October 2019

Exploring the Equity Performance of Bike-Sharing Systems with Disaggregated Data: A Story of Southern Tampa

Zhiwei Chen

*University of South Florida*

Follow this and additional works at: [https://scholarcommons.usf.edu/etd](https://scholarcommons.usf.edu/etd)

Part of the [Public Policy Commons](https://scholarcommons.usf.edu/etd), and the [Urban Studies and Planning Commons](https://scholarcommons.usf.edu/etd)

**Scholar Commons Citation**


This Thesis is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact scholarcommons@usf.edu.
Exploring the Equity Performance of Bike-Sharing Systems with Disaggregated Data:

A Story of Southern Tampa

by

Zhiwei Chen

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Civil Engineering
Department of Civil and Environmental Engineering
College of Engineering
University of South Florida

Major Professor: Xiaopeng Li, Ph.D.
Amy L. Stuart, Ph.D.
Yu Zhang, Ph.D.

Date of Approval:
October 28, 2019

Keywords: Equity Analysis, Shared Mobility, Individual Accessibility, Tour-based Analysis, Coast Bike Share System

Copyright © 2019, Zhiwei Chen
Dedication

I dedicate this thesis to my beloved parents and sisters.
Acknowledgments

I sincerely thank my advisor Dr. Xiaopeng Li for his guidance and assistance during my studies. I am also very thankful for my thesis committee members Dr. Amy L. Stuart, and Dr. Yu Zhang for their invaluable advices and constructive comments during the preparation of this thesis. Also, I would like to thank Mr. Yujie Guo, who has worked closely with me on the project resulting in this thesis and assisted me with GIS application and R programming. This thesis would not have been completed without their assistance.

I gratefully acknowledge support provided by the Center for Transportation, Environment, and Community Health (CTECH), which is funded by a grant from the U.S. Department of Transportation’s University Transportation Centers Program, Grant No. 69A3551747119. I would also like to express my appreciation to Dr. Sashikanth Gurram for providing the disaggregated activity and demographic data necessary for this thesis.
# Table of Contents

List of Tables ................................................................................................................................. iii

List of Figures ................................................................................................................................ iv

Abstract ............................................................................................................................................v

Chapter 1: Introduction ....................................................................................................................1
  1.1 Background ....................................................................................................................1
  1.2 Contribution Statement ..................................................................................................3
  1.3 Thesis Organization .......................................................................................................4

Chapter 2: Literature Review ...........................................................................................................5
  2.1 Review of Accessibility Modeling in Equity Analysis ..................................................5
  2.2 Review of Transportation Equity Analysis Approaches ..............................................10
  2.3 Review of Studies on Bike-sharing Equity ..................................................................13

Chapter 3: Methodology ................................................................................................................16
  3.1 Analysis Methodology .................................................................................................16
    3.1.1 Bike-sharing Accessibility for a Single Trip ................................................16
    3.1.2 Tour-based Bike-sharing Accessibility for Individuals ................................19
    3.1.3 Analysis of Horizontal Equity ..........................................................................22
    3.1.4 Analysis of Vertical Equity ..............................................................................24
  3.2 Case Study Application ................................................................................................26
    3.2.1 Bike-sharing System ........................................................................................26
    3.2.2 Study Area for Analysis of Accessibility and Equity ......................................27
    3.2.3 Data Sources ....................................................................................................28
  3.3 Verification of the Proposed Methodology ................................................................29
    3.3.1 Determining the Impacts of Improved Single Trip Accessibility ....................29
    3.3.2 Determining the Impacts of Using a Tour-based Approach .........................30

Chapter 4: Results and Discussion .................................................................................................31
  4.1 Descriptive Statistics ....................................................................................................31
  4.2 Horizontal Equity ........................................................................................................32
  4.3 Vertical Equity .............................................................................................................36
  4.4 Impacts of the Proposed Methodology ....................................................................41
    4.4.1 Impacts of “Walking-cycling-walking” for Single Trips ...............................41
    4.4.2 Impacts of Tour-based Approach ....................................................................44
    4.4.3 Impacts of Using Disaggregated Data in Equity Analysis .........................45

Chapter 5: Conclusion ....................................................................................................................46
List of Tables

Table 4.1: Distribution summary statistics of study data and accessibility results ..........................32
Table 4.2: Gini index values resulting from our approach and the benchmark approaches ..............35
Table 4.3: Summary statistics for bike-sharing accessibility by population subgroup .....................38
Table 4.4: Result from the ANOVA tests ......................................................................................39
List of Figures

Figure 3.1: The walking-cycling-walking process of a bike-sharing trip.................................17

Figure 3.2: An illustrative example for the necessity for the tour-based analysis......................19

Figure 3.3: The study area .............................................................................................................27

Figure 4.1: Lorenz curves with their Gini index values calculated at different levels ...............34

Figure 4.2: Bike-sharing accessibility using different approaches at different analysis levels ....37

Figure 4.3: Subgroup inequality index versus cumulative accessibility level..............................40
Abstract

The rising adoption of bike-sharing systems brings significant benefits to individuals and society as a whole. However, whether these benefits are distributed throughout society in a fair manner is still an open question. This study presents a methodological framework for assessing the equity performance of bike-sharing systems, with Coast Bike Share system in southern Tampa as a case study. The framework integrates three different datasets: bike-sharing infrastructure, individual travel itineraries and individual sociodemographic attribute data. With these datasets, we model individual accessibility to activity locations using bike-sharing as the mode of transportation by analyzing the “walking-cycling-walking” process of a bike-sharing trip and the trip-chaining behavior in individuals’ travel itineraries. Further, a series of equity measures are used to analyze equity performance from the perspective of both horizontal and vertical equity. To assess horizontal equity, we apply the Lorenz curve, Gini index and mapped geographic differences. For vertical equity, we compare descriptive statistics by group, perform analysis of variance, and calculate the subgroup inequality index using disaggregated data. Results reveal that accessibility to bike-sharing is not evenly distributed among individuals in the population, within the geographic space, and among different sociodemographic groups of southern Tampa. Further, our results also highlight the need to incorporate high-resolution disaggregated data and tour-based analysis to address the equity performance of bike-sharing systems.
Chapter 1: Introduction

1.1 Background

The first bike-sharing system appeared in Amsterdam in 1965, followed by a coin-deposit system in Denmark in 1991 and IT-based bike-sharing in England in 1996. The latest generation, free-floating bike-sharing, further promotes the adoption of bike-sharing systems (Shaheen et al., 2010). Hence, bike-sharing has become one of the fast-growing transportation modes throughout the world (Liu et al., 2018; Schmidt, 2018; Lei and Ouyang, 2018; Chen et al., 2019). As of the end of 2016, the number of cities that were operating a bike-sharing system had increased to around 1000 globally (Wikipedia, 2018), with China owning the largest bike fleet. In the United States, the number of shared bikes grew from 42,500 to around 100,000 from the end of 2016 to the end of 2017 (NACTO, 2019). There was also a significant increase in trips commenced with shared bikes, from less than 1 million in 2010 to almost 35 million at the end of 2017.

Along with these successes, bike-sharing systems bring substantial benefits to individuals and to society as a whole. By providing stand-alone service or solving the first/last mile problem in public transit, bike-sharing systems can reduce dependence on private automobiles and bolster public transit usage, therefore reducing fossil fuel consumption and tailpipe emissions (Zhang and Mi, 2018). As an active transportation mode, bike-sharing also induces individual physical activity, which brings positive health impacts overall (Woodcock et al., 2014). Further, bike-sharing contributes substantially to economic development in various ways such as saving travel time,

---

1 Portions of this chapter has been previously published in Guo et al. (2018) and Chen et al. (2019). Permission is included in Appendix A.
creating job opportunities, reducing household transportation expenses and improving tourism (Castro 2011).

Against the proliferation of bike-sharing systems and their positive impacts, however, more people have come to question their equity impacts, specifically, whether benefits brought by the bike-sharing system are distributed fairly among society, and particularly whether disadvantaged population groups are benefitting equally. Indeed, surveys have shown that equity impacts are a real problem in some bike-sharing systems. For example, in Washington D.C., black people accounted for around 50% of the population but only 4% of the Capital Bikeshare membership in 2016 (Benjamin, 2017). In New York, the number of hourly trips by men is almost three times more than by women (Wang and Akar, 2019). Barajas and Drive (2018) surveyed the largest 29 bike-sharing systems in the United States and found that bike-share systems disproportionately serve residential areas that are whiter, less poor and more proficient in English.

To address this issue, many operators and administrators of bike-sharing programs have initiated efforts to overcome the user barriers and address inequality (Buck, 2013; NACTO, 2019; Mcneil, et al., 2017). Meanwhile, although research funding from the Better Bike Share Partnership (PeopleForBikes, 2018) has been awarded to explore the equity question, studies remain very limited. Studies on equity analysis usually adopt statistical and econometric methods to analyze aggregated user demographic data (e.g., Qian and Niemeier, 2019; Wang and Lindsey, 2018; Mooney et al., 2019; Goodman and Cheshire, 2014). Please refer to Chapter 2 for a detailed review on relevant studies. These aggregated data-based methods might hinder our understanding of the equity impacts since individual disparities are absorbed after aggregation. Also, results from these approaches is dependent on the spatial scale for the analysis, and thus may lead to different analytical results with different spatial scales (Mooney et al., 2019). Further, existing studies have
focused on comparing user demographics rather than analyzing the distribution of benefits (e.g. accessibility) that bike-sharing systems have brought to people, which is a crucial perspective of equity analysis.

To fill these gaps, this thesis proposes a methodological framework for quantitatively accessing equity performance of bike-sharing systems using disaggregated data, with Coast Bike Share system in southern Tampa as the case study. Different from previous studies, this framework takes into account disaggregated (activity and demographic) data and individual accessibility that people obtain from bike-sharing systems, enabling analysis of the distribution of accessibility among individuals and groups in society. Transportation equity can be broadly divided into two categories: horizontal equity and vertical equity (Delbosc and Currie, 2011). In this thesis, we define equity in a similar way as Litman (2002) did. Specifically, horizontal equity requires each similar individual or group to have the same distribution of bike-sharing accessibility and to bear costs proportionate to the accessibility they receive (Litman, 2002). For vertical equity, it is more equitable if bike-sharing systems favor economically- and socially-disadvantaged groups (Pettit, 1974) and/or those with impaired mobility (Litman, 2002). The proposed framework can be applied for assessing both horizontal and vertical equity.

1.2 Contribution Statement

The contributions of this thesis are summarized as follows.

First, we propose a bike-sharing equity assessment framework where a full synthetic population of the study area, rather than a small sample or aggregated zonal-level data, was also used for the analysis. Use of a full population of disaggregated data unveils individual heterogeneity that might be absorbed by existing methods with aggregated data and thus avoids misleading our understanding of equity.
Second, to measure the benefits of bike-sharing systems, we present an individual bike-sharing accessibility model that incorporates the unique multi-modal operational characteristics of bike-sharing (i.e., walking-cycling-walking) and the trip chaining of an individual’s daily travel itineraries. The consideration of these factors avoids overestimation of bike-sharing accessibility, improving our understanding of equity impacts.

