Measuring Impact of Emerging Transportation Technologies on Community Equity in Economy, Environment and Public Health

or

Equity Assessment for Emerging Transportation Technologies: A Comprehensive Literature Review and Case Study

Center for Transportation, Environment, and Community Health
Final Report

by
Yujie Guo, Zhiwei Chen, Amy Stuart, Xiaopeng Li, Yu Zhang

October 23, 2018

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated in the interest of information exchange. The report 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.

1. Report No.	2.Government Accession No.	3. Recipient's Catalog No.	
4. Title and Subtitle Equity assessment for emerging transportation technologies: A comprehensive literature review and case study		5. Report Date 9/30/2018	
		6. Performing Organization Code	
7. Author(s) Yujie Guo Zhiwei Chen Amy Stuart Xiaopeng Li Yu Zhang		8. Performing Organization Report No.	
9. Performing Organization Name and Address Department of Civil and Environmental Engineering University of South Florida Tampa, FL 33620		10. Work Unit No.	
		11. Contract or Grant No. 69A3551747119	
12. Sponsoring Agency Name and Address U.S. Department of Transportation 1200 New Jersey Avenue, SE Washington, DC 20590 Department of Civil and Environmental Engineering University of South Florida Tampa, FL 33620		13. Type of Report and Period Covered Final Report 12/01/2017 – 8/31/2018	
		14. Sponsoring Agency Code US-DOT USF	
15. Supplementary Notes			

16. Abstract
Emerging transportation technologies (e.g. connected vehicles) and services (e.g. shared mobility) provide efficient, sustainable and cost-effective alternatives to traditional travel modes. However, whether these innovative technologies bring benefits to different population groups in an equal and reasonable manner is still an open question. This report aims to tackle this question and is divided into the following two parts: transportation equity literature review and a case study on bike-sharing systems. The first part of the report comprehensively surveys the literature about methodologies for analyzing transportation equity for traditional and emerging transportation technologies in terms of economy, environment and public health. It is found that existing methodologies can be unified into a 3-step equity analysis framework. Research gaps and future research directions are also discussed. The second part of the report closes one of the research gaps mentioned in the first part of report by using disaggregated data for equity measurement. This part of report develops a comprehensive equity assessment framework on bike sharing accessibility in southern Tampa with individual-level data. The report compares the equity outcomes of the proposed approach and serveral benchmarks, and interpretes results of horizontal equity and vertical equity analysis. The results justify the importance of using disaggregated tour data, also reveal some equity issues in southern Tampa.

17. Key Words Equity assessment framework, Emerging transportation technologies, Economy, Environment, Public health, Bike sharing, Disaggregated data, Accessibility		18. Distribution Statement Public Access		
19. Security Classif (of this report) Unclassified	20. Security Classif. (of this page) Unclassified		21. No of Pages	22. Price

Form DOT F 1700.7 (8-69)

Part 1

Equity assessment for economic, environmental, and public health outcomes of transportation: From conventional to emerging technologies

1   Introduction

In recent years, various innovative transportation technologies (e.g., autonomous, electric, and connected vehicles) and services (e.g., bike-sharing, car-sharing) have emerged as alternatives to traditional travel modes and are becoming increasingly popular all over the world. For example, car-sharing companies such as Zipcar operate in 500+ cities around the world, with 1 million memberships in urban areas in 2016 (https://www.zipcar.com/press/releases/millionmembers). In addition, in 2015, more than 800 bike-share programs operated around the world. Hangzhou, with the largest bike share program, has 78,000 bikes and 3,131 stations (Source: https://uli.org/wp-content/uploads/ULI-Documents/Bicycle-Sharing.pdf). These innovative services bring significant economic, environmental, and health benefits to society, including improved transportation services (Li, Ma, Cui, Ghiasi, & Zhou, 2016), reduced emissions of pollutants (Pal & Zhang, 2017; H. Yu & Stuart, 2017), and induced physical activities (Woodcock, Tainio, Cheshire, O'Brien, & Goodman, 2014). For example, 1 car-sharing vehicle removed 9–13 private vehicles from roadways in North America in 2008 (Martin, Shaheen, & Lidicker, 2010). Buck (Buck, 2013) summarized the benefits of bike-sharing, including increased overall physical activity, increased accessibility to public transit by solving the first/last mile problem, and reduced greenhouse gas (GHG) emissions through mode shifting from vehicles. However, the majority of these new transportation systems are designed to follow an efficiency-oriented paradigm that aims to maximize overall efficiency; more specifically, they are designed to pursue objectives such as saving total travel cost (Deng & Cardin, 2018; Li et al., 2016; Ma, Li, Zhou, & Hao, 2017), reducing total travel time (Levin, 2017), conserving energy consumption (Ma, Li, Zhou, Hu, & Park, 2017), decreasing GHG emissions (Lee & Madanat, 2017), etc. While such guidelines ensure that the resulting systems can bring benefits to the society as a whole, they also naturally raise the question of whether these systems benefit different demographic/socioeconomic groups equally or in a fair and reasonable manner, also known as the equity issue in urban transportation planning.

Conceptually, "equity" refers to the fairness of distribution of impacts among populations (Litman, 2002) and is a multi-disciplinary term widely used in various fields such as economics, environmental science, and politics; thus, its specific definition varies among different contexts. For example, environmental equity is defined as "a public policy goal of ensuring that the adverse human health or environmental effects of government activities do not fall disproportionately upon minority populations or low-income populations" (Forkenbrock & Schweitzer, 1999; Griffin & Sener, 2015). In transportation, the equity issue first emerged with the Civil Rights Act of 1964, which requires federal agencies to distribute federal resources in the fairest and least discriminatory manner (Welch & Mishra, 2013). According to (Litman, 2002), transportation equity can be divided into three categories: horizontal equity, vertical equity with regard to income and social class, and vertical equity with regard to mobility need and ability. Horizontal equity is the most popular measurement in the literature, requiring that each individual or group be treated with the same distribution of costs or benefits and should bear costs proportionate to the benefits they receive (Litman, 2002). For vertical equity with regard to income and social class, transportation policies are more equitable if they favors economically- and socially-disadvantaged groups (Pettit, 1974). Vertical equity with regard to mobility need and ability is concerned with ensuring that the needs of individuals or groups with impaired mobility are met (Litman, 2002).

Based on these definitions, a number of studies have been conducted to assess the distribution of benefits in various traditional transportation systems, ranging from unimodal systems (e.g., automobile, bus) to multimodal systems (e.g., transit network consisting of buses, trains, trams) in terms of economy, environment, and public health. Despite these studies, to date there is no established standard to assess the equity performance of traditional transportation systems, including those involving emerging transportation technologies and services. A better understanding of past attempts in traditional transportation systems can lay a solid foundation for proposing an appropriate equity assessment methodology for emerging transportation systems.

This paper reviews state-of-the-art equity assessment methodologies for traditional transportation systems and identifies the existing challenges in developing such methodologies for emerging transportation technologies. A comprehensive survey of the literature concerning equity assessment of the benefit distributions of traditional transportation systems in terms of economy, environment, and public health was conducted, leading to both basic definitions and a taxonomy of the current methodologies in developing the population measurement, cost/benefit measurement, and equity assessment approaches. This review is different from others in the following ways. First, it offers a comprehensive review of the equity assessment methodologies in a transportation system from an integrated perspective considering economy,

environment, and public health. Existing works usually study the equity performance of a transportation system from these aspects separately, but very few have offered an integrated view. Second, a new taxonomy system is developed based on a three-step assessment framework that was generally adopted in existing studies. This new system summarizes the measurements that can be applied in each step when evaluating the equity performance of a transportation system, allowing transportation planning agencies to find the measurements they desire more efficiently. Third, how these methodologies can be adapted to emerging transportation technologies and existing challenges to such adaptations is discussed.

