Though the average trip distance has reduced for the shared bike user group, it is unclear whether shared bike users are choosing more proximate destinations because of the attractiveness of these nearby destinations or because they have a pro-environment attitude and prefer to use active or transit modes with limited geographic accessibility. The use of active travel modes is known to be more of a preference, resulting in higher satisfaction (De Vos, 2018; Singleton, 2017). Hence, shared bike users may use these services because of their personal choice. As shared bike users are accessing proximate destinations in a sustainable manner with multi-modal travel, transportation planners need to work towards better integrating shared bike services with transit services. The reduced average trip distance of shared bike users is likely an outcome of their preference and not because of their restricted geographic accessibility due to their travel mode since shared bike users.

Under an ideal scenario, the change in the average trip distance of shared bike users and non-users of these services between 2008 and 2018 should be the same if the services weren’t available in 2018. Apart from the presence of these services, other factors could also affect trip distance. These could be the travel costs, growth in dedicated bike lane infrastructure, etc., influencing the distance people travel to reach destinations. However, it is extremely difficult to get accurate information on bike infrastructure and its classifications (protected bike lanes, separate bikeways, buffered bike lanes, etc.). As the literature shows, different types of bike infrastructure have different impacts on users (Clark et al., 2019). As such, we could not find any reliable information on dedicated bike lanes for Chicago from 2008 to compare it to the 2018 bike infrastructure. Hence, we did not use this factor in our analysis, as incomplete information might lead to wrong conclusions and policy implications. In addition, at a system level, a decrease in distance can also increase the frequency of travel, negating any effect but the behavior is not explored in this study, which is a limitation of this study. From a methodological perspective, PSM is useful in the absence of true longitudinal data to leverage repeated cross-sectional datasets for analyzing change over time and to ensure the treatment and control groups are similar. However, the demographic and socioeconomic characteristics of the two periods may have changed. As a result, we have chosen cross-sectional matching in the PSM matching and only used longitudinal matching to match 2018 shared bike users to potential shared bike users in 2008, which we assume to be the next best alternative to having true panel data.

7 Conclusion

In this study, we have explored the impact of shared bike services beyond their impact on mobility aspects, investigating the impact on trip distance. We have taken a causal inference approach to understand the change in average trip distance. We have leveraged repeated cross-sectional household travel datasets, which are more easily available, unlike panel datasets, using PSM and DiD regression methods to understand the change in trip distance for shared bike users. Our results indicate that with the presence of shared bike services, the average trip distance has reduced for shared bike users, which could be a result of their multimodal travel and the users being residents of urban areas within Chicago city limits that have better access to transit, shared bike services, and walking infrastructure. The focus of planning should be to integrate shared bike services with other sustainable modes to increase access to destinations, while also reducing activity space of people with mixed land-use planning leading to an increase in proximity to multiple destinations in urban areas. Further analysis is needed to examine the impact of destination characteristics on destination choice in conjunction with mode availability and choice to understand better how they influence each other. Also, Divvy started electric shared bike services in 2020, around the beginning of the pandemic. The change in the Divvy bike share fleet and the shift in travel behavior due to the pandemic may have influenced trip distance and mode share for shared bike users. Hence, the analysis in this study can be further extended by adding a third time period to understand the impact of electric-shared bike services on trip distance by analyzing changes in average trip distance over the three time periods.

Acknowledgments

This research was made possible thanks to the statutory boundary data shared by the Chicago Metrpolitan Agency for Planning.

Author contribution

The authors confirm contribution to the paper as follows: Study conception and design: Shubhayan Ukil and Aditi Misra; data collection: Shubhayan Ukil; analysis: Shubhayan Ukil; draft manuscript preparation: Shubhayan Ukil and Aditi Misra.

References

Aman, J. J. C., Zakhem, M., & Smith-Colin, J. (2021). Towards equity in micromobility: Spatial analysis of access to bikes and scooters amongst disadvantaged populations. *Sustainability, 13*(21), 11856. https://doi.org/10.3390/su132111856

Angrist, J. D., & Pischke, J.-S. (2009). Differences-in-differences: Pre- and post-treatment and control. In *Mostly harmless econometrics: An empiricist's companion*. Princeton, NJ: Princeton University.