Third, the proposed methodological framework is applied to the Coast Bike Share system in southern Tampa. Results confirm the need for such a framework and also provide interesting managerial insights that can assist bike-sharing operators to implement future expansion plans that improve equity in southern Tampa and beyond.

Overall, this framework provides bike-sharing service providers and policy makers with a simple approach to analyze the equity performance of bike-sharing systems using disaggregated data. Besides, the proposed individual bike-sharing accessibility model also extends the transit accessibility modeling literature from single-modal trip-based methods for conventional transit systems to multi-modal tour-based analysis in bike-sharing systems.

1.3 Thesis Organization

The remainder of this thesis is organized as follows. Chapter 2 reviews relevant literature and explains in detail how this thesis can contribute to the existing literature. Chapter 3 presents a new approach to evaluating the equity performance of bike-sharing systems using disaggregated data and its application in southern Tampa. Chapter 4 presents and discusses the analysis results, including equity analysis, and verification of the proposed methodology. Finally, Chapter 5 provides conclusions and potential future research directions.
Chapter 2: Literature Review

This chapter provides a comprehensive review of the relevant literature. First, a review of accessibility modeling in equity analysis is presented in Section 2.1. Then, Section 2.2 presents a review of existing transportation equity analysis approaches. Finally, Section 2.3 reviews existing studies on bike-sharing equity.

2.1 Review of Accessibility Modeling in Equity Analysis

Accessibility reflects the extent to which a transportation system enables individuals to reach activities or destinations by means of transport modes or a combination thereof (Welch and Mishra, 2013). It is a fundamental element in evaluating the equity performance of a transportation system, no matter from which aspect the evaluation is being carried out. For example, if one wants to assess whether a transportation system brings equal opportunities for individuals to be employed, accessibility to jobs should be calculated. If equity in public health is being analyzed, access to health-related facilities (e.g., parks, food grocery stores, health-care facilities, community and social activities, recreation activities) should be calculated. Following is a summary of common accessibility-related cost/benefit measures that have been used in equity analysis. Note that there are other categories of accessibility measures proposed for various purposes in the literature (e.g., space-time accessibility measures, Kwan, 1998; digital accessibility, Tranos et al., 2013) but they have not been considered in transportation equity analysis and thus are out of the scope of this

---

2 Portions of this chapter has been previously published in Guo et al. (2018) and Chen et al. (2019). Permission is included in Appendix A.
thesis. Interested readers are referred to Geurs and Van Wee (2004) for a comprehensive review of general accessibility modeling.

The simplest accessibility-related cost/benefit measure used in equity analysis is Currie (2010) to identify the spatial need gap in public transportation supply in Melbourne, Australia. This measure evaluates the population of a zone’s accessibility to transportation facilities (e.g., bus stops, train stations, tram stops, etc.) by calculating the amount of transportation services the population can receive. Given a zone \( i \) with a total area of \( a_i \) and a set of transit station \( \mathcal{M}_i := [1, 2, \ldots, M_i] \), if the intersection area between the service range (or walk catchment) of a station \( m \in \mathcal{M} \) and the zone is \( a_m \) and the service level of that station is \( l_m \) (i.e., service capacity, service frequency), then the transport provision of zone \( i \) is defined as

\[
s_i = \frac{\sum_{m \in \mathcal{M}_i} a_m l_m}{a_i}, i \in \mathcal{I},
\]

where \( \mathcal{I} \) is the set of zones in the study area. This measure accounts for the spatial coverage of a transportation system taking into account its service level in a simple and intuitive manner. Thus, it is called “coverage-based measure” in the following analysis. Due to its simplicity, this measure has been applied to studies (Delbosc and Currie, 2011; Ricciardi et al., 2015) that investigated the horizontal and vertical equity of the public transport systems in Melbourne and Perth, respectively. However, several significant drawbacks exist in this measure. First, although service frequency has been used to weigh different stations, many other aspects of service quality are not considered, such as the number of lines passing through a station, vehicle capacity, running speed, land use, and so on. Thus, this simplified measure cannot capture many significant details in a transportation system, which leads to its inability to accurately reflect the quality of service of a transportation system. Second, it measures only the population’s accessibility to a transportation system (or service) in its own zone rather than describing the ability to reach activities or destinations within
the studied area. Thus, it fails to reach the ultimate goal of accessibility assessment: to determine to what extent a transportation system enables people to reach other activities or destinations.

To address the first drawback of the simplest accessibility measure, Welch and Mishra (Mishra et al., 2012; Welch and Mishra, 2013) proposed a refined measurement that focuses on capturing more details about the operations of a transportation system so that its service quality can be more accurately evaluated. Different from the previous measure that merely adopts service frequency to measure service quality, this measure defines a set of attributes $\mathcal{F}$, indexed by $f \in \mathcal{F}$ and assigns each $f \in \mathcal{F}$ a weight $w_f$. Generally speaking, these attributes can include various factors that can reflect the service quality of a transit system, such as frequency, speed, distance, capacity, required transfers, and activity density of the land around the transit station. This measure also considers that there are multiple bidirectional transit lines passing through a single station $m \in \mathcal{M}$, denoted as $\mathcal{L}_m := \{1, 2, \ldots, L_m\}$, indexed by $l \in \mathcal{L}_m$. Then, the value of attribute $f \in \mathcal{F}$ along the inbound direction of line $l \in \mathcal{L}_m$ passing through station $m \in \mathcal{M}_i$ can be denoted as $f^{in}_{lm}$.

With the above settings, the metric “connecting power” was used to describe service capacity and quality in both the inbound and outbound directions. For the inbound direction, the inbound connecting power of line $l \in \mathcal{L}_m$ passing through station $m \in \mathcal{M}$ is formulated as

$$p_{lm}^{in} := \prod_{f \in \mathcal{F}} w_f f^{in}_{lm}, \forall l \in \mathcal{L}_m, m \in \mathcal{M}_i, i \in \mathcal{I}. \quad (2)$$

The outbound connecting power of line $l \in \mathcal{L}_m$ passing through station $m \in \mathcal{M}$ is

$$p_{lm}^{out} := \prod_{f \in \mathcal{F}} w_f f^{out}_{lm}, \forall l \in \mathcal{L}_m, m \in \mathcal{M}_i, i \in \mathcal{I} \quad (3)$$

Then, the connecting power of station $m \in \mathcal{M}$ is defined as the sum of the average of the inbound and outbound connecting power of all $l \in \mathcal{L}_m$. 


\[ p_m = \sum_{l \in L_m} \frac{p_m^{out} + p_m^{in}}{2}, \forall m \in M_i, i \in J. \] (4)

A parameter representing people’s accessibility to a transit station is defined as
\[ \delta_m = \alpha e^{-\beta t_{im}}. \] (5)

where \( \alpha, \beta \) are parameters that need calibration and \( t_{im} \) is the average time for the population living in the service area of station \( m \) walking from their household to station \( m \). Note that this parameter not only captures the coverage of a station, it more accurately reflects the basic rule that people’s accessibility decreases as access time to transportation services increases. Also note that this measure still adopts the concept of coverage, so it falls into the category of coverage-based measurement. With the connecting powers of all \( m \in M_i \), the connecting power of zone \( i \) is formulated as
\[ p_i = \sum_{m \in M_i} \frac{p_m \delta_m}{|M_i| - 1}, \forall i \in J. \] (6)

The revised measure overcomes the first drawback in the simplest coverage-based measure; however, it still cannot reveal how many activities or destinations the population in a zone can access within the investigated area. Further, the coverage-based measures are built on the service radii of the transit stations, so they cannot be adapted to transportation modes without stations, especially for emerging transportation technologies such as free-floating bike-sharing, free-floating car-sharing, ride-sourcing, and so on. In light of these issues, some scholars propose reachability-based measures to identity the population of a zone \( i \in J \)’s ability to reach the activities or destinations in all other zones \( j \in J \setminus \{i\} \) within the investigated area given the monetary and (or) time budget. The basic idea of reachability-based measures is to count how many zones the population within a specific zone can reach with the given budget; the more zones one can reach, the larger its accessibility. Intuitively speaking, the accessibility between two zones
decreases as the travel cost increases. The first step to formulate a reachability-based measure is to define a function to capture the “accessibility-cost” relationship mathematically. Denote the accessibility and travel cost from zone $i \in I$ to $j \in J$ as $r_{ij}$ and $c_{ij}$, respectively, then this relationship can be generally described as

$$(r_{ij} - r_{ik})(c_{ij} - c_{ik}) \leq 0, \forall i, j, k \in J.$$  

(7)

Any functions that satisfy this property can be applied. One common example in the literature is the cumulative accessibility function (used to formulate the cumulative-opportunity measure in the literature; Kwan, 1998)

$$r_{ij} = \begin{cases} 1 & \text{if } c_{ij} \leq \bar{c}_{ij}, \forall i, j \in J, i \neq j. \\ 0 & \text{if } c_{ij} > \bar{c}_{ij} \end{cases}$$  

(8)

where $\bar{c}_{ij}$ denotes the travel cost budget of the population in zone $i \in I$. In this function, a zone is accessible to another zone if the travel cost between them is less than a pre-defined threshold (El-Geneidy et al, 2016; Golub and Martens, 2014). Another example (used to formulate the gravity-type measure; Kwan, 1998) is

$$r_{ij} = e^{-w_i c_{ij}}.$$  

(9)

where $w_i$ is a calibrated parameter determined by the origin zone $i$ (Guzman et al., 2017). Note that the travel cost here is not just limited to the travel time that has been adopted in many studies; it is actually a generalized travel cost. For example, in El-Geneidy et al. (2016) and Guzman et al. (2017), the generalized travel cost is obtained by summing the travel time and the ratio between the monetary cost and the value of time.

With this, we can formulate the accessibility of a zone $i \in J$ as the sum of its accessibility to any other zone $j \in J \setminus \{i\}$, i.e.,
\[ r_i = \sum_{j \notin \{i\}} h_j r_{ij}, \forall i \in J. \]

where \( h_j \) denotes the number of activities or destinations in zone \( j \in J \) of interest.

### 2.2 Review of Transportation Equity Analysis Approaches

Various approaches have been proposed in the literature to assess the equity performance of transportation systems. These approaches can be broadly divided into three categories. Below is a brief summary of common equity analysis approaches for transportation systems.