The remainder of this paper is organized as follows: Section 2 reviews the general procedure of transportation equity analysis, and Section 3 reviews population measurements. Section 4 presents a review of cost/benefit measurements in terms of accessibility, environment, and safety, and Section 5 reviews the equity assessment approaches. Section 6 summarizes current literature that measures equity of emerging transportation technologies and discusses the application of an equity analysis framework to emerging transportation systems. Section 7 concludes the paper by summarizing research findings and research gaps to evaluate the equity of emerging transportation systems.

2   Transportation Equity Analysis Procedure

As noted, there is no established standard to assess the equity performance of traditional transportation systems. Despite diverse assessment methodologies, the majority of studies concerning transportation equity analysis follows a general three-step framework, as shown in Figure 1.

- **Population Measurement** – defines the population characteristics (e.g., geospatial distribution, race, income, education) and sets reference against which population groups are compared with specified cost/benefit measure.
- **Cost/Benefit Measurement** – quantifies the costs/benefits of relevant topics (e.g., accessibility, traffic emission exposure, traffic accidents) for the population and subgroups defined by the population measure.
- **Inequality Measurement** – compares costs/benefits among the population and its subgroups. Conclusions about whether a system's costs/benefits distributions are fair, (equity assessment) are based on comparison of costs/benefits among the population.
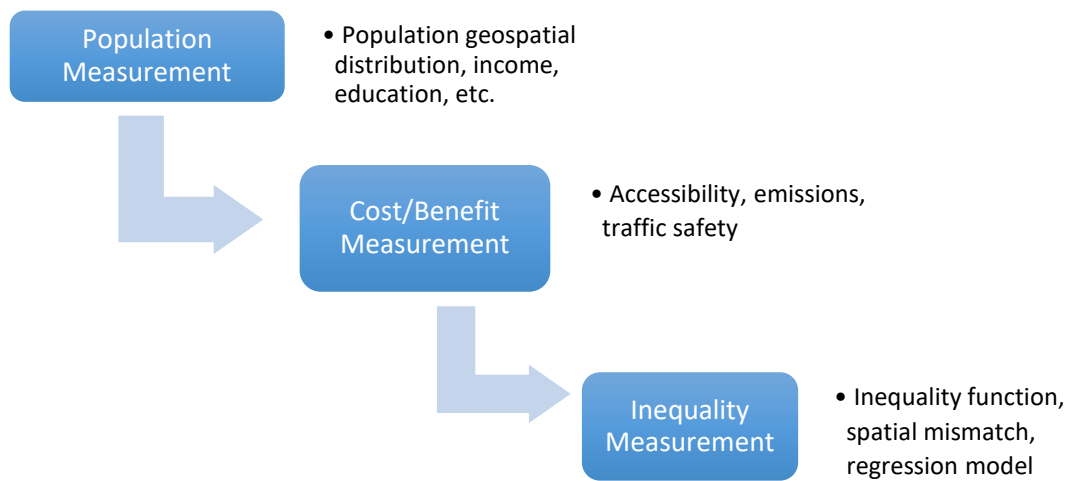
Figure 1: Three-step framework of equity assessment

Population measurement defines the research scope of population groups or individuals for whom the cost/benefit is measured and compared. The distribution of benefits or risks can be compared for populations between delineated unit areas such as census areas or Traffic Analysis Zone (TAZ); it also can be compared between populations with different social status. Cost/benefit measurement calculates the topic of interest (e.g., accessibility, traffic emission) for specified population groups defined by population such as transportation accessibility index for populations between different census tract areas or traffic exposure for different income groups. Inequality can be evaluated by visualization (e.g., using GIS mapping), inequality equations (e.g., Gini index) and regression modeling (multivariate regression).

It is worth noting that population and cost/benefit measures must have a consistent scope of geospatial data resolution. Often, population data are available only at a standard geography level such as census tract depending on the sophistication level of the cost/benefit measures. For example, human exposure to traffic pollution can be estimated at the individual level, which requires the cost/benefit measure results to be aggregated to population data resolution so inequality can be measured.

3   Population Measurement

Population measurement defines the characteristics of human population in which the cost/benefit measure is being compared, basically answering the question "equity for whom?" Population can be populations that are distributed spatially or individually if horizontal equity is measured and also can be different socioeconomic groups or people with different mobility needs if vertical equity is measured.

7

Horizontal equity evaluates the distribution of benefits and costs among each individual or among spatially-distributed population groups. In transportation studies, the equity measure of the individual level is usually impractical due to the number of populations, the complexity of human travel behavior in a multimodal transportation system, and the lack of individual-level demographic data due to privacy issues. Thus, equity is usually evaluated using data on the residential population of standard geographic entities such as TAZ and census tract. A common source of population data is the U.S. Census Bureau, which has standard hierarchy of geographic entities for the provision of population data for residents, including census block, block groups, census tracts, etc., as shown in Figure 1. A census block is the smallest geographic unit and serves as a valuable source of data for small-area geographic studies, but some blocks may not contain any population. A block group is the cluster of census blocks that contains 600–3,000 people. Census tracts are relative permanent geographic entities within counties, generally have 2,500–8,000 residents, and are designed to be as homogeneous as possible in terms of residential population characteristics, economic status, and living conditions. Beyond this standard geographic hierarchy, the Census Bureau developed other geographic entities to support specific data uses, such as a TAZ, which is most commonly used for travel demand models in transportation planning process (Source: https://www.census.gov).
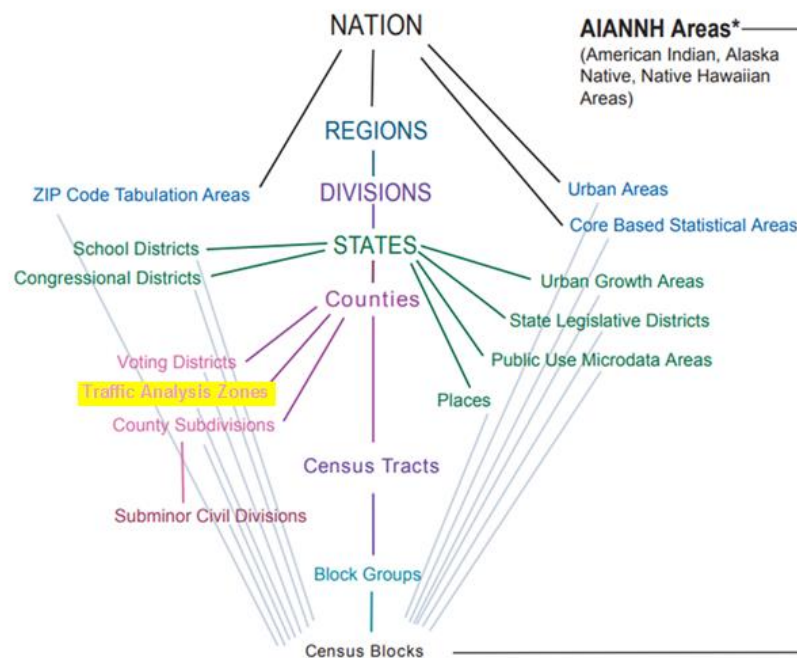

Figure 2: Standard hierarchy of U.S. Census Bureau geographic entities

(https://www.census.gov/mso/www/training/pdf/GEO_Webinar_3-13-13.pdf)