Austin, P. C. (2011). An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behavioral Research, 46*(3), 399–424. https://doi.org/10.1080/00273171.2011.568786

Bachand-Marleau, J., Lee, B. H. Y., & El-Geneidy, A. M. (2012). Better understanding of factors influencing likelihood of using shared bicycle systems and frequency of use. *Transportation Research Record, 2314*(1), 66–71. https://doi.org/10.3141/2314-09

Beal, S., & Kupzyk, K. (2014). An introduction to propensity scores: What, when, and how. *The Journal of Early Adolescence, 34,* 66–92. https://doi.org/10.1177/0272431613503215

Becker, S. O., & Hvide, H. K. (2013). *Do entrepreneurs matter?* IZA Discussion Papers, No. 7146. Bonn, Germany: Institute for the Study of Labor.

Bieliński, T., Kwapisz, A., & Ważna, A. (2021). Electric bike-sharing services mode substitution for driving, public transit, and cycling. *Transportation Research Part D: Transport and Environment, 96,* 102883. https://doi.org/10.1016/j.trd.2021.102883

Chicago Metropolitan Agency for Planning. (2008).Travel tracker survey. Chicago: Chicago Metropolitan Agency for Planning.

Chicago Metropolitan Agency for Planning. (2019). My daily travel survey. Chicago: Chicago Metropolitan Agency for Planning.

Clark, C., Mokhtarian, P., Circella, G., & Watkins, K. (2019). User preferences for bicycle infrastructure in communities with emerging cycling cultures. *Transportation Research Record, 2673*(12), 89–102. https://doi.org/10.1177/0361198119854084

De Vos, J. (2018). Do people travel with their preferred travel mode? Analyzing the extent of travel mode dissonance and its effect on travel satisfaction. *Transportation Research Part A: Policy and Practice, 117,* 261–274. https://doi.org/10.1016/j.tra.2018.08.034

EPA. (2011). Smart Location Database (Version 1), Geodatabase. Retrieved from https://www.epa.gov/smartgrowth/smart-location-mapping#SLD

EPA. (2021). Smart Location Database (Version 3), Geodatabase. Retrieved from https://www.epa.gov/smartgrowth/smart-location-mapping#SLD

Faghih-Imani, A., & Eluru, N. (2015). Analyzing bicycle-sharing system user destination choice preferences: Chicago's Divvy system. *Journal of Transport Geography, 44,* 53–64.

Fishman, E. (2016). Bikeshare: A review of recent literature. *Transport Reviews, 36*(1), 92–113. https://doi.org/10.1080/01441647.2015.1033036

Fishman, E., Washington, S., & Haworth, N. (2014). Bike share's impact on car use: Evidence from the United States, Great Britain, and Australia. *Transportation Research Part D: Transport and Environment, 31,* 13–20. https://doi.org/10.1016/j.trd.2014.05.013

Fukushige, T., Fitch, D. T., & Handy, S. (2023). Estimating vehicle-miles traveled reduced from dock-less e-bike-share: Evidence from Sacramento, California. *Transportation Research Part D: Transport and Environment, 117,* 103671. https://doi.org/10.1016/j.trd.2023.103671

Fukushige, T., Fitch, D. T., Mohiuddin, H., Andersen, H., & Jenn, A. (2022). *Micromobility trip characteristics, transit connections, and COVID-19 effects* (UC-ITS-2021-32). Berkeley, CA: The University of California Institute of Transportation Studies. https://escholarship.org/uc/item/2pk6t2cz

Griffin, G. P., & Sener, I. N. (2016). Planning for bike share connectivity to rail transit. *Journal of Public Transportation*, *19*(2), 1–22. https://doi.org/10.5038/2375-0901.19.2.1

Huntington-Klein, N. (2022). Difference-in-differences. In *The effect: An introduction to research design and causality*. Boca Raton, FL: CRC Press, Chapman & Hall. https://theeffectbook.net/