Mismatch analysis of simple descriptive statistics of the cost/benefit measures is a typical approach to study the equity performance of a transportation system; its history can date back to the very earliest study in measuring the performance of public transportation systems in meeting the transport needs for different demographic groups (Currie, 2004). Basically, this method presents the distribution of measures in maps or tables and then manually compares the distributions. With these maps or tables, an intuitive understanding of the equity performance of each zone or group can be obtained. A very simple approach following this idea is to map the statistical metrics of the measures in maps (usually in GIS), where each zone has one color on a scale from the lowest to the highest quintile (Kaplan et al., 2014). The statistical metrics are usually the mean values (of the measures) of a zone but, in some situations, median, maximum, minimum, and standard deviation also can be used (El-Geneidy et al., 2016). This simple mapping approach can present much macroscopic information in a very intuitive and compact way such that it has been used extensively in assessing both horizontal and vertical equity. If one wants to assess horizontal equity, the cost/benefit measures in each zone within the studied area need to be plotted (e.g., Golub and Martens, 2014). To assess vertical equity, both the population sociodemographic measure and the cost/benefit measure can be plotted in two maps and then the distributions of these two maps can be compared manually (e.g., Kaplan et al., 2014). Note that when the number of
zones or groups is not large, a map is not needed; the statistical metrics can be summarized in tables and compared directly (e.g. El-Geneidy et al., 2016). Though the mismatch analysis is quite simple and intuitive, it cannot offer quantitative information of the equity performance. Thus, quantitative analysis approaches based on exploratory statistical and econometric models and inequality index formulation have been proposed.

Popular exploratory statistical and econometric modeling include analysis of variance (ANOVA testing) and regression modeling. ANOVA testing is a statistical method to determine if there are (significant) differences between the measures of different groups. For example, El-Geneidy et al. (2016) applied ANOVA testing to compare the descriptive statistics (mean, minimum, maximum, range, and standard deviation) of accessibility measures of different population groups and found that socially vulnerable groups enjoy higher levels of accessibility in Montreal, Canada. Nevertheless, results from ANOVA testing do not offer any information on the magnitude of the disparity between different population groups and they cannot evaluate multiple sociodemographic attributes simultaneously. To address these issues, regression modeling methods can be applied to investigate possibly quantitative relationships between multiple sociodemographic attributes and transportation costs/benefits. For instance, Ogilvie and Goodman (2012) built linear regression models to analyze the inequality in usage of the bike-sharing system in London, using multiple demographic attributes (e.g., gender, place of residence) as independent variables and the mean number of trips made via bike-sharing per month as the dependent variable. The sign of the estimated coefficient of an attribute indicates whether the population with that attribute is over- or under-represented in terms of bike-sharing usage while the absolute values of the coefficient reveal how far the over- or under-representation is. Similar approaches have also been used in studying other transportation systems in other cities, e.g., Chavis et al., (2018);
Shellooe (2013). These exploratory statistical and econometric approaches are particularly appropriate for analyzing vertical equity due to its capacity of describing the relationship between transportation costs/benefits and sociodemographic attributes. Also, insights into horizontal equity can be obtained since with the sociodemographic attributes of each investigated geographic units (e.g., census block, traffic analysis zones) and the estimated regression model, one can easily infer how costs/benefits are distributed among these similar geographic units of analysis.

Although the exploratory statistical and econometric models offer profound information on the relationship between different sociodemographic attributes and transportation costs/benefits, an overall level of inequality cannot be obtained. A solution to this drawback is the application of inequality indexes that have been extensively used in social science and public health. It offers a quantitative evaluation of the degree of inequality such as income disparities among populations. Popular inequality indicators include Gini index, Atkinson index, Theil’s entropy index, etc. In transportation, the most popular method is Gini index. The Gini index traditionally has been used to evaluate the distribution of wealth or income among a population. Delbosc and Currie (2011) first applied this tool along with the Lorenz curve to analyze equity performance in transit supply in Melbourne. Due to the computational tractability and intuition of the Gini index, many studies have followed the pioneering work and used the Gini index as an overall index of the equity performance in transportation systems (e.g., Kalpan et al., 2014; Guzman et al., 2017). The Gini index can be easily adapted to various contexts and used for evaluating the horizontal equity for both the entire population and a targeted population group. Despite its wide application, the Gini index has several inherent limitations, such as the subgroup inconsistency, the lack of subgroup decomposability (Levy, 2006). To address these issues, other inequality indexes have been proposed but most of them have not been applied in transportation equity analysis. For instance,
the Atkinson index (Atkinson, 1970) was initially derived for income inequality and has been popularly applied in environmental justice analysis (Levy, 2006). Apart from being subgroup decomposable, an inequality aversion parameter is introduced in the formation of the Atkinson index, making it possible for decision makers to analyze the equity performance under various inequality aversion scenarios. Another popular indicator is the Theil’s entropy index derived from Shannon’s measure of information entropy (Theil, 1976). This index is subgroup decomposable and offers different weighting options for different subgroups but is criticized for the lack of intuition (Levy, 2006). In order to evaluate vertical equity, Stuart et al. (2009) proposed a subgroup inequality index based on a ratio of subgroup population fractions—specifically the ratio of the fraction of population that belongs to a subgroup with a given accessibility level to the fraction that subgroup comprises of the total population. The sign of the subgroup inequality index indicates if a population subgroup is over-/under-represented with respect to the entire population. Another popular indicator for assessing vertical equity is the comparative environmental risk index (CERI, Harner, 2002). The CERI is also based on a ratio of subgroup population fractions but the fractions are defined in a different way. It calculates the ratio of the fraction of population that belongs to a subgroup with a given risk level to the fraction of population that does not belong to this subgroup with a given risk level. Whether the value of the CERI is greater than 1 indicates if a population subgroup is over-/under-represented with respect to the rest of the population.

2.3 Review of Studies on Bike-sharing Equity

In the literature, three main approaches have been proposed to analyze the equity performance of bike-sharing systems. The first approach is to analyze sociodemographic profiles of registered users, e.g., Goodman and Cheshire (2014); Wang and Akar (2019), or residents in geographic units where bike-sharing infrastructure is deployed, e.g., Smith et al. (2015), Chavis et
al. (2018), McNeil et al. (2017), Mooney et al. (2019), Wang and Lindsey (2018), Qian and Niemeier (2019) through simple statistical comparison or regression analysis. The other approach compares descriptive statistics of sociodemographic data of residents within and outside bike-sharing service areas (defined as a buffer around a bike-sharing station), e.g., Hosford and Winters (2018), Ursaki and Aultman-Hall (2015). Another popular approach involves a comparison of sociodemographic characteristics of registered users with those of census residences within bike-sharing service areas, e.g., Gavin et al. (2016), Ogilvie and Goodman (2012).

In existing approaches, data are typically aggregated to a specific geospatial scale due to the lack of individual-level sociodemographic data; scale units have included census block groups (e.g., Ursaki and Aultman-Hall, 2015; Qian and Niemeier, 2019), census blocks (Wang and Lindsey, 2018), census tracts (Smith et al., 2015), neighborhoods (Mooney et al., 2019), and dissemination areas (Hosford and Winters, 2018). The use of aggregated data might hinder our understanding of the equity impacts since individual disparities are absorbed after aggregation. For example, in a census tract where white people have high accessibility to bike-sharing facilities, we cannot conclude that all white people living in this tract have high accessibility. Further, these aggregated approaches are vulnerable to the modifiable area unit problem; i.e., different levels of aggregation lead to different analytical results (Mooney et al., 2019). Therefore, with the availability of high-resolution data and advancements in modelling techniques (e.g., activity-based travel demand modeling) in recent years, some scholars have argued for the importance of introducing individual data into transportation equity analysis (Bills and Walker, 2017), but this problem has still not been well addressed in the literature.

Additionally, focusing on accessibility to bike-sharing infrastructure rather than treating bike-sharing as a mode of transport to activity locations, these approaches can essentially be
classified as the coverage-based approach in transit accessibility modeling (Currie, 2010; El-Geneidy et al., 2010; Murray, 2001). They offer a simple and intuitive metric to evaluate the structure of a transit network, but cannot capture traveler’s travel demand (Nassir et al., 2016). To address this issue, the reachability-based transit accessibility modeling approach is needed. This method considers travelers’ origin–destination pairs and models transit accessibility as a decreasing function of the travel impedances with estimated travel time (Kawabata and Shen, 2006; Liu and Zhu, 2004), time-dependent travel time (Church et al., 2005), generalized travel cost (El-Geneidy et al., 2016; Guzman et al, 2017), transit service quality (Mishra et al, 2012; Welch and Mishra, 2013), passenger choice behaviors (Nassir et al., 2016), etc. However, existing reachability-based methods for traditional transits has not considered two special challenges in bike-sharing systems. First, trips via shared bikes involve three consecutive steps, i.e., walking-cycling-walking. Studies have found that the distance someone will walk to use a shared bike is about 1,000 feet or 5 minutes walking (NACTO, 2015). Thus, the walking phase is an essential component for determining whether someone will use bike sharing. Therefore, the multimodal process must be taken into account in accessibility modeling for bike-sharing. Second, individual decisions on whether to use a shared bike for a trip is affected by its neighboring trips; i.e., an individual does not likely choose biking for the current trip if driving is the only choice for the trip followed. Thus, a reachability-based approach for modelling individual bike-sharing accessibility is needed.
Chapter 3: Methodology

This chapter proposes a new approach to evaluating the equity performance of bike-sharing systems. We first discuss the analysis methodology, which includes approaches for trip- and tour-based modeling of individual bike-sharing accessibility as well as methods to analyze both horizontal and vertical equity using disaggregated individual data. Finally, we present a case study application of the methodology.

3.1 Analysis Methodology

This section presents an innovative methodological framework for equity analysis. First, accessibility modeling using trip-based and tour-based analysis is presented. Then, statistical and econometric methods that can be applied to analyze equity performance of bike-sharing systems from both the horizontal and vertical perspectives are discussed.

3.1.1 Bike-sharing Accessibility for a Single Trip

To estimate individual accessibility to activity locations via bike-sharing, we first start with modeling accessibility for a single trip during an individual’s day. Let \( J \), indexed as \( i \in J \), be the set of all individuals in the study area, \( N_i \) be the number of trips that individual \( i \in J \) commences over a day, and \( N_i := \{1, 2, \ldots, N_i\} \), indexed as \( n \in N_i \), be the set of all consecutive trips in individual \( i \)’s daily itinerary. We define a trip, \( n \in N_i \), as one-way travel with a defined origin \( (p_{in}^-) \) and destination \( (p_{in}^+) \). Further, we denote the set of all bike-sharing facilities in the study area (bike-sharing stations for station-based systems, and designated parking spots for free-floating systems).

---

3 Portions of this chapter has been previously published in Guo et al. (2018) and Chen et al. (2019). Permission is included in Appendix A.
as $\mathcal{B} := [1, 2, \ldots B]$, where $B$ is the number of bike-sharing facilities. We index the bike-sharing facility for picking up and returning bikes as $b^- \in \mathcal{B}$ and $b^+ \in \mathcal{B}$, respectively. As shown in Figure 3.1, a bike-sharing trip is essentially comprised of three consecutive steps: (1) walking to pick up a bicycle at a bike-sharing facility $b^- \in \mathcal{B}$ near origin $p_{in}^-$, (2) cycling from $b^-$ to another bike-sharing facility $b^+ \in \mathcal{B}$ near destination $p_{in}^+$ to return the bicycle, and (3) walking from $b^+$ to $p_{in}^+$. As an improvement over existing bike-sharing equity analyses, we consider all three of these steps in our measure of accessibility, rather than assuming that only the population residing within the service area of a bike-sharing facility can access bike-sharing (Gavin et al., 2016). This allows capture of the impact of distance to a bike-sharing facility and travel distance of individual trips on accessibility.