Vertical equity compares distributions of risks and benefits between populations with different demographic and socioeconomic status. Common measures of population characteristics used in equity studies include race, ethnicity, and income and education levels. Race and ethnicity are similar classifications that group people by common ancestry and physical characteristics. Groups often considered include Black/African American, White, and Hispanic (Tian, Xue, & Barzyk, 2013), and more detailed division also is considered that includes more Census Bureau categories, including American Indian or Alaskan Native, Asian or Pacific Islander, Multiracial, etc. (Stuart & Zeager, 2011). Income is usually defined by median household income (Buzzelli & Jerrett, 2007; Kravetz & Noland, 2012; Sider, Hatzopoulou, Eluru, Goulet-Langlois, & Manaugh, 2015; Tian et al., 2013) or average household income quintiles (Morency, Gauvin, Plante, Fournier, & Morency, 2012). For education level, the percentage of low education (adults with less than high school education) (Buzzelli & Jerrett, 2007), (Tian et al., 2013) are used to represent low education groups, and subgroups of Less than High School, High School, Some College, and College Graduate are used to capture more details (Harper, Charters, & Strumpf, 2015). Age categories that distinguish children and older adult populations have been used in some studies (Gurram, Stuart, & Pinjari, 2015). In addition to these commonly-used population characteristics, some researchers also consider groups that are explicitly based on disadvantaged status, such as unemployment rate (Buzzelli & Jerrett, 2007; Sider et al., 2015), deprivation index (also termed social disadvantage indicator) (Havard, Deguen, Zmirou-Navier, Schillinger, & Bard, 2009; Sider et al., 2015), and percentage of car ownership. Deprivation index is a measure of cumulative disadvantage that integrates various socioeconomic factors such as average household income, percentage of car ownership, unemployment rate, and ethnicity. Some studies found that the factors included in the deprivation index should be tailored to country-specific conditions (Sánchez-Cantalejo, Ocana-Riola, & Fernández-Ajuria, 2008). The motivation for using deprivation index is that the combination of socioeconomic factors includes both material and social elements that are more representative of a population's disadvantage (Sider et al., 2015).

To consider vertical equity based on mobility need and ability, Currie (Currie, 2010) defined a set of need indicators $\mathcal{E}$, indexed by $e \in \mathcal{E}$ and the weights of an attribute $e \in \mathcal{E}$ as $w_e, \forall e \in \mathcal{E}$. Possible elements in set $\mathcal{E}$ include adults without cars, persons over age 90, persons on a disability pension, and low-income households, to name a few. With this, the transport need index d_i for zone i is formulated as the weighted sum of all indicators in $\mathcal{E}$, i.e.,

$$d_i := \sum_{\forall e \in \mathcal{E}} e w_e, i \in \mathcal{I},$$

Previous studies of transportation equity have measured these population characteristics data at different scales. Measures of residential population characteristics for a census tract (Boyce, Zwickl, & Ash, 2016; Levy, Greco, Melly, & Mukhi, 2009; Tian et al., 2013), census block (Havard et al., 2009; Kravetz & Noland, 2012), TAZ (Jang, An, Yi, & Lee, 2017; Mortazavi & Akbarzadeh, 2017) and ZIP code (Goodman, Wilkinson, Stafford, & Tonne, 2011) are commonly used in transportation engineering studies. Additional units of analysis for population data have been used in other studies of the equity effects of transportation, including enrollment data for elementary schools (Stuart & Zeager, 2011) and, most recently, individual-level demographic data (Gurram, 2017; Gurram et al., 2015; Gurram, Stuart, & Pinjari, 2018). However, very few studies explored the suitability of scale, which is likely to depend on the cost/benefits measure of interest. Gurram et al. (Gurram et al., 2015; Stuart, Mudhasakul, & Sriwatanapongse, 2009; H. Yu & Stuart, 2013, 2016) analyzed disparities in exposures to measures of traffic pollution in the Tampa area using block group level population data and found that disadvantaged groups are exposed to higher levels of traffic pollution. (Rowangould, 2013) used census block data to analyze near-road population throughout the U.S. and found greater shares of minority residences in higher traffic density areas in most counties in the northeast U.S.; Tian (Tian et al., 2013) used census tract data and found similar results. Rowangould (Rowangould, 2013) explained that the difference might stem from different scales used for data analysis. (Tian, Goovaerts, Zhan, & Wilson, 2010) investigated racial disparities in breast cancer mortality by using census tract, ZIP code, and county-level data and found that census tract is the optimal scale to assess socioeconomic status (SES) and health disparities due to its homogeneous population characteristics and SES. However, Tian (Tian et al., 2010) also noted that additional research is needed before generalization of the conclusion. Thus, more research is needed to evaluate the choice of scales and impacts on transportation economy, health, and safety analysis.

4   Cost/Benefit Measurement

Based on research interests, cost/benefit measurement quantifies the benefits and cost of transportation system to population groups. For example, if equity assessment activities are carried out to explore whether a transportation system has economic impacts on different demographic groups equally, cost/benefit measurements must be able to quantify the economic benefits among these groups, e.g., their accessibility to employment within the investigated area. However, if transportation planning agencies are studying the benefit distribution of a transportation system in terms of public health, some health-related measurements should be selected, such as traffic safety, air quality, and active transportation (Boehmer et al., 2017; Singleton & Clifton, 2017). For active transportation, zonal-level research is usually too coarse to accurately capture non-motorized modes (Iacono, Krizek, & El-Geneidy, 2010). Thus, it is not the focus of this paper.

4.1   Accessibility-related cost/benefit measures

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 & 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 measurements that have been used in equity analysis.

The simplest accessibility-related cost/benefit measurement was proposed by (Currie, 2010) to identify the spatial need gap in public transportation supply in Melbourne, Australia. This measurement 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 $\mathcal{M}_i := [1, 2, \cdots, M_i]$, 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

$$s_i := \frac{\sum_{m \in \mathcal{M}_i} a_m l_m}{a_i}, i \in \mathcal{I}.$$

This measurement 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 measurement" in the following analysis. Due to its simplicity, this measurement has been applied to studies (Delbosc & Currie, 2011; Ricciardi, Xia, & Currie, 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 measurement. 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 measurement 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, this measurement 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 measurement, Welch and Mishra (Mishra, Welch, & Jha, 2012; Welch & 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 measurement that merely adopts service frequency to measure service quality, this measurement 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 measurement also considers that there are multiple bidirectional transit lines passing through a single station $m \in \mathcal{M}$, denoted as $\mathcal{L}_m := \{1, 2, \cdots L_m\}$, 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_{lm}^{in}. 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

$$p_{lm}^{in} := \prod_{f \in \mathcal{F}} w_f f_{lm}^{in}, \forall l \in \mathcal{L}_m, m \in \mathcal{M}_i, i \in \mathcal{I}$$

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

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

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 $l \in \mathcal{L}_m$

$$p_m = \sum_{l \in \mathcal{L}_m} \frac{p_{lm}^{out} + p_{lm}^{in}}{2}, \forall m \in \mathcal{M}_i, i \in \mathcal{I}$$

A parameter representing people's accessibility to a transit station is defined as

$$\delta_m = \alpha e^{-\beta t_{im}}$$

where α, β 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 measurement still adopts

the concept of coverage, so it falls into the category of coverage-based measurement. With the connecting powers of all $m \in \mathcal{M}_i$, the connecting power of zone i is formulated as

$$p_i = \frac{\sum_{m \in \mathcal{M}_i} p_m \delta_m}{|\mathcal{M}_i| - 1}, \forall i \in \mathcal{I}.$$

The revised measurement overcomes the first drawback in the coverage-based measurement; 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 measurements 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 measurements to identity the population of a zone $i \in \mathcal{I}$'s ability to reach the activities or destinations in all other zones $j \in \mathcal{I} \setminus \{i\}$ within the investigated area given the monetary and (or) time budget. The basic idea of reachability-based measurements 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 measurement is to define a function to capture the "accessibility- cost" relationship mathematically. Denote the accessibility and travel cost from zone $i \in \mathcal{I}$ to $j \in \mathcal{I}$ as r_{ij} and c_{ij}, respectively, then this relationship can be generally described as

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

Any functions that satisfy this property can be applied. One common example in the literature is the cumulative accessibility function $r_{ij} = \begin{cases} 1 & \text{if } c_{ij} \leq \bar{c}_i \\ 0 & \text{if } c_{ij} > \bar{c}_i \end{cases}, \forall i, j \in \mathcal{I}, i \neq j$, where $\bar{c}_{ij}$ denotes the travel cost budget of the population in zone $i \in \mathcal{I}$. 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 & Martens, 2014). Another example is $r_{ij} = e^{(-w_i c_{ij})}$, where w_i is a calibrated parameter determined by the origin zone i (Guzman, Oviedo, & Rivera, 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 $i \in \mathcal{I}$ as the sum of its accessibility to any other zone $j \in \mathcal{I} \setminus \{i\}$, i.e.,

$$r_i = \sum_{j \in \mathcal{J} \setminus \{i\}} h_j r_{ij}, \forall i \in \mathcal{J}.$$

where h_j denotes the number of activities or destinations in zone $j \in \mathcal{J}$ of interest.