Maat, K., & Timmermans, H. J. P. (2009). A causal model relating urban form with daily travel distance through activity/travel decisions. *Transportation Planning and Technology, 32*(2), 115–134. https://doi.org/10.1080/03081060902861285

Martin, R., & Xu, Y. (2022). Is tech-enhanced bikeshare a substitute or complement for public transit? *Transportation Research Part A: Policy and Practice, 155*, 63–78. https://doi.org/10.1016/j.tra.2021.11.007

Mix, R., Hurtubia, R., & Raveau, S. (2022). Optimal location of bike-sharing stations: A built environment and accessibility approach. *Transportation Research Part A: Policy and Practice, 160,* 126–142. https://doi.org/10.1016/j.tra.2022.03.022

Mohiuddin, H., Fitch, D., & Handy, S. (2022). *Examining market segmentation to increase bike-share use: The case of the greater Sacramento region* (NCST-UCD-RR-22-29). Davis, CA: Institute of Transportation Studies.

Oakes, J., & Johnson, P. J. (2006). *Propensity score matching for social epidemiology,* Vol. 1 (pp. 364–386). San Francisco: Jossey Bass-Wiley.

Reck, D. J., & Axhausen, K. W. (2021). Who uses shared micro-mobility services? Empirical evidence from Zurich, Switzerland. *Transportation Research Part D: Transport and Environment, 94,* 102803. https://doi.org/10.1016/j.trd.2021.102803

Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*(1), 41–55. https://doi.org/10.1093/biomet/70.1.41

Scheiner, J. (2010). Interrelations between travel mode choice and trip distance: Trends in Germany 1976–2002. *Journal of Transport Geography, 18*(1), 75–84. https://doi.org/10.1016/j.jtrangeo.2009.01.001

Shaheen, S. A., Martin, E. W., Cohen, A. P., & Finson, R. S. (2012). *Public bikesharing in North America: Early operator and user understanding* (MTI Report 11-26). San Jose, CA: Mineta Transportation Institute.

Singleton, P. (2017). *Exploring the positive utility of travel and mode choice.* Portland, OR: Portland State University. https://doi.org/10.15760/etd.5664

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science: A Review Journal of the Institute of Mathematical Statistics, 25*(1), 1–21. https://doi.org/10.1214/09-STS313

Stuart, E. A., Huskamp, H. A., Duckworth, K., Simmons, J., Song, Z., Chernew, M., & Barry, C. L. (2014). Using propensity scores in difference-in-differences models to estimate the effects of a policy change. *Health Services & Outcomes Research Methodology*, *14*(4), 166–182. https://doi.org/10.1007/s10742-014-0123-z

Tatsuya Fukushige, D. F. (2021). *Bike-share in the Sacramento region primarily substitutes for car and walking trips and reduces vehicle miles traveled.* Berkeley, CA: University of California, Institute of Transportation Studies. https://escholarship.org/uc/item/18q404xb

U.S. Census Bureau. (2023) Longitudinal Employer-Household Dynamics (LODES 5). [dataset]. Retrieved from https://lehd.ces.census.gov/data/

Wooldridge, J. M. (2021). Two-way fixed effects, the two-way Mundlak regression, and difference-in-differences estimators. *SSRN Electronic Journal*. Retrieved from https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3906345

Zhang, J., Pan, X., Li, M., & Yu, P. S. (2016). Bicycle-sharing system analysis and trip prediction. *2016 17th IEEE International Conference on Mobile Data Management* (MDM), 174–179. https://doi.org/10.1109/MDM.2016.35

Zhong, H., Li, W., & Boarnet, M. G. (2021). A two-dimensional propensity score matching method for longitudinal quasi-experimental studies: A focus on travel behavior and the built environment. *Environment and Planning B: Urban Analytics and City Science, 48*(7), 2110–2122. https://doi.org/10.1177/2399808320982305

Zhou, Y., Yu, Y., Wang, Y., He, B., & Yang, L. (2023). Mode substitution and carbon emission impacts of electric bike sharing systems. *Sustainable Cities and Society, 89*, 104312. https://doi.org/10.1016/j.scs.2022.104312