![Figure 3.1: The walking-cycling-walking process of a bike-sharing trip](image)

As the trip is a consecutive process, the bike-sharing accessibility for a single trip, denoted as $a_{in}$, can be formulated as the product of the accessibilities of each step. For the walking steps, the individual’s distance-dependent willingness to walk between a bike-sharing facility and the trip origin or destination can be used as the measure of accessibility. An individual’s willingness
to walk to a specific bike-sharing facility is calculated with a distance decay function, $f_w(d_{p^-_in,b^-})$, where $d_{p^-_in,b^-}$ is the distance between $p^-_in$ and $b^-$. An individual’s willingness to walk to pick up a shared bike, denoted as $w_{p^-_in}$, can then be calculated as $w_{p^-_in} := \max_{b^- \in B} \{ f_w(d_{p^-_in,b^-}) \}$. Note that we use a maximum here to allow consideration of multiple facilities in the overall calculation of trip accessibility. Additionally, the facility selected for bike pick-up, denoted as $b^-_{in}$, is the one that maximizes $w_{p^-_in}$, i.e. $b^-_{in} := \arg\max_{b^- \in B} f_w(d_{p^-_in,b^-})$. Willingness to walk after returning a shared bike, $w^+_p$, and the facility selected for bike return, $b^+_in$, are calculated similarly. For the cycling process, the individual’s willingness to cycle, $c_{in}$, can also be calculated using the distance between $b^-_{in}$ and $b^+_in$, denoted as $d_{b^-_{in},b^+_in}$, with the distance decay function for cycling, denoted as $c_{in} = f_c(d_{b^-_{in},b^+_in})$. One example of the distance decay function that has been used for willingness to travel is $f(d) = \alpha_1 e^{-\alpha_2 d}$, where $\alpha_1$ and $\alpha_2$ are calibrated parameters (Hochmair, 2015) and $d$ is distance. This is the general form of the distance decay function applied here for both willingness to walk $f_w(d_{p^-_in,b^-})$, $f_w(d_{p^-_in,b^+_in})$ and willingness to cycle $f_c(d_{b^-_{in},b^+_in})$, but with parameter values calculated separately for walking versus cycling and $d$ replaced by the corresponding distance notation. For walking we used $\alpha_1 = 1.0126$, and $\alpha_2 = 0.0013$/mile (Zhao et al., 2003); for cycling, we used $\alpha_1 = 80.7908$ and $\alpha_2 = 0.0002$/mile (Hochmair, 2015), respectively. With this, we can compute the bike-sharing accessibility for a single trip as $a_{in} = w^-_{p^-_in}c_{in}w^+_p$. Note that with these distance decay functions, the willingness to walk and to cycle will range from 0 to different values of $\alpha_1$. To make them affect the resulting accessibility value comparably, we divided the willingness values obtained from these functions by their corresponding $\alpha_1$ to obtain accessibility values ranging from 0 to 1. Further, the geographic unit of analysis was the parcel (geographic units whose boundaries are defined in deed to a property, National Research Council, 2007) for
this work; hence, all individual origin and destination locations (and related distance end points) are at parcel centroids.

3.1.2 Tour-based Bike-sharing Accessibility for Individuals

Although the accessibility of single trips via bike-sharing is important, it cannot capture the interactions between trips that affect accessibility in an individual’s full daily travel itinerary. To illustrate the need for considering tours in modeling of individual bike-sharing accessibility, consider the daily itinerary shown in Figure 3.2. Nodes 1 through 5 represent activity locations of home, a convenience store, a work place, a restaurant and a shopping mall, respectively, with the sequence of trips between nodes indicated by arrows. Consider a scenario where bike-sharing facilities are located at each node with the accessibility via bike-sharing for each trip shown in the figure. If a trip-based approach is applied, the bike-sharing accessibility for trips (1, 2) and (5, 1) will be 0.8 and 0.6, respectively, which are relatively high values. However, these results may not be realistic in practice. If the accessibility to work, i.e., for trips (2, 3) and (3, 5), via bike-sharing (or public transit) is small, the traveler likely also drives for trip (1, 2) and (5, 1). Therefore, the resulting bike-sharing accessibility would be overestimated by the trip-based approach. To address this drawback, we propose a tour-based approach that considers each traveler’s trip chaining for more realistic evaluations.

Figure 3.2: An illustrative example for the necessity for the tour-based analysis
The first step of the tour-based approach is to break each individual’s daily travel itinerary, \( N_i \), into a set of subtours indexed as \( m \in M_i, \forall i \in J \). Subtours are determined using a subtour generation algorithm that begins by generating the sequence of all visited locations in individual \( i \) ’s daily itinerary, \( \mathcal{X}_i \). Subsequently, her set of activity locations \( O_i \) is generated using unique values from \( \mathcal{X}_i \). Afterwards, five variables are defined and initialized, including the individual \( i \) ’s set of tours \( M_i \), the tour counter \( m \), the current location index \( s \), the number of visits to each activity location \( E_o, \forall o \in O_i \), and the set of indices for each visit to each activity location \( F_o, o \in O_i \). With these, the algorithm iterates through the sequences \( \mathcal{X}_i \) and \( O_i \) in an outer and inner loop, respectively, to divide the travel itinerary into multiple subtours as follows. At each iteration of the outer loop, the algorithm takes the first \( x \in \mathcal{X}_i \) as the current visit location and then proceeds to the inner loop. In the inner loop, for each activity location \( o \), it first checks if the current visit location \( x \) is the same as an activity location \( o \). If yes, it increases the number of times that \( o \) has been visited \( (E_o) \) by 1 and adds the index of the visit location \( (s) \) into set \( F_o \). Next, it checks if \( o \) has been visited more than once. If yes, a subtour has been completed, so it updates the tour index \( m \) by 1, collects the set of visited locations in the tour \( (N_{im}) \) (i.e., locations in \( \mathcal{X}_i \) with indexes ranging between \( \min(F_o) \) and \( \max(F_o) \)), and adds this tour to the set of tours \( M_i \). Next, it removes all trips in \( N_{im} \) from \( \mathcal{X}_i \), updates the current visit location index \( s \) to account for the removed locations, and reinitializes \( E_o \) and \( F_o \). Specifically, these are reinitialized to as 0 and \( \emptyset \), respectively for all activity locations visited in tour \( m \) except the last location. For the last location, \( E_o \) and \( F_o \) are reinitialized as 1 and the start location index, respectively, because the next tour starts here. When a tour has been found or all \( O_i \) have been checked, the algorithm updates the index to the current visit location \( (s) \) by 1 and then moves on to the next iteration of the outer
loop. The iteration process continues until all visited locations, $\mathcal{X}_i$, are checked. Pseudocode for the subtour generation algorithm is provided in Algorithm 1.

<table>
<thead>
<tr>
<th>Algorithm 1. Subtour Generation Algorithm</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Input:</strong> $\mathcal{J}, N_i, i \in \mathcal{J}$</td>
</tr>
<tr>
<td>1. for $i \in \mathcal{J}$</td>
</tr>
<tr>
<td>2. $\mathcal{X}<em>i \leftarrow {p</em>{in}^i} \cup {p_{im}^i</td>
</tr>
<tr>
<td>3. $O_i \leftarrow \text{unique values from } \mathcal{X}_i$; // generate set of activity locations</td>
</tr>
<tr>
<td>4. $\mathcal{M}_i \leftarrow \emptyset$; // initialize set of tours as an empty set</td>
</tr>
<tr>
<td>5. $m \leftarrow 0$; // initialize tour counter</td>
</tr>
<tr>
<td>6. $s \leftarrow 0$; // initialize current location index</td>
</tr>
<tr>
<td>7. $E_o \leftarrow 0, \forall o \in O_i$; // initialize # of visits to each activity location $o$</td>
</tr>
<tr>
<td>8. $F_o \leftarrow \emptyset, \forall o \in O_i$; // initialize set of indices for each visit to $o$</td>
</tr>
<tr>
<td>9. while $\mathcal{X}_i$ is not empty</td>
</tr>
<tr>
<td>10. $s \leftarrow s + 1$; // update the current location index</td>
</tr>
<tr>
<td>11. for $o \in O_i$</td>
</tr>
<tr>
<td>12. if $x = o$; // if current location $x$ and activity location $o$ are the same</td>
</tr>
<tr>
<td>13. $E_o \leftarrow E_o + 1$; // increment the number of visits to $o$</td>
</tr>
<tr>
<td>14. $F_o \leftarrow F_o \cup s$; // add the current location index to set $F_o$</td>
</tr>
<tr>
<td>15. end if</td>
</tr>
<tr>
<td>16. if $E_o &gt; 1$; // if $o$ visited twice, a tour is complete</td>
</tr>
<tr>
<td>17. $m \leftarrow m + 1$; // increment the tour counter</td>
</tr>
<tr>
<td>18. $N_{im} \leftarrow {n</td>
</tr>
<tr>
<td>19. $\mathcal{M}_i \leftarrow \mathcal{M}<em>i \cup {N</em>{im}}$; // add tour to set of tours $\mathcal{M}_i$</td>
</tr>
<tr>
<td>20. $\mathcal{X}_i \leftarrow \mathcal{X}_i \setminus {n</td>
</tr>
<tr>
<td>21. $s \leftarrow s - (\max(F_o) - \min(F_o))$; // update the current location index to account for removed locations</td>
</tr>
<tr>
<td>22. $E_o \leftarrow 1, F_o \leftarrow \max(F_o), o @ \min(F_o)$; // reset the number of visits to the tour start activity location to 1 and its location index to that from the end of tour $m$</td>
</tr>
<tr>
<td>23. $E_o \leftarrow 0, F_o \leftarrow \emptyset, o \in {n</td>
</tr>
<tr>
<td>24. break</td>
</tr>
<tr>
<td>25. end if</td>
</tr>
<tr>
<td>26. end for</td>
</tr>
<tr>
<td>27. end while</td>
</tr>
<tr>
<td><strong>Output:</strong> $\mathcal{M}_i, \forall i \in \mathcal{J}$</td>
</tr>
</tbody>
</table>
Once the itinerary has been broken into subtours, we can compute the bike-sharing accessibility for each subtour $m \in M_i$ as the product of the accessibility of all included trips, i.e.,

$$a_{im} = \prod_{n \in N_{im}} a_{in}, \forall i \in I, m \in M_i. \quad (11)$$

The accessibility of bike-sharing for each individual $i \in I$, denoted as $a_i, \forall i \in I$, can be formulated as the average of the accessibility of all her subtours, i.e.

$$a_i = \frac{\sum_{m \in M_i} a_{im}}{|M_i|}, \forall i \in I. \quad (12)$$

Note that here we adopt the simplest approach to getting a summary statistic for the central tendency of individual accessibility, given there is no evidence in the literature for a more complicated weighting. Hence, we have a measure of individual accessibility of bike-sharing that considers the chaining of trips within subtours.