4.2   Traffic pollution-related cost/benefit measures

Traffic is a major source of air pollution worldwide. Traffic emissions include carbon monoxide (CO), oxides of nitrogen (NO$_x$), volatile organic gases (VOCs), particulate matter (PM) and its constituents (fine particles PM$_{2.5}$), black or elemental carbon (BC/EC), organic carbon (OC), and metals such as lead. According to the U.S. Environmental Protection Agency (EPA) (2014), traffic contributes to more than 55% of NO$_x$ and more than 60% of CO of total air pollution in the U.S. Exposure to traffic emissions can cause adverse effects on human health. A panel of the Health HEI panel (Health Effects Institute Panel, 2010) comprehensively reviewed the literature on emissions from, exposures to, and health effects of traffic-related air pollution and concluded that exposure to traffic emissions can exacerbate asthma as well as contribute to the development of childhood asthma and other respiratory symptoms, impaired lung function, cardiovascular mortality and morbidity, and overall mortality. However, the data were not sufficient to clearly support a causal relationship for these latter outcomes.

Direct assessment of human exposures and health impacts of motor vehicle emissions is challenging as ambient pollution is a complex mixture of several emitted pollutants. Secondary formation of pollutants through chemical reactions in the air also make attribution of pollutant levels to traffic sources more difficult. Hence, surrogates such as concentrations of specific pollutants emitted from vehicles (e.g., CO, NO$_x$, EC, PM$_{2.5}$) and measures of traffic itself (such as traffic density) are commonly used to characterize traffic pollution in research. Nitrogen dioxide (NO$_2$), a component of NO$_x$, has many advantages as a surrogate of traffic pollution; its level is influenced by traffic counts (Rijnders, Janssen, Van Vliet, & Brunekreef, 2001; Stuart & Zeager, 2011), it has greater spatial heterogeneity than some other air pollutants (Jerrett et al., 2005), and exposure varies substantially among socioeconomic groups (Stroh et al., 2005). Although no surrogate was found to be ideal by the HEI panel (Health Effects Institute Panel, 2010), NO$_2$ often is used to evaluate equity of transportation pollution.

Measuring human exposure can generally be divided into two steps: 1) estimate the distribution of pollution concentrations in the study area for each unit of analysis (such as the census tract, TAZ, individual residence address, etc.), and 2) match these to the spatiotemporal location of human activities.

4.2.1 Traffic pollution measurement

Broadly, four methods are used to estimate traffic emission concentration: 1) pollution monitoring, 2) use of nearby traffic as a surrogate, 3) use of a traffic emission model and dispersion model, and 4) use of a land use regression (LUR) model. Each method is discussed in detail in following paragraphs.

a) Pollution monitoring

Monitoring is the most straightforward way to measure ambient pollution concentrations. It uses samplers or monitoring instruments to collect pollution samples at specified sites or on mobile vehicles. Concentrations can be determined through *in-situ* or laboratory analysis. A few traffic- related pollutants (CO, NO_2, $PM_{2.5}$) are routinely monitored throughout the U.S., but monitoring networks are very sparse and generally inadequate for intra-urban equity analyses. Hence, dedicated fixed-site or mobile monitoring campaigns often are used to capture variability in concentrations at high spatial resolution. Fixed-site samplers usually are mounted 2–3 meters from the ground for a few to several consecutive weeks to capture average pollution concentrations. As good indicators for traffic pollution, NO_2/ NO_x are popularly sampled to estimate traffic emission concentration. (Stuart & Zeager, 2011) used Ogawa passive samplers to measure NO_2 concentration near elementary schools to evaluate their relationship to the racial, ethnic, and income distribution of students. (Buzzelli & Jerrett, 2007; Wang, Henderson, Sbihi, Allen, & Brauer, 2013) also used Ogawa passive samplers, and (Habermann, Billger, & Haeger-Eugensson, 2015) used IVL passive samplers to measure NO_2 concentrations. As monitoring data are rarely available for each specific unit of analysis needed, aggregation models are often applied. For example, (Habermann et al., 2015) used their measurement data to determine an LUR model to estimate NO_2 levels at unsampled locations.

b) Measures of nearby traffic

Using nearby traffic characteristics such as traffic density as a surrogate for traffic pollution is the easiest method when no traffic emission data are available. (Rijnders et al., 2001; Stuart & Zeager, 2011) both found that traffic density can influence NO_2 levels near roadways. Traffic density and road density are commonly calculated for a particular spatial unit of analysis (e.g., census tract or block) as

$$Traffic\ density = \frac{\sum Length * AADT}{Area}$$

$$Road\ density = \frac{Road\ area}{Area}$$

Where length is the length of the roadway segment within the spatial unit of analysis, AADT is the average annual daily traffic of the roadway segment, road area is the road area within the spatial unit of analysis,

and area is the area of the spatial unit of analysis. Other measures of traffic used in some studies of traffic pollution include highest AADT, distance to a high-volume roadway, and number of roads (Health Effects Institute Panel, 2010).

(Tian et al., 2013) used road density and traffic density as indicators to estimate traffic pollution within a census tract area; a buffer distance along the road is created for sensitivity analysis.(Rowangould, 2013) also used traffic density as surrogate but at a finer spatial scale (census block). The improved resolution of analysis provides better alignment with spatial scale of roadway emission gradients. However, the problem of using traffic density is that AADT does not record traffic counts on small local roads and so cannot fully represent all traffic counts within a measured area. Also, the link between particular traffic density levels and pollutant concentrations is less understood (Rowangould, 2013), as it does not capture the effect of metrological effects such as wind, temperature, etc.

c) *Modeling of traffic emissions and dispersion*

Modeling traffic emission and dispersion effects is the most comprehensive approach to estimate traffic emission concentration, providing very detailed pollution levels within the research area. Generally, the process can be divided into three stages, as shown in Figure 3:

Figure 3: Traffic emission and dispersion modeling process

The first step aims to estimate characteristics of roadway travel condition such as vehicle volume and travel speeds on a road segment within a time interval. Two main methods are used to estimate these values: a travel demand model (e.g., activity-based travel demand model) or direct calculation based on available data (e.g., AADT). For example, (Sider et al., 2015) used origination-destination (OD) survey data conducted by Agence Metropolitaine de Transport (AMT), which was expanded to account for full population and generate TAZ OD matrices. The OD demand is finally allocated to road network by stochastic user-equilibrium path assignment. (Hatzopoulou & Miller, 2010) used Travel Activity Scheduler for Household Agents (TASHA) to generate trips, with network flows assigned by a traffic assignment model. (H. Yu & Stuart, 2016) used an approach that does not rely on travel demand modeling. AADT data were extracted from Florida Department of Transportation data and an empirical function from the Bureau of Public Roads (Gannett Fleming Inc., 2010a) was applied to derive traveling speed on each link:

$$S_{l,h} = S_{l,f}/[1 + \alpha_l (T_{l,h}/C_l)^{\wedge} \beta_l]$$

Where,

$S_{l,h}$: $travel\ speed\ at\ link\ l\ and\ hour\ h$

$S_{l,f}$: $free\ flow\ speed\ at\ link\ l$

$T_{l,h}$: $traffic\ volume\ at\ link\ l\ and\ hour\ h$

C_l: $road\ capacity\ at\ link\ l$

α_l, β_l: $empirical\ parameters$

AADT, however, has a disadvantage in that it does not consider many variabilities such as traffic count differences between weekdays and weekends (Fujita et al., 2003), although an activity- based travel demand model creates a sample of highly-resolved sequential spatial-temporal record of each person-day travel activity. (Hatzopoulou, Miller, & Santos, 2007) noted that travel demand modeling might be a better method for spatiotemporal allocation of traffic activity.

The second step aims to obtain link-based or area-based emission inventories through complex emission models. The most popular traffic emission model is the USEPA MOVES model (U.S. EPA, 2010a, 2010b), although the previous mobile source emission estimator (Mobile 6.2) also has been used extensively. The majority of studies used a single pollutant as an indicator to evaluate emission inventories, such as NO_2, $PM_{2.5}$, etc. MOVES requires the input of vehicle populations, vehicle mileage, and meteorological parameters and outputs emission factors (grams of pollutant per km). The emission factors are matched with vehicle speed, vehicle types, conditions, road types, etc. In addition to emissions factors, MOVES also can output an "emissions inventory" of masses (grams) from different processes such as running emissions, idling, hot-soak and cold-start, etc. The link-level emissions are calculated by multiplying link length and emission factors (Sider et al., 2015). Similar to MOVES, (Hatzopoulou & Miller, 2010) used Mobile6.2C, a Canadian version of USEPA Mobile6.2, and considered both link-based traffic emission and hot soak emission (while engine is off and mostly composed of VOC). (Gurram et al., 2015; H. Yu & Stuart, 2013) used a hybrid model to estimate traffic emission, which is composed of bottom-up and top-down approaches. Bottom-up is the same as MOVES or MOBILE 6.2 and is used to estimate the link-level emission on major roads. For the remainder of roads, top-down methods split the study area into spatial zones; with each area, emissions are allocated using a surrogate such as roadway density. Using the hybrid approach, emissions from all roadways for a large metropolitan area can be represented with reduced computational costs.

With the data of emission inventories on a road or area, other factors such as wind, land use, etc. can impact the flow of traffic emission. A dispersion model, the most popular model to estimate concentrations at an

intra-urban scale, integrates these factors to produce a high-resolution map of pollutant concentrations within study areas. It requires detailed information such as emission sources from roadway links, geographic location, and release parameters (e.g., velocity, temperature). Some research also includes emission sources from stationary points and areas (H. Yu & Stuart, 2013, 2016). The output of a dispersion model is pollutant concentration at receptor points in the study area. The spatial resolution of the network or receptors is specified by users. Hence, pollutant concentrations can be produced for the centroids of standard geographic entities (e.g., census block, tract, TAZ), or values from a regular network can be aggregated to these areas. However, a dispersion model has the limitation of significant computing efforts, which might restrict the feasibility of widespread use. CALPUFF and RLINE are dispersion models popularly used in transportation context. (Gurram et al., 2015; H. Yu & Stuart, 2013) used CALPUFF to estimate emission concentrations on a receptor grid with 1km spatial resolution covering Hillsborough County, Florida. To overcome the computation limitation issue, (Rowangould, 2015) developed an approach that breaks up the computational domain into parallel simulations to enable 100 meter spatial resolution near roadways, using the AERMOD dispersion model. Without using a dispersion model, (Sider et al., 2015) used a rounded buffer around road links to approximate dispersion effects.

d) Land use regression (LUR) model

LUR is the newest addition to pollution modeling and has become very popular in the last 10 years. It is used to estimate the spatial variation of small-scale pollution concentration in an intra-urban area. (de Hoogh et al., 2014) compared the land use regression model, dispersion model, and measured pollution data for 13 study areas; results indicate that both methods are useful to study small-scale variations of traffic pollution. LUR has the basic form:

$$Pollution\ Concentration = (\beta_0 + \beta_1 X_1 + \beta_2 X_2 \cdots \beta_x X_x) + \varepsilon$$

LUR models variables and coefficients by regressing measured ambient pollution levels and against independent land-use and other variables such as nearby traffic counts, elevation, etc. Once the variables and coefficients are determined, the model can be applied to estimate pollution levels for unsampled locations. The process of modeling can be divided into four steps: 1) pollution sampling, 2) independent variable selection, 3) regression modeling, and 4) model validation and evaluation.

Pollution sampling was described at Section 4.2.1 Part a. Usually the concentration of NO_2, NO_x is used as the indicator for traffic emissions and as the dependent variable for the regression model. The independent variables are unique to different cities and usually are characterized for buffers areas surrounding each concentration measurement site location. (Ryan & LeMasters, 2007) grouped these variables into four

classes: road type, traffic count, elevation, and land cover. (Dirgawati et al., 2015) considered 124 potential environmental predictors generated by GIS categorized by land use, population/ household density, and traffic-related variables. The final model shows that statistically significant variables are traffic intensity on nearest road, household density, industry and commercial area, and road length. (Habermann et al., 2015) assessed 31 independent variables; in the final multivariate analysis, two (elevation, traffic count) were statistically significant in Gothenburg, Sweden; the R^2 of final model to predict NO_2 was 0.594. To select proper independent variables, (Dirgawati et al., 2015; Habermann et al., 2015; Wang et al., 2013) conducted univariate regression to assess the correlation between pollutant concentration and each predictor variable. Correlations with highest values were reserved for multivariate regression. For the remainder of the predictors, (Wang et al., 2013) used stepwise linear regression, and (Dirgawati et al., 2015) proposed three criteria that need to be satisfied and combined with AIC and BIC to select remaining variables. (Dirgawati et al., 2015) also used a Variance Inflation Factor (VIF) to ensure that the final model was not affected by multicollinearity. After building the multivariable regression model, the performance can be evaluated by cross validation (Dirgawati et al., 2015; Wang et al., 2013) and root mean square error (RMSE) (Brauer et al., 2003) The final model can be used to predict NO_2 concentration for unsampled locations (e.g., residential address (Dirgawati et al., 2015)) with available predictor data. Long-term NO_2 predictability of LUR is explored by (Cesaroni et al., 2012; Eeftens et al., 2011; Wang et al., 2013), who found that local LUR models perform very well for long-term forecasting, whereas hindcasting has lower accuracy. LUR has some limitations. Most LUR studies used only short-term monitoring campaign data, which ignore temporal variation (Habermann et al., 2015). Also, there is an issue of the availability of predictors data. For example, household density and industrial area data are sourced from different years from the year the pollution concentration was measured, which might also influence the quality of the LUR model (Dirgawati et al., 2015). Recent work has begun to explore combinations of approaches for determining pollutant concentration, including the use of temporally varying pollutant concentrations and dispersion model outputs as predictor variables in LUR modeling.

4.2.2 *Traffic exposure measurement*

A person's total daily exposure to pollutants depends on where they travel, time spent in vehicles, being indoors and outdoors, ventilation of buildings, etc. (Rowangould, 2013). Due to the natural complexity of travel behavior, the individual or population exposure estimation is usually simplified and can be grouped into two categories: 1) fixed location-based approach (Rowangould, 2015; Sider et al., 2015; Stuart et al., 2009; H. Yu & Stuart, 2013, 2016), and 2) activity-based approach (Hatzopoulou & Miller, 2010).