We note that in the present model development, we only consider trips by bike-share (with walking) in the calculation of the component trip accessibilities ($a_{in}$). However, this tour-based approach could easily be applied to the case where bike-sharing serves as access/egress to public transit use for longer trips. Specifically, willingness to take public transit could be incorporated as a step of a multi-modal trip travel itinerary that involved walking-transit-walking or also included cycling at either end of transit use in the calculation of each $a_{in}$. Additionally, stand-alone bike-sharing service has been shown to account for a substantial portion of bike-sharing trips in many cities (Bachand-marleau et al., 2012; Martin and Shaheen, 2014; Zhang et al., 2015; Zhang et al., 2018). Hence, the more limited approach demonstrated here remains important.

### 3.1.3 Analysis of Horizontal Equity

To evaluate horizontal equity in this study, we calculate the Lorenz curve and the Gini index, and also apply geographic mapping. The Lorenz curve is a graphical analysis tool from
economics (Lorenz, 1905) that has been applied in many fields to study inequality. Here, it provides a visual comparison of the cumulative distribution of accessibility with that of population share (ranked by accessibility). The Gini index is a quantitative measure of inequality that is derived from the Lorenz curve, and has also been applied broadly (Delbosc and Currie, 2011; Guzman et al., 2017; Kaplan et al., 2014; Lucas et al., 2016; Welch and Mishra, 2013). We use it to quantitatively assess overall inequality in the distribution of accessibility. Gini index values range from 0 to 1, with 1 indicating the most skewed distribution of the bike-sharing accessibility and 0 the most even distribution. Although there are several methods of approximation, the Gini index (G) is classically formulated as (Allison, 1978):

\[
G = \frac{1}{2I^2 \mu} \sum_{i=1}^{I} \sum_{j=1}^{I} |a_i - a_j|, 
\]

(13)

where \( I \) is the number of individuals, and \( \mu \) is the mean value of individual accessibility in the study area. We calculated the Gini index using the ineq package in R statistical software (Zeileis and Kleiber, 2015). Finally, similar to other authors (Kaplan et al., 2014), we map the spatial distribution of accessibility for improved insights into inequality in bike-sharing accessibility. For spatial mapping of bike-sharing accessibility within the studied area, we used the Kernel Density Tool in the Arcgis Toolbox; this tool has been frequently used for hotspot identification such as crash hotspots (Thakili et al., 2015). The geographic mapping offers a simple and intuitive way to explore the spatial distribution of the bike-sharing accessibility, while the Gini index provides a quantitative measure of overall inequality.

Because individual data are rarely available, evaluation of horizontal equity using the approaches outlined above is commonly accomplished by comparing benefits (or costs) among groups of the population that are defined based on geographic location (such as zip code areas).
Here, we present an approach that leverages the availability of individual-level data. Specifically, we calculate the Lorenz curve and Gini index directly from individual-level accessibility; we call this the analysis of population equity.

Additionally, for comparison to previous work, we also calculate these measures and perform spatial mapping based on individual data that have been aggregated to defined geographic areas; we call this spatial equity analysis. For this thesis, we aggregate the individual data into both parcel-level and TAZ-level data (TAZ is an acronym for traffic analysis zone, a geographic unit delineated by state and/or local transportation organizations for tabulating transportation-related data, U.S. Census Bureau, 2002). For example, we calculate the aggregate accessibility for each parcel using the individual bike-sharing accessibility \( a_i, \forall i \in I \). Since the population varies across parcels, we sum and normalize the accessibility of all individuals within a parcel. More specifically, let \( P := [1,2,\ldots,P] \), indexed as \( p \in P \), be the set of parcels in the investigated area and \( a_p \) be the aggregate accessibility of parcel \( p \in P \). Further, let \( p_i = 1 \) if individual \( i \in I \) resides in parcel \( p \in P \) and 0 otherwise. Then the accessibility in parcel \( p \in P \) can be formulated as

\[
a_p = \frac{\sum_{i \in I} a_i p_i}{\sum_{i \in I} p_i}, \forall p \in P.
\]  

(14)

Similar aggregations can be applied to larger geographic units, including TAZs. With these aggregated accessibility measures, we perform both geographic mapping analysis of spatial equity and also calculate the Lorenz Curve and Gini index from the aggregate-level data.

3.1.4 Analysis of Vertical Equity

Individual-level accessibility measures can also be used to study vertical equity. Following previous studies using aggregated data (e.g. El-Geneidy et al., 2016), we first calculate and compare descriptive statistics of accessibility by sociodemographic group. Analysis of variance (ANOVA) is used to assess the statistical significance of differences in mean accessibility between
multiple groups within a demographic category. Finally, we calculate a subgroup inequality index (after Stuart et al. (2009), Yu and Stuart (2016, 2013)) for a detailed assessment of inequality in accessibility.

The subgroup inequality index is based on a ratio of subgroup population fractions—specifically the ratio of the fraction of population that belongs to a subgroup with a given accessibility level to the fraction that subgroup comprises of the total population. Let $q$ denote an accessibility level and $r$ a population subgroup (e.g. white, male, Hispanic, or poor). Then, the subgroup inequality index can be mathematically formulated as

$$F_{rq} = \log \left( \frac{Y_{rq}}{Y_r} \right),$$

(15)

where $F_{rq}$ quantifies the degree to which members in subgroup $r$ are disproportionally distributed among the population with bike-sharing accessibility above level $q$, $Y_{rq}$ is the fraction of population that belongs to subgroup $r$ with an accessibility level above $q$, and $Y_r$ is the fraction of population that is subgroup $r$. The sign of this index indicates the direction of the disparity. A positive $F_{rq}$ indicates that members in subgroup $r$ are disproportionally more present among the population with bike-sharing accessibility above level $q$, while a negative $F_{rq}$ indicates the opposite. Finally, an index value of 0 reveals that members in subgroup $r$ are not more or less present among the population with bike-sharing accessibility above level $q$. Consider a simple example with subgroups male and female, where males constitute 50% of the population in the study area and 60% among those with accessibility above 0.6. From the above formulation, we can obtain $F_{\text{male},0.6} = \log \left( \frac{0.6}{0.5} \right) = 0.08$, indicating that males are overrepresented among the population with accessibility above 0.6, which can be easily verified from the original example.
percentages. Similarly, we can obtain $F_{\text{female}, 0.6} = \log \left( \frac{0.4}{0.5} \right) = -0.10$, indicating that females are underrepresented among the population with accessibility above 0.6.

### 3.2 Case Study Application

To demonstrate the above methodology for determining the equity performance of bike-sharing systems and to reveal planning implications, we applied the analysis to calculate accessibility for a bike-sharing system in the Tampa area. A description of the bike sharing system, study area, and other data sets used for the analysis are provided below.

#### 3.2.1 Bike-sharing System

Since its inception in late 2014, Coast Bike Share (a for-profit bike-sharing service provider) has been running an independent bike-sharing system in downtown Tampa, the central business district of Tampa (see Fig. 3.3), with the dual purposes of serving both local residents and tourists. Although it is a small system compared with counterparts in other large cities (e.g. Divvy in Chicago, Citi Bike in New York), it is the largest of the two independent systems in the county, with a total fleet size of around 300 at 42 stations (Danielson, 2014). Furthermore, locals comprise a significant portion of system users (McKenna, 2016), making it an appropriate demonstration case for analyzing residents’ accessibility to activity locations using the methodology proposed here. Additionally, it is interesting from an equity perspective because the sociodemographic profile of residents in the downtown area shows high diversity compared with that of the other system in the county. Furthermore, although reports reveal that the system has brought substantial benefits to the city, including improved accessibility to activities, mitigated traffic congestion, and reduced crowding of parking spaces (Irwin, 2017), the distribution of benefits from both a horizontal and vertical perspective are unknown. This is particularly important because equity has become a crucial factor in transportation planning in Hillsborough County (Kristine and ACIP,
2017), and more broadly throughout the US (Litman, 2002). Thus, this case study can help inform the equitable expansion of both this particular system and other bike-sharing systems more broadly.

Figure 3.3: The study area. (a) shows the Coast Bike Share system in downtown Tampa (with dots at the location of bike-share hubs), (b) shows the area used for analysis of accessibility and equity, and (c) provides the location of Tampa in its larger geographic context.

3.2.2 Study Area for Analysis of Accessibility and Equity

To study accessibility to the Coast Bike Share system and the distribution of that accessibility among the population, we focus on an analysis area that surrounds the bike-share system. The area, which is the southern part of Tampa (see Figure 3.3 (b), (c)), is 57.7 square miles in size with 167,992 people in 2017. The area is an excellent testbed for investigating the benefit distribution among sociodemographic groups because of its diversity. According to the US Census Bureau (2017), the white, black and Asian racial categories composed 59%, 10% and 4.0% of the population, respectively in 2017, while 25% of the population self-identified as Hispanic or Latino origin. 15.3% of the households in the area lived below the poverty line while 31% earned more than $100K per year. Females accounted for 48% of the total population and the age distribution consisted of 18.0% under 18 years, 68% between 18 and 64, and 11.8% over 65.
3.2.3 Data Sources

To investigate accessibility to bike sharing in the study area using a disaggregated approach, we used three distinct types of data. These include information on the bike-sharing facility locations, individual travel demand, and individual demographics. We obtained the coordinates of each bike-sharing facility from Coast Bike Share, as shown in Figure 3.3 (a). Data on individual travel itineraries and individual characteristics are described below.

To estimate travel demand in the study area, we applied daily activity and travel itineraries from a previous agent-based simulation of travel demand for the study area. Modeled data rather than survey data are used because a survey of the full population would be rather time- and money-consuming. Specifically, we used data for all individuals in the study region for a typical weekday, as simulated by Gurram et al. (2019) using the Person Day Activity and Travel Simulator (Daysim, Bradley et al, 2010). The simulated itinerary data contain parcel-level information on the locations of origins and destinations of trips, the trip sequence, and the timing of each hypothetical individual’s daily itinerary. Hypothetical individuals represent the real population, and were generated using an iterative proportional fitting approach (Beckman et al., 1996) from the Census Public Use Microdata Sample (PUMS) of the American Community Survey (US Census Bureau, 2010). (Real individual data cannot be used due to privacy concerns.) Additional details of the Daysim simulation can be found in Gurram et al. (2019). From the simulation results, we used the travel distances and the parcel-level location of origins and destinations to compute the willingness to walk and cycle for each trip, as discussed in Section 3.1.1.

The data from the Daysim simulation also provided the individuals’ residence locations and sociodemographic attributes that are needed for equity analyses. We consider both the parcel and the traffic analysis zone (TAZ) as residence locations for separate equity analyses. We
consider two scales to analyze the impacts of different aggregation levels on equity. Parcel centroid coordinate locations and TAZ centroid locations were obtained from the Daysim data and PlanHillsborough (2017), respectively. For sociodemographic attributes, we consider age (18-35, 35-45, 45-65, above 65), gender (male, female), household income level (below poverty, middle income defined as above the 2009 poverty level but with an annual household income below $75,000, and upper income with an annual household income above $75,000), race (white, black, Asian, other), and ethnicity (Hispanic and non-Hispanic), after Gurram (2019) and Gavin et al. (2016).