A fixed location-based approach estimates individual or population exposure based on residential address, school location, etc., by estimating the concentration at that location. For equity analysis, the exposure often is determined at the spatial level of the population data (such census block, tract, or TAZ). In (H. Yu & Stuart, 2016), the exposure for small block groups (less than 1 km^2) was estimated by the concentration at a model receptor (from a dispersion model) at a block group centroid; for large block groups, the exposure was calculated by averaging concentrations of all gridded receptors within the block group. (Sider et al., 2015) calculated average exposures within a TAZ by normalizing the total concentration within the land area of the TAZ to generate average emission density (kg/km^2). If the concentration data are time varying (e.g., hourly concentrations from dispersion modeling), then any temporal statistic of exposures (such as daily-average exposure, annual-average exposure, or maximum hourly exposure) can be calculated for any receptor location or spatial unit of interest (e.g., census block) (H. Yu & Stuart, 2013). Finally, exposure for each population subgroup can be calculated by summing the population weight of the subgroup in the spatial area times pollution exposure for each spatial area divided by total subgroup populations. The disadvantage of using a fixed location proximity approach is that it does not capture population activity patterns.

The activity-based approach tracks individuals throughout daily activities and estimates exposure for each individual using spatiotemporally varying activity and concentration data. (Gurram et al., 2015)) estimated the individual activity-based exposures using

$$C_A = \left(\sum c_\sigma \Delta t_\sigma\right)/T$$

where,

C_A: average activity-based exposure concentration estimate

σ: spatiotemporal location (e.g., latitude, longitude, time)

c_σ: concentration at each activity location

Δt_σ: time spent at each activity location (or resolution of the activity data)

T: averaging time (e.g., 24 hours for average daily exposure concentration)

Concentration data were obtained from dispersion modeling. Fixed activity locations of individuals were obtained from the National Household Travel Survey (NHTS), and locations of during travel were determined using a shortest-time path model. To prepare the data for vertical analysis, the individual exposures were grouped by age, race, and household income with the population information provided by the NHTS. Other studies (Gurram, 2017; Gurram et al., 2018; Hatzopoulou & Miller, 2010) used activity-based travel demand models to obtain individual activity locations. (Hatzopoulou & Miller, 2010) did not

consider exposures during travel in their formula to calculate individual daily exposure, arguing that in-vehicle exposures are significantly different from ambient air concentrations.

Comparing individual exposure results from an activity-based analysis with a fixed-location analysis (based on residential location) (Gurram et al., 2015) suggests that residence-based exposure may misclassify the exposures of social disadvantaged population groups less than that of others.

4.3   Traffic safety-related cost/benefit measures

Traffic safety mainly refers to injuries or death caused by motor vehicle accidents. Factors that can influence traffic injuries/casualties has been studied extensively, including population characteristics and demographics (e.g., race, income), physical development (e.g., employment density), and environment factors (e.g., road density, traffic density) in an area (Kravetz & Noland, 2012). Population characteristics and demographics is the main focus of traffic safety equity, as it provides insights into the distribution of crash injuries/causalities among population groups, although physical development and environmental factors may help explain the cause of inequity. The cost/benefit measure of traffic safety mainly refers to crash rate or crashes and injuries/fatalities indices. For example, crash rate can be the number of White people injured in a census area divided by total number of the White population in that area. Equity between races commonly adopt this measure. For indices, the example can be simply number of pedestrian injuries in a census area or number of crashes per 100,000 populations, etc. Unlike cost/benefit measurements for accessibility and traffic pollution, which may involve formula calculation and modeling, the cost/benefit measure of traffic safety is mainly obtained through data source and processing with focus on pedestrian, cyclist, and driver injuries and casualties. (Kravetz & Noland, 2012; Noland, Klein, & Tulach, 2013) extracted crash data from the Plan4Safety database, which has comprehensive records of crashes occurring in New Jersey, including date of time, place, severity, vehicle actions and directions, environmental and surface conditions, occupants and pedestrians involved, driver characteristics, etc. The geocoded crash data are aggregated at the census block group level for further analysis. Similarly, (Steinbach, Green, Edwards, & Grundy, 2010; Steinbach, Green, Kenward, & Edwards, 2016) obtained child injury data STATS19 from the London Road Safety Unit that includes all reported causalities and traffic collisions in London. The crash data location is assigned to a lower super output area (LSOA) for equity analysis purposes. However, frequently-used administrative data has the limitation of under-reporting, especially related to slight injuries. Police may focus on accidents that legally must be reported but may ignore accidents not involving motor vehicles. For example, STATS19 data cover accidents on public highways but exclude single bicycle incidents and pedestrian falls (Aldred, 2018). As summarized by (Ahmed, Sadullah, & Yahya, 2017), the percentage error reporting (under-reporting) for developed countries is low for fatal injuries, but it can be

more than 40 percent for slight injury accidents. There are many methods to rectify the data, the most popular of which is the capture-recapture method. Other methods including comparison with health sector data and probabilistic linkage. But it is not the focus of this paper to discuss data enhancement. (Kraemer & Benton, 2015) adopted the capture-recapture method to assemble the data, which combines data from two independent registries of fatal crashes of wheelchair users. One registry used the National Highway Traffic Safety Administration (NHTSA) Fatality Analysis Reporting System (FARS); preliminary assessment, however, indicates incompleteness of data. A second registry was constructed by searching keywords in a LexisNexis U.S. newspaper database. The unmatched data from the news were sought in the FARS database; almost all unmatched cases were identified in FARS, as those data failed to be coded as wheelchair users. Excluding administrative data, (Morency et al., 2012) used ambulance service data to collect information of road traffic injuries, although ambulance service data naturally neglect more light injuries. (Aldred, 2018) extracted data from the National Travel Survey (NTS) in the UK, in which respondents are asked to recall if they had involvement in road accidents over the previous three years; details such as injury, treatment, police involvement, and mode used were recorded. Although the data do not include the geospatial information of accidents, the detailed personal information (e.g., income, gender, age) and accident data still provide insights of equity issues.

5   Equity Assessment Approach

Various equity assessment approaches have been proposed in the past few decades, which are broadly divided into three categories in this paper. The following paragraphs discuss each equity assessment method.

5.1   *Spatial mismatch analysis-based approach*

Mismatch analysis is the most traditional approach to study the equity performance of a transportation system within an investigated area; its history can date back to the very earliest study in measuring the performance of public transport in meeting the transport needs for different demographic groups (Currie, 2004). Basically, this method presents the distributions of both the population and cost/benefit measurements in maps or tables and then manually compares the distributions of these two measurements. 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 population and cost/benefit measurements in two different maps (usually in GIS), where each zone has one color on a scale from the lowest to the highest quintile (Kaplan, Popoks, Prato, & Ceder, 2014). The statistical metrics are usually the average values (of the population and cost/benefit measurements) 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 measurement in each zone within the studied area needs to be plotted. For example, Golub (Golub & Martens, 2014) analyzed the distributions of accessibility across the San Francisco Bay Area by different modes (automobile, transit) and for different destination types (manufacturing jobs, service jobs) with the same method. To assess vertical equity, both the population measurement and the cost/benefit measurement can be plotted in two maps and then the distributions of these two maps can be compared manually. For instance, Kalpan et al. (Kaplan et al., 2014) evaluated the vertical equity of the public transit system in the Greater Copenhagen area by mapping the distributions of the accessibility measurements and those of the population density, average income, number of jobs, etc. Tian (Tian et al., 2013) investigated whether socioeconomic status and racial differences have correlation with road/ traffic density for Rhode Island by mapping spatial distributions of road, traffic density, race/ ethnicity, and socioeconomic status on a GIS map. 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 (Boarnet, Giuliano, Hou, & Shin, 2017; El-Geneidy et al., 2016; Stuart & Zeager, 2011).

Although the mapping approach can offer intuitive information, it might be cumbersome to identify the gaps from the maps manually for all zones in a large metropolitan area. Thus, simplified approaches have been proposed in the literature. Currie (Currie, 2010)plotted the supply-demand relationship of each zone within the investigate area in a two-dimensional Cartesian coordinate system. In this way, zones with different degrees of inequality are clustered in different regions in the coordinate system and, thus, can be very efficiently identified. For instance, zones that fall into the southeast corner in the coordinate system have the highest demand but the lowest transit supply, which reflects the highest degree of inequality. Another method is to combine the population and cost/benefit measurements to obtain a new measurement (if possible) and then plot the combined measurement. For instance, to analyze transit provision with respect to social needs in Melbourne, Currie (Currie, 2010) defined the need-gap of zone $i \in \mathcal{I}$ as the difference between its transport provision and need index, i.e.,

$$g_i := s_i - d_i, i \in \mathcal{I}.$$

where s_i denotes the transit provision and d_i the transport need index in zone i, respectively. This measurement links transport provisions and the population's heterogeneity in terms of their travel needs and offers an opportunity to analyze the vertical equity with respect to the mobility need and ability. A similar method was applied to investigate the disparity between transport provisions and social exclusion in Cali, Colombia, by taking transport social needs into account. Later, Ricciardi (Ricciardi et al., 2015) replaced the need indexes with potential demands generated from Distribution Fit Tool in the Matlab studio, based on which the equity distributions of three separate social disadvantaged groups (i.e., older adults, low-income, and households without a car) were investigated.

Note that although the mismatch analysis is quite simple and intuitive, it is very subjective and cannot offer quantitative information of the equity performance. Therefore, some quantitative analysis approaches based on inequality index formulation or statistical model have been proposed.

5.2   Inequality indicator-based approach

The first mainstream quantitative analysis approach is based on inequality indexes that have been extensively used in social science to offer a quantitative indicator of the degree of inequality among populations. Popular inequality indicators include Gini, Atkinson, Thiel, and concentration. In the following paragraphs, only the Gini index and Atkinson index are discussed, as they are used extensively in a transportation context.

The Gini index traditionally has been used to evaluate the distribution of wealth or income among a population. Delbosc and Currie (Delbosc & Currie, 2011) first applied this tool along with the Lorenz curve to analyze equity performance in transit supply in Melbourne. The approximated mathematical formulation of Gini index is

$$G_\alpha = 1 - \sum_{k=1}^{K}(x_k - x_{k-1})(y_k - y_{k-1})$$

where α denotes the specific zone or demographic group for which we are computing the Gini index; G_α is the Gini index for that zone or demographic group α; x_k is the cumulative proportion of the population measurement ($x_0 = 0, x_K = 1$), and y_k is the cumulative proportion of the cost/benefit measurement ($y_0 = 1, y_K = 1$. Note that the Gini index ranges from 0 to 1, with 0 indicating perfect equality and 1 perfect inequality. Due to the computational tractability and intuition of the Gini index, many studies have followed the pioneering work of Currie and used the Gini index as an overall index of the equity performance in transportation systems. For instance, Kalpan et al. (Kaplan et al., 2014) found that the Gini index for the

public transit systems in the Greater Copenhagen area was 0.33 and concluded that the studied area was broadly equitable from the spatial perspective (horizontal equity). Guzman et al. (Guzman et al., 2017) adopted the Gini index to compare the economic benefit distributions of different travel modes, i.e., automobile and public transport, among different demographic groups in Bogotá (vertical equity with respect to income and social class). Welch and Mashira (Welch & Mishra, 2013)compared the Gini indexes for several subzones in Washington D.C. and Baltimore, which led to their conclusion that the indexes can reflect whether the transit supplies are concentrated or scattered within the studied zones; whether it is equitable across different demographic groups depends on how transit planning agencies define equity. As can be seen from these examples, the Gini index can be easily adapted to various contexts and used for evaluating both horizontal and vertical equity.

The Atkinson index was initially derived for income inequality (Atkinson, 1970) and has been popularly applied in the environmental justice context. The Atkinson index has many desirable features such as sub-group decomposable and explicit value judgement of distribution. It is formulated as:

$$AI = 1 - [\frac{1}{n}\sum_{i=1}^{n}[\frac{y_i}{\bar{y}}]^{1-\varepsilon)}]^{\frac{1}{1-\varepsilon}}$$

where,

y_i: $Income\ of\ individual$

$\bar{y}$: $average\ income$

n: $population$

ε: $inequality\ aversion\ parameter$

Like the Gini index, the Atkinson index ranges between 0 and 1, with 0 being the complete equal and 1 being complete unequal. The ε ranges from 0 to infinity; 0 indicates no societal concern with inequality, and a higher value means a higher concern for the low benefit (e.g., income) group. However, for health risk, it is more concerned with high exposure values, and the typical value of ε ranges from 0.25 to 2 according to (Levy, Chemerynski, & Tuchmann, 2006).

Applying this formula to the transportation health context. y_i refers to the risk of individual i and n is the number of individual. The Atkinson index is suggested by (Levy et al., 2006) to evaluate health inequality after comparing with a variety of indicators. As the Atkinson index does not violate the Pigou-Dalton transfer principle (for income equity evaluation, the indicator decreases when income is transferred from richer to poorer and does not decrease when income is transferred from poorer to richer), it is subgroup-

decomposable and has an explicit inequality aversion parameter that allows equality to be evaluated from different societal viewpoints. Levy (Levy et al., 2009) applied the Atkinson Index to quantify the changes of equality benefits for baseline and different mobile source control strategies in Boston.

In addition to borrowing concepts of inequality indicators from social science, transportation and environmental professionals also developed an inequality index that intuitively quantifies disparities between populations groups. Harner et al. (Harner, Warner, Pierce, & Huber, 2002) developed an environmental justice index (Comparative Environmental Risk Index, CERI) that captured if racial or social minority groups are more likely to be exposed to environmental hazards than the rest of population. For example, to compare pollution exposure between White and non-White populations, the formula can be written as:

$$CERI_{nw} = \frac{(\frac{at\ risk\ nonwhites}{total\ MSA\ nonwhites})}{(\frac{at\ risk\ whites}{total\ MSA\ whites})}$$

Where MSA is the Metropolitan Statistical Area. The index also can be used to compare poor, non-poor, etc. If $CERI_{nw}$ is greater than 1, non-Whites are exposed to higher environmental risk, and vice versa. Similarly, Stuart et al. (Stuart et al., 2009) and Yu and Stuart (H. Yu & Stuart, 2016) developed a quantitative subgroup index of inequity that is formulated as

$$F_i = \log(\frac{Z_i}{T_i})$$

Where i is the subgroup, F quantifies the degree that subgroup i is disproportionately exposed to environmental pollution, and Z is the fraction of the total subgroup population of a larger area (such as a county) that is exposed to a particular exposure level. The exposure level used for the calculation can be based on proximity to sources or modeled concentrations (for example, living within a buffer zone of a pollution source or living in an area with ambient concentrations higher than a healthy threshold level). T is the overall fraction of a subgroup in the large-area population. A good feature with this formulation is easy interpretation. Positive F means members of subgroup i tend to be exposed to the threshold level of pollution, and negative F means they tend not to be exposed to this level. Applying subgroup index of inequity, Stuart et al. (Stuart et al., 2009)found Blacks, Hispanics, and people living in poverty tends to live closer to sources of air pollution and further from ambient monitoring sites. As concentration or traffic pollution increased, Yu and Stuart (H. Yu & Stuart, 2013) found the disparity between different groups also increased. For secondary pollutants such as formaldehyde and acetaldehyde, Yu and Stuart (H. Yu & Stuart,

2016) found that exposure was largely reversed among groups and suggested that disparities in exposure depend on pollutant type.

5.3 Statistical approach

Statistical methods such as ANOVA, correlation, regression modeling, etc., are widely used in the literature to assess transportation equity. Stuart and Zeager (Stuart & Zeager, 2011)used bivariate linear correlations to investigate relationships between NO_2 concentration levels and population subgroup percentage enrollments of elementary school students. Results showed positive associations between Black enrollment versus NO_2. To determine if levels of emissions and exposures were significantly different between groups with varying social disadvantage index Sider et al. (Sider et al., 2015) used one-way ANOVA and found the difference to be statistically significant. Further, it was found that more socially-disadvantaged groups are exposed to more pollutions.