3.3 Verification of the Proposed Methodology

To determine the impacts of the proposed methodology on estimates of accessibility and equity, we performed two analyses comparing results from this framework to traditional methods. The approach we used for that comparison is described below.

3.3.1 Determining the Impacts of Improved Single Trip Accessibility

We first compared the proposed bike-sharing accessibility measure with two popular accessibility modeling methods from the literature: the coverage-based approach and the reachability-based approach. The purpose of this comparison was to illustrate the need to consider all three consecutive steps (i.e., walking-cycling-walking) in a bike-sharing trip. To perform this comparison, we calculated the accessibility for a single trip, \( a_{in} \), in two additional ways, as described below.

In public transit studies, the coverage-based approach simply considers how much of the population in a geographic unit can access a public transit system (Currie, 2010; Ricciardi et al., 2015), which can be regarded as the first step in a bike-sharing trip and is usually quantified as the portion of the area or population that is within the service area of the transit system. In bike-sharing
systems, however, an individual’s willingness to walk to the bike-sharing facilities should also be considered. Therefore, we defined the coverage-based comparison measure for the bike-sharing system as \( a_{in} = w_{in}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i. \)

The other modelling method, i.e., the reachability-based approach, considers the population’s ability to travel with a transportation mode (such as based on distance or time), which can be regarded as the second step in a bike-sharing trip and measured as either a binary function (El-Geneidy et al., 2016) or a distance decay function (Guzman et al., 2017). Because we used the distance decay function as our proposed measure, we also formulated the comparison reachability-based measure using the distance decay function as \( a_{in} = c_{in}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i. \)

Using these definitions of accessibility for a single trip, we calculated the overall accessibility for the population of the study area using the same approach described in Section 3.1. Distributions of accessibility and horizontal equity were then compared using the methods of Section 3.2 to determine the impact of our walking-cycling-walking approach.

3.3.2 Determining the Impacts of Using a Tour-based Approach

To illustrate the need for tour-based analysis when modeling bike-sharing accessibility for an individual, we also compared the proposed method with a method using trip-based analysis. For the trip-based analysis comparison measure, the individual bike-sharing accessibility was revised to \( a_i = \frac{\sum_{n \in \mathcal{N}_i} a_{in}}{N_i}, \forall i \in \mathcal{I}. \) All other operations remained the same as those described in Section 3.1. As with the single trip method comparison, results on accessibility and horizontal equity were compared to demonstrate the impact of the proposed approach.
Chapter 4: Results and Discussion

This chapter discusses the results of our analyses. Section 4.1 presents descriptive statistics of study data and accessibility results. Results on horizontal equity, or on the distribution of bike-sharing accessibility among individuals in the population and within the geographic space of southern Tampa, are presented in Section 4.2. Section 4.3 presents the vertical equity results on the distribution of bike-sharing accessibility among different sociodemographic groups. Finally, Section 4.4 compares results of our framework method against several benchmark measures adapted from the existing literature, to highlight the necessity and importance of the proposed methodology.

4.1 Descriptive Statistics

Table 4.1 presents the descriptive statistics of the study data and accessibility results for the population in the study area.

As seen in Table 4.1, the average number of trips and tours that individuals travel over a weekday in southern Tampa are 4.52 and 1.72, respectively. The individual data produce a mean bike-sharing accessibility of 0.0027 for the population in this area, with accessibility ranging from 0 to 0.41. The low mean bike-sharing accessibility in this area is attributable to the small service area of the Coast Bike Share system. Additionally, the 75th percentile value of bike-sharing accessibility is zero, suggesting at least 75% of the population has no bike-sharing accessibility. Using parcels as the geographic unit of analysis, the mean bike-sharing accessibility is lower than

---

4 Portions of this chapter has been previously published in Guo et al. (2018) and Chen et al. (2019). Permission is included in Appendix A.
the result from individual-level analysis, with accessibility values ranging from 0 to 0.35. In contrast, with TAZ as the geographic unit of analysis, the mean bike-sharing accessibility is higher, while the maximum value is reduced to 0.09. The 75th percentiles in both the parcel-level and TAZ-level accessibility indicate that at least 75% parcels and TAZ’s have no bike-sharing accessibility in the study area.

Table 4.1: Distribution summary statistics of study data and accessibility results

<table>
<thead>
<tr>
<th></th>
<th>Number of trips</th>
<th>Number of tours</th>
<th>Accessibility individual-level</th>
<th>parcel-level</th>
<th>TAZ-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maximum</td>
<td>19</td>
<td>7</td>
<td>0.4100</td>
<td>0.3500</td>
<td>0.0900</td>
</tr>
<tr>
<td>75th %ile</td>
<td>6</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean</td>
<td>4.52</td>
<td>1.72</td>
<td>0.0027</td>
<td>0.0017</td>
<td>0.0042</td>
</tr>
<tr>
<td>Median</td>
<td>4</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>25th %ile</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Minimum</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Sample Size</td>
<td>274371</td>
<td>274371</td>
<td>274371</td>
<td>92498</td>
<td>291</td>
</tr>
</tbody>
</table>

4.2 Horizontal Equity

The Lorenz curves of the cumulative distribution of accessibility versus that of population share are shown in Figure 4.1. Subplots provide curves based on disaggregated individual data (a), and data aggregated to the parcel level (b) and the TAZ level (c). The Gini index values are summarized in Table 4.2. Finally, Figure 4.2 (top row) provides the geographic distribution of bike-sharing accessibility among parcels and TAZs in the study area.

As can be seen from Figure 4.1 (a), the distribution of bike-sharing accessibility is highly unequal among the population in southern Tampa, with over 90% of individuals having no bike-sharing accessibility at all and around 2% of individuals enjoying 50% of the bike-sharing
accessibility. This inequality is also reflected in a large value for the Gini index of 0.964. This value is higher than those recently reported by Wang and Lindsey (2018) for the Nice Ride program in the Twin Cities, who calculated Gini index values of 0.67 for the residential population and 0.32 for job locations. This indicates that bike-sharing accessibility may be less evenly distributed among the population in southern Tampa than in the Twin Cities. The higher Gini index value here may also be attributable to the smaller service area of the Coast Bike Share system in southern Tampa. However, Wang and Lindsey (2018) used a coverage-based approach that measured the distribution of bike-sharing stations, rather than calculating individual accessibility. Hence, their result may overestimate bike-sharing accessibility. Please see Section 4.4.1 for an additional discussion on this issue. Further, compared with Gini index values from studies focusing on inequity of public transportation systems, e.g., 0.52 in Perth (Ricciadi et al., 2015), 0.68 in Melbourne (Delbosc and Currie, 2011), and 0.51 in Baltimore (Welch and Mishra, 2013), the Gini index value for the bike-sharing system in southern Tampa is also relatively high. This indicates that public transit accessibility is more evenly distributed among the population in these cities than bike-sharing accessibility is distributed among the population in southern Tampa. These public transit systems feature relatively dense networks across the cities, suggesting that the limited service area and low density in southern Tampa may contribute to the high Gini index value. From a planning perspective, this comparison suggests that integrating bike-sharing services with public transit may improve transportation equity, both in southern Tampa and more broadly. Explicit consideration of equity during the design of an individual system or integration of systems may be helpful.

Figure 4.1 (b) reveals similar results from the view of spatial equity, with parcels as the geographic unit of analysis. Specifically, we see an extremely skewed Lorenz curve and a very
similar Gini index value of 0.960. These results indicate that more than 90% of the parcels in southern Tampa do not have bike-sharing accessibility, while 50% of the bike-sharing accessibility is concentrated in around 2% of the parcels. A map of the geographic distribution of bike-sharing accessibility, Fig. 5 (top row, left column) further shows that most of southern Tampa’s parcels have no bike-sharing accessibility. Areas with bike-sharing accessibility are either within or on the edge of downtown Tampa, where bike-sharing stations are deployed.

Figure 4.1: Lorenz curves with their Gini index values calculated at different levels. (a) individual level, (b) parcel level, and (c) TAZ level.

Alternatively, using the TAZ as the unit of spatial analysis, the Lorenz curve distribution, Figure 4.1 (c), appears less skewed, with about 75% of the TAZs having no bike-sharing accessibility. The Gini index value was also found to be somewhat less unequal, at 0.854. A map of the geographic distribution of accessibility from TAZ-level analysis, Figure 4.2 (top row, right column), shows a similar overall pattern as that from the parcel-level analysis, with no accessibility except in the downtown region. However, bike-sharing accessibility values are generally lower in parcels with some accessibility, in all but the most central parcels. This is because at the edge of
downtown Tampa, more individuals without bike-sharing accessibility are taken into account when computing the mean bike-sharing accessibility for TAZ’s.

Overall, these analyses show that the benefits of the bike-sharing program are concentrated in only a small portion of the population and parcels. This result is not very surprising considering the existing bike-sharing deployment scheme. Specifically, Coast Bike share only deployed stations in downtown Tampa, leaving a large portion of the study area, and thus its population, unserved by the bike-sharing system. Additionally, bike-sharing accessibility values are relatively low throughout the study area of southern Tampa, even in the downtown area. This may be attributable to both the low density of bike-sharing stations (0.73 stations/km² in southern Tampa vs. 7 stations/km² in London (Goodman and Cheshire, 2014)) and impacts of residents’ full travel itineraries. First, previous studies (Du and Cheng, 2018) show that the distance between bike-sharing stations and traveler’s origins/destinations plays an important role in determining the adoption of a bike-sharing system. Although stations have been deployed in southern Tampa, they are not dense enough for substantial willingness of travelers to walk to these stations. Second, the majority of the residents’ daily travel demands in southern Tampa are long-distance trips that biking cannot cover. This may lead to lower values for individuals’ willingness to cycle, and thus for bike-sharing accessibility. We will provide further evidence for these two arguments in Section 4.4.

Table 4.2: Gini index values resulting from our approach and the benchmark approaches

<table>
<thead>
<tr>
<th>Approach</th>
<th>Individual-level</th>
<th>Parcel-level</th>
<th>TAZ-level</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our approach</td>
<td>0.964</td>
<td>0.960</td>
<td>0.854</td>
</tr>
<tr>
<td>Coverage-based</td>
<td>0.908</td>
<td>0.888</td>
<td>0.763</td>
</tr>
<tr>
<td>Reachability-based</td>
<td>0.620</td>
<td>0.518</td>
<td>0.206</td>
</tr>
<tr>
<td>Trip-based</td>
<td>0.883</td>
<td>0.858</td>
<td>0.748</td>
</tr>
</tbody>
</table>
4.3 Vertical Equity

This section presents the distribution of the bike-sharing accessibility among individuals in the study area based on their sociodemographic attributes. Table 4.3 presents group summary statistics for bike-sharing accessibility by attribute. Table 4.4 reports result from ANOVA testing for differences between the group means. Finally, Figure 4.3 plots the subgroup inequality index versus accessibility level.

As seen in Table 4.3, bike-sharing accessibility for each sociodemographic group follows an extremely skewed distribution, with the majority of the members having no bike-sharing accessibility (the 75th percentile value for all subgroups is 0). Also, the means for all population subgroups are small values that seem relatively similar to each other. However, ANOVA tests for differences in means tell a different story. As can be seen from Table 4.4, the p-values from all tests are less than 0.002, indicating at least 99.8% confidence in rejecting the null hypothesis that the mean bike-sharing accessibility is the same across different population subgroups within each attribute category. Thus, for all five sociodemographic attributes considered in this study, there are statistically significant differences between the population subgroups in terms of the average bike-sharing accessibility. Specifically, as shown in Table 4.3, Asians and non-Hispanics have higher (by 11.1% and 3.7%, respectively) bike-sharing accessibility than the population mean on average, while blacks, other races, and Hispanics have 11.1% less bike-sharing accessibility. In terms of gender, the average bike-sharing accessibility of males is better than the population mean by 3.7%. Among income categories, the below poverty and upper income subgroups have better average bike-sharing accessibility than the population mean by 25.9% and 14.8%, respectively. Finally, both adults aged between 18 and 35 and senior citizens over 65 enjoy higher bike-sharing accessibility than the population on average by 51.9% and 7.4%, respectively. Yet, the average
Figure 4.2: Bike-sharing accessibility using different approaches at different analysis levels. Analyses results were obtained using our approach (top row), the coverage-based approach (2nd row), the reachability-based approach (3rd row), and the trip-based approach (bottom row) at the parcel level (left column) and TAZ level (right column).
bikesharing accessibility is lower than the population mean by 18.5% for people aged between 45 and 65. Overall, these results provide a general picture of the distribution of bikesharing accessibility among different population subgroups.

Table 4.3: Summary statistics for bike-sharing accessibility by population subgroup

<table>
<thead>
<tr>
<th>Population group</th>
<th>minimum</th>
<th>25th %ile</th>
<th>median</th>
<th>mean</th>
<th>75 %ile</th>
<th>maximum</th>
<th>Std. dev.</th>
</tr>
</thead>
<tbody>
<tr>
<td>population</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0027</td>
<td>0</td>
<td>0.41</td>
<td>0.0159</td>
</tr>
<tr>
<td>race</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>white</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0027</td>
<td>0</td>
<td>0.41</td>
<td>0.0162</td>
</tr>
<tr>
<td>black</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0024</td>
<td>0</td>
<td>0.40</td>
<td>0.0147</td>
</tr>
<tr>
<td>Asian</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0030</td>
<td>0</td>
<td>0.21</td>
<td>0.0127</td>
</tr>
<tr>
<td>other</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0024</td>
<td>0</td>
<td>0.41</td>
<td>0.0150</td>
</tr>
<tr>
<td>ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0024</td>
<td>0</td>
<td>0.41</td>
<td>0.0148</td>
</tr>
<tr>
<td>Non-Hispanic</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0028</td>
<td>0</td>
<td>0.41</td>
<td>0.0162</td>
</tr>
<tr>
<td>gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>male</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0028</td>
<td>0</td>
<td>0.41</td>
<td>0.0166</td>
</tr>
<tr>
<td>female</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0026</td>
<td>0</td>
<td>0.41</td>
<td>0.0152</td>
</tr>
<tr>
<td>income</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>below poverty</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0034</td>
<td>0</td>
<td>0.41</td>
<td>0.0185</td>
</tr>
<tr>
<td>middle income</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0020</td>
<td>0</td>
<td>0.38</td>
<td>0.0132</td>
</tr>
<tr>
<td>upper income</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0031</td>
<td>0</td>
<td>0.41</td>
<td>0.0171</td>
</tr>
<tr>
<td>age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-35</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0041</td>
<td>0</td>
<td>0.41</td>
<td>0.0120</td>
</tr>
<tr>
<td>35-45</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0027</td>
<td>0</td>
<td>0.40</td>
<td>0.0161</td>
</tr>
<tr>
<td>45-65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0022</td>
<td>0</td>
<td>0.41</td>
<td>0.0147</td>
</tr>
<tr>
<td>&gt; 65</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0029</td>
<td>0</td>
<td>0.38</td>
<td>0.0167</td>
</tr>
</tbody>
</table>

Results on the subgroup inequality index, shown in Figure 4.3, unveil additional details regarding how different levels of the bike-sharing accessibility are distributed among population groups as the accessibility level changes. For instance, for the racial category (Figure 4.3 (a)), we observe disproportionally high bike-sharing accessibility for the whites (with subgroup inequality
index values greater than 0) for all but the highest accessibility levels (greater than 0.37). For Asian people, who were found to be disproportionally over-represented on average in Table 4.3, the index values indicate over-representation when accessibility is lower than 0.15, but underrepresentation at higher accessibility. Indeed, the inequality index value for Asian people reaches minus infinity when the accessibility is greater than 0.2, indicating that there are no Asian people in the study area with bike-sharing accessibility higher than 0.2. The black subgroup has disproportionally lower bike-sharing accessibility at most accessibility levels, but is over-represented at the highest accessibility levels (greater than 0.39). For the combined other races subgroup, results are particularly interesting with disproportionately low representation at accessibility less than 0.35 but very substantial overrepresentation at higher accessibilities (with the inequality index value reaching pas 1.5).

Table 4.4: Result from the ANOVA tests

<table>
<thead>
<tr>
<th>Sociodemographic attributes</th>
<th>Source</th>
<th>SS</th>
<th>Df</th>
<th>MS</th>
<th>F</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Race</td>
<td>Intergroup</td>
<td>0.0038</td>
<td>3</td>
<td>0.0013</td>
<td>5.51</td>
<td>0.0017</td>
</tr>
<tr>
<td></td>
<td>Intragroup</td>
<td>39.56</td>
<td>156424</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethnicity</td>
<td>Intergroup</td>
<td>0.0031</td>
<td>1</td>
<td>0.0031</td>
<td>12.21</td>
<td>0.0005</td>
</tr>
<tr>
<td></td>
<td>Intragroup</td>
<td>39.56</td>
<td>156426</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gender</td>
<td>Intergroup</td>
<td>0.0028</td>
<td>1</td>
<td>0.0028</td>
<td>11.17</td>
<td>0.0008</td>
</tr>
<tr>
<td></td>
<td>Intragroup</td>
<td>39.56</td>
<td>156126</td>
<td>0.0003</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income</td>
<td>Intergroup</td>
<td>0.0535</td>
<td>2</td>
<td>0.0267</td>
<td>105.7</td>
<td>&lt; 2E-16</td>
</tr>
<tr>
<td></td>
<td>Intragroup</td>
<td>39.51</td>
<td>156125</td>
<td>0.0003</td>
<td>16</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>Intergroup</td>
<td>0.0661</td>
<td>3</td>
<td>0.0220</td>
<td>76.42</td>
<td>&lt; 2E-16</td>
</tr>
<tr>
<td></td>
<td>Intragroup</td>
<td>34.40</td>
<td>119321</td>
<td>0.0003</td>
<td>16</td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.3: Subgroup inequality index versus cumulative accessibility level. (Dots at the edges of the figure represents values outside the range displayed).
For the other sociodemographic attributes we considered, the results are relatively consistent with the mean comparison results. Figure 4.3 (b) and (c) confirm that the bike-sharing accessibility is skewed towards non-Hispanic people and males at most accessibility levels, with some variability with level. Among income groups (Figure 4.3 (d)), the below-poverty group consistently has disproportionately high bike-sharing accessibility, while the middle-income group has disproportionately low accessibility. Interestingly these disparities grow as accessibility increases. Finally, Figure 4.3 (e) shows that people aged 18–35 yrs are over-represented while those aged between 45 and 65 are under-represented at most accessibility levels. However, the values of the index fluctuate dramatically with accessibility level, particularly at the highest levels.

4.4 Impacts of the Proposed Methodology

To highlight the importance of the proposed methodology, below we present and discuss the results from our framework method and several benchmark methods. The result of geographic mapping analysis from the three benchmark methods (i.e., coverage-based, reachability-based, and trip-based) for calculating bike-sharing accessibility are shown in Figure 5.2. The corresponding Gini index values are summarized in Table 4.2.

4.4.1 Impacts of “Walking-cycling-walking” for Single Trips

Figure 5.2 (2nd row) shows that with the coverage-based approach, a large portion of the geographic areas are estimated to have higher bike-sharing accessibility values compared with our approach (especially for areas shaded with red). This is because, the coverage-based approach only considers how an individual can access bike-sharing facilities at their origins. Thus, many long-distance trips that cannot be completed with a bicycle are calculated as having high accessibility. Also, trips whose destinations have no bike-sharing facility (which means that the shared bike cannot be returned and thus travel would not occur by bike) are calculated as having high
accessibility. This suggests that the coverage-based approach may lead to substantial overestimation of the bike-sharing accessibility. A direct consequence of the overestimation is a slightly larger proportion of the geographic area (and population) estimated as enjoying bike-sharing accessibility in southern Tampa. Specifically, more spatial units (parcels or TAZs) are shaded with non-blue colors (Figure 5.2, 2nd row) and lower Gini index values are obtained for the individual-, parcel- and TAZ-level analyses (Table 4.2). These results suggest that only considering the walking phase at the origin might contribute to an underestimation of the concentration of the bike-sharing accessibility. Indeed, this is the limitation of coverage-based accessibility models: i.e. they reflect individual accessibility to transportation facilities but cannot describe how well transportation systems fulfill the expected travel demand (Nassir et al., 2016). Therefore, the coverage-based approach is appropriate for assessing the distribution of bike-sharing facilities but not the performance of bike-sharing systems. That said, the cycling phase and walking phase at destinations should be considered when transportation planners assess the accessibility to activity locations via bike-sharing; otherwise overestimation may happen. Note that this result also reflects that the coverage of the Coast Bike system is not very high in southern Tampa because the accessibility of most areas remains less than 0.01 even if only the walking phase at the origin is considered. Hence it indicates a low density of bike-sharing stations and confirms our argument in Section 4.2 that the low density of bike-sharing facilities is partially responsible for the overall low bike-sharing accessibility in the study area. To further improve the bike-sharing accessibility in southern Tampa, a well-designed expansion of the Coast Bike Share system is needed.

From Figure 5.2 (3rd row), we can see that the reachability-based approach results in a completely different distribution of bike-sharing accessibility in southern Tampa. With this
approach, only a small portion of geographic units (less than 5% for parcels and almost 0% for TAZs) were estimated to have no bike-sharing accessibility. Further, geographic units with bike-sharing accessibility are estimated as scattered across much of the studied area, with downtown Tampa and areas to the west of the downtown estimated as enjoying relatively high bike-sharing accessibility. As a result, the inequality index values are all substantially reduced (e.g., 0.206 for the TAZ level), suggesting that the distribution of the bike-sharing accessibility is not as concentrated as that shown by our approach. In essence, the reachability-based approach measures individuals’ willingness to commence trips with shared bikes. However, without considering the walking phases, many short-distance trips that cannot be served by the bike-sharing system due to the unavailability of bike-sharing facilities within walking distance are estimated to have relatively high bike-sharing accessibility. For instance, parcels near the middle-west edge of the study area are estimated as having similar accessibility as parcels in downtown Tampa (Figure 5.2, 3rd row, left column). Thus, only considering individuals’ willingness to cycle may result in a misleading conclusion about the distribution of bike-sharing accessibility. Indeed, the reachability-based approach works when bike-sharing facilities are densely sited so that population in each geographic unit of analysis can access bike-sharing facilities. Note that even in this case the estimation results are biased because the approach ignores differences in individual access to bike-sharing facilities due to different distributions of these facilities in each geographic unit of analysis. Otherwise, the walking phases should be taken into account when transportation planners assess the accessibility to activity locations via bike-sharing. Besides, the low values of bike-sharing accessibility obtained from this approach (between 0.04 and 0.27 for most parcels) imply that only a small portion of the individuals’ daily travel is comprised of short-distance trips, so their willingness to cycle is relatively low. This observation supports our statement in Section 4.2 that
the dominance of long-distance trips is partially responsible for the low bike-sharing accessibility in southern Tampa. These results also reveal planning implications on designing future expansion plan for Coast Bike Share system. Specifically, high priority for station siting should not be given to spots with intensive long-distance travel demands (unless access to public transit is available there); bike-sharing accessibility will be low regardless of the number of bike-sharing facilities. This is very important for areas like southern Tampa where long-distance trips are dominant and thus methods are needed to properly predict short-distance trips.

4.4.2 Impacts of Tour-based Approach

As we can see from Figure 5.2 (bottom row), the trip-based analysis also generates a similar distribution of the bike-sharing accessibility as the tour-based analysis does. As expected, this approach overestimates bike-sharing accessibility because it does not consider the interdependence between the mode choices of several consecutive trips (i.e. a tour). This way, short-distance trips that are not likely to be served by the Coast Bike share system (because of the existence of long-distance trips in the trip chain) are estimated as having high bike-sharing accessibility. Similar to what was observed for the coverage-based approach, a natural consequence of such overestimation is more areas estimated to have relatively high bike-sharing accessibility and lower Gini index values (see Table 4.2). Thus, access to bike-sharing is estimated to be more evenly distributed among both the population and geographic units. This observation implies the importance of tour-based analysis in designing future expansion plans for the Coast Bike Share system in southern Tampa and beyond. While guidelines for planning bike-sharing systems are available (NACTO, 2015), they simply suggest placing bike-sharing stations at places where heavy pedestrian or visitor flows are present. Yet, analysis results from the tour-based approach reveal that all short-distance trips may not be counted as demand for bike-sharing systems due to the interrelated
decisions of consecutive trips in a tour. Thus, it would be a waste of investment to site bike-sharing stations at locations with many short-distance trips, but where shared bikes are not actually needed to commence these trips because of the following long-distance trips not served by public transit. To address this issue, the proposed tour-based approach can be applied to filter short-distance trips that are not likely to be served by shared-bikes and thus assist in developing a better sitting plan for future bike-sharing stations.

4.4.3 Impacts of Using Disaggregated Data in Equity Analysis

Finally, we want to note the importance of incorporating disaggregated data for equity analysis. As discussed above, one of the benefits of using disaggregated data is that they enable us to analyze inequity in the population considering individual-level data. Our results for horizontal inequality show that such a seemingly simple methodological change is non-trivial. We can see from Figure 5.2 that as the geographic unit of analysis increases in size, the estimated bike-sharing accessibility decreases. Further, Table 4.2 shows that the estimated Gini index also decreases as the unit of analysis increases from the individual to the TAZ. These observations indicate that data aggregation tends to absorb the disparities in bike-sharing accessibility between individuals. Moreover, the higher the aggregation level (i.e., the larger the unit of analysis), the more the disparity is hidden. As a result, the distribution of accessibility seems to be less uneven, which can mislead our understanding of horizontal inequality. This suggests that it is better to use disaggregated data, when available, for equity analysis.
Chapter 5: Conclusion

This thesis closes a gap in the research literature by describing a comprehensive equity assessment framework that uses disaggregated data to analyze the distribution within the population of accessibility to a bike-sharing system. With datasets containing information on individual travel demand and bike-sharing provision, the framework first models accessibility to bike-sharing for individuals by accounting for the walking-cycling-walking process in a bike-sharing trip and the trip-chaining behavior in an individual’s travel itinerary. Then, we combine the resulting individual accessibility indicators and a sociodemographic dataset to perform a series of inequity analyses from both the horizontal and vertical perspective. In addition to applying traditional analysis methods including geographic mapping, the Gini index, comparison of distribution means, and ANOVA testing, a subgroup inequality index is applied to measure vertical equity quantitatively. To demonstrate the approach, the proposed methodological framework was applied to the Coast Bike Share system in southern Tampa. The main findings are:

- From the horizontal perspective, the distribution of bike-sharing accessibility was found to be very unevenly distributed among both the population and the geographic space in southern Tampa, with Gini index values higher than 0.95. Geographic mapping revealed that the accessibility is concentrated in areas within and around downtown Tampa.

- From the vertical perspective, bike-sharing accessibility was also found to be

---

5 Portions of this chapter has been previously published in Guo et al. (2018) and Chen et al. (2019). Permission is included in Appendix A.
unevenly distributed between different sociodemographic groups. Overall, bike-sharing accessibility is higher for whites, Asians, non-Hispanic people, males, individuals from lower- and upper-income households, and people aged between 18 to 35 and over 65. However, disparities change substantially with accessibility level for some groups, including those categorized by race, income level and age.

- Bike-sharing accessibility in southern Tampa was found to be relatively low due to the low density of bike-sharing stations and the large portion of long-distance daily travels. By considering the “walking-cycling-walking” process in a bike-sharing trip and the trip chaining of individuals’ travel itineraries, the proposed method avoids overestimation of bike-sharing accessibility. This finding demonstrates the necessity and importance of the proposed tour-based modeling approach.

- Disaggregated data enable us to analyze horizontal and vertical equity at the individual level, which unveils many important messages that might be absorbed with existing methods that use aggregated data. Indeed, aggregated data (e.g., means) may dilute the disparities among individuals, which might mislead our understanding of inequity from both the horizontal and vertical perspectives. Thus, it is helpful to incorporate disaggregated data into transportation equity analysis.

Finally, we want to note several avenues through which this work can be extended. First, this study considers bike-sharing as a stand-alone service, but serving as an access/egress mode for public transportation is also important for bike-sharing systems. By also including availability of public transit and willingness to take transit in the calculation of individual trip accessibility, the proposed tour-based analysis approach could be used to improve understanding of how bike-sharing systems interact with other transportation modes in a multimodal transportation system.
This would paint a more complete picture of the transportation equity in a city. Additionally, as transportation is increasingly regarded as a service in modern society, extending the bike-sharing accessibility measure proposed in this thesis by taking into account factors that might affect the service quality of the bike-sharing systems, such as the number of bikes at stations and repositioning activities, is necessary. Another interesting avenue for future work would be to apply the proposed method to bike-sharing systems in other cities with different sizes and/or structures, which would offer an opportunity to obtain a better understanding of the interaction between system size, structure, and equity performance. Finally, it would be interesting to assess the accessibility in concert with other environmental and public health benefits of the system as a whole.
References


PeopleForBikes. (2018). Better bike share partnership grants more than $410,000 for research and more equitable bike shares. Retrieved from https://pfpitches.com/peopleforbikes/bike-share-grants/


Schmidt, C. (2018). Active Travel for All? The Surge in Public Bike-Sharing Programs. Environmental Health Perspectives, 126(8), 082001. https://doi.org/10.1289/EHP3754


Appendix A: Copyright Permissions

The permission from ELSEVIER for material in Chapters 1, 2, 3, 4, and 5 is below.

Copyright

Describes the rights related to the publication and distribution of research. It governs how authors (as well as their employers or funders), publishers and the wider general public can use, publish and distribute articles or books.

Journal author rights

In order for Elsevier to publish and disseminate research articles, we need publishing rights. This is determined by a publishing agreement between the author and Elsevier. This agreement deals with the transfer or license of the copyright to Elsevier and authors retain significant rights to use and share their own published articles. Elsevier supports the need for authors to share, disseminate and maximize the impact of their research and these rights, in Elsevier proprietary journals* are defined below:

<table>
<thead>
<tr>
<th>For subscription articles</th>
<th>For open access articles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Authors transfer copyright to the publisher as part of a journal publishing agreement, but have the right to:</td>
<td>Authors sign an exclusive license agreement, where authors have copyright but license exclusive rights in their article to the publisher**. In this case authors have the right to:</td>
</tr>
<tr>
<td>• Share their article for Personal Use, Internal Institutional Use and Scholarly Sharing purposes, with a DOI link to the version of record on ScienceDirect (and with the Creative Commons CC-BY-NC-ND license for author manuscript versions)</td>
<td>• Share their article in the same ways permitted to third parties under the relevant user license (together with Personal Use rights) so long as it contains a CrossMark logo, the end user license, and a DOI link to the version of record on ScienceDirect.</td>
</tr>
<tr>
<td>• Retain patent, trademark and other intellectual property rights (including research data).</td>
<td>• Retain patent, trademark and other intellectual property rights (including research data).</td>
</tr>
<tr>
<td>• Proper attribution and credit for the published work.</td>
<td></td>
</tr>
</tbody>
</table>

---

* Elsevier proprietary journals include but are not limited to: ScienceDirect, Nature, and The Lancet.

** If the article is licensed under a Creative Commons license, authors retain some rights and do not transfer copyright to Elsevier.
Quick definitions

Personal use

Authors can use their articles, in full or in part, for a wide range of scholarly, non-commercial purposes as outlined below:

- Use by an author in the author's classroom teaching (including distribution of copies, paper or electronic)
- Distribution of copies (including through e-mail) to known research colleagues for their personal use (but not for Commercial Use)
- Inclusion in a thesis or dissertation (provided that this is not to be published commercially)
- Use in a subsequent compilation of the author’s works
- Extending the Article to book-length form
- Preparation of other derivative works (but not for Commercial Use)
- Otherwise using or re-using portions or excerpts in other works

These rights apply for all Elsevier authors who publish their article as either a subscription article or an open access article. In all cases we require that all Elsevier authors always include a full acknowledgement and, if appropriate, a link to the final published version hosted on Science Direct.
Appendix B: Disclaimer

The contents of this thesis reflect the views of the author, who is responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The thesis is funded, partially or entirely, by a grant from the U.S. Department of Transportation’s University Transportation Centers Program. However, the U.S. Government assumes no liability for the contents or use thereof.