Previous methods present a degree of inequality and are measured with regard to one population measure. However, inequality also can be evaluated by contribution of each population measure and additional environmental variables (independent variable) to the cost/benefit measure (dependent variable) and multiple independent variables can be considered. For example, the disparity of traffic accidents in different community areas could be explained by average household income, population density, race proportion, road infrastructure, etc., and each independent variable might explain different proportions of traffic accidents. The level of inequality can be evaluated by the sign and magnitude of coefficients of independent variables. The goal of regression modeling is to integrate more than one independent variable to quantify causal relationships for each independent variable and cost/benefit measures. Regression modeling is popularly applied for analyzing equity of traffic safety and pollution.

For traffic safety, negative binomial regression and multilevel Poisson regression are commonly used, as traffic accidents do not occur frequently, which follows Poisson or negative binomial distribution. Morency et al. (Morency et al., 2012) used a multilevel Poisson regression model to examine the roadway environment impacts on traffic injuries across wealthy and poor urban areas. Results found that there were significantly higher traffic injuries in poor areas than wealthy areas and that traffic volume, intersection geometry, pedestrian and cyclist volumes can explain the substantial portion of road traffic injuries in poor areas. Zhang and Lin (Zhang & Lin, 2013) also used a Poisson regression approach to examine the likelihood of traffic injuries by driver race, age, sex, and residence status (urban/ rural) in Nebraska. Results indicated that Black drivers residing in rural areas have significantly higher traffic injuries. Interaction terms such as race-sex and race-age should be considered based on their significant contribution to disparity. A

negative binomial regression model is applied if data are over-dispersed. (Harper et al., 2015; Kravetz & Noland, 2012) used negative binomial regression to evaluate income and minority disparities of pedestrian crash rates in New Jersey and found that low median income and high percentage of Black and Latino populations are highly associated with a higher number of traffic crashes. Harper et al. (Harper et al., 2015) also used a negative binomial regression model and looked at long-term motor vehicle accident deaths from the education disparity perspective. Results indicated that overall MVA death rates declined from 1995 to 2000 while disparities persisted or worsened.

For traffic pollution, (Buzzelli & Jerrett, 2007) simplified the exposure to binary level, e.g., low and high exposure. Logistic regression was used to capture the factors (number of dwelling, low income, low education, medium income, race, etc.) that discriminate among high and low exposures in Toronto. Results indicated that education is the most important variable. (Buzzelli & Jerrett, 2007; Goodman et al., 2011; Havard et al., 2009) found a nonlinear relationship between NO_2 levels and the deprivation index by trying both the linear regression model and the spatial autoregressive (SAR) model. The SAR model had a better fit with decreased association between NO_2 levels and deprivation index. He suggested considering geospatial factors for epidemiology studies.

6   Applications to emerging transportation technologies and future opportunities

This section presents some existing applications of the three-step framework to analyze equity performance in emerging transportation systems. Although emerging transportation technologies have sprung up in many cities worldwide, studies on their equity performance are still in their infancy, and only limited research has been conducted on bike-sharing and car-sharing services.

6.1   *Equity analysis of bike-sharing services*

Among various emerging transportation technologies, the equity issue in bike-sharing systems has received the most attention. To evaluate whether the provision of bike-sharing services is equal in the U.S., Ursaki and Aultman-Hall (Ursaki & Aultman-Hall, 2016) defined the service area of bike-sharing at 500 meters within bike-sharing stations (accessibility). The study used a T-student test to compare social and economic characteristics of census block groups that are within a service area to outside a service area for eight U.S. bike-sharing programs. Results indicated that percent of the White, college-educated, higher-income population are higher than the results of groups in Chicago, Denver, Seattle, and New York city. Washington, DC and Arlington, Virginia, are the most equitable, although they still show differences for

household income variables. Likewise, Gavin et al. (Gavin, Bennett, Auchincloss, & Katenta, 2016) compared bike-share membership survey data with census residence characteristics within bicycle service areas (0.5 mile within bike stations) for three cities and concluded that users are more likely to be residents who are male, young, White, affluent, and educated. Although there are great disparities of accessibility to bike-sharing, Saviskas and Sohn (Saviskas & Sohn, 2015) surveyed populations in Berkeley and concluded that low-income and high-income people had similar levels of interest in using bike-sharing; also, the primary purpose for bike-sharing was social activities, whereas fewer than 50 percent used it for commuting or shopping. However, Shaheen, Guzman and Zhang (Shaheen, Guzman, & Zhang, 2012) found that bike-sharing users were mostly likely to use it for commuting purpose, then social activities.

6.2   *Equity analysis of car-sharing services*

To evaluate the equity of distribution of car-sharing service in New York City, Shellooe (Shellooe, 2013) used a regression model that integrated dependent variable car-sharing density and independent variable population density, number of rail stops, percentage of zero car households, income, race, education degree, etc., in community districts. Car-sharing density is defined as the number of Zipcars per square mile in a community district in New York City. Results indicated that Zipcar density increased with high-income, highly-educated White populations.

7   Conclusion

This paper summarizes a general framework that evaluates transportation equity with regard to economy, environment, and health. The equity assessment framework is divided into three components—population measure, cost/benefit measure, and equity measure—which are commonly applied in the literature. Following is a summary of equity findings from the literature and indications of research gaps of applying the investigated framework to emerging transportation technologies.

7.1   *Findings of transportation equity*

Generally speaking, disadvantaged groups enjoy the highest accessibility to public transportation, such as the lowest income groups in Inner Melbourne (Delbosc & Currie, 2011), Perth (Ricciardi et al., 2015), the San Diego Metropolitan Area (Boarnet et al., 2017), and the San Francisco Bay Area (Golub & Martens, 2014)); the youngest groups in Inner Melbourne (Delbosc & Currie, 2011); older adults in outer Melbourne (Delbosc & Currie, 2011) and Perth (Ricciardi et al., 2015); and groups with the highest vulnerability indicator (El-Geneidy et al., 2016). Further, households without cars are also advantaged in terms of public

transportation accessibility (Melbourne (Delbosc & Currie, 2011), Perth (Ricciardi et al., 2015), Bogota (Guzman et al., 2017)). However, as automobiles can bring better accessibility to activities and destinations within an area (Golub & Martens, 2014) and high-income groups have higher car accessibility (Guzman et al., 2017), the accessibility of low-income and minority groups is lower than other groups if both public transport and automobile are taken into account (El-Geneidy et al., 2016; Golub & Martens, 2014; Grengs, 2010; Kawabata, 2009). To reduce the accessibility gap caused by the difference between public transport and automobile, changing the access and egress to and from stations may be an effective solution (Boarnet et al., 2017).

For environment and safety, findings are consistent on distribution of equity among population groups. Overall, populations that are White (Gurram et al., 2015; Stuart et al., 2009; Tian et al., 2013; H. Yu & Stuart, 2013), high-income (Gurram et al., 2015; Stuart et al., 2009; Tian et al., 2013; H. Yu & Stuart, 2013), and urban (Gurram et al., 2015) are less exposed to traffic pollution, and populations that have high income (Kravetz & Noland, 2012; Morency et al., 2012), high education (Harper et al., 2015), and high White population density (C.-Y. Yu, 2014; Zhang & Lin, 2013) have fewer vehicle accident injuries and deaths. Also, long-term equity might be worse for socially-disadvantaged populations with regard to vehicle accident deaths (Harper et al., 2015) and disadvantaged communities are exposed to higher levels of emissions but contribute very little pollution in Montreal (Sider et al., 2015).

Most socially-advantaged groups enjoy greater benefits from transportation accessibility, health, and safety. The results depend on locality and other factors. Yu and Stuart (H. Yu & Stuart, 2016) measured multiple traffic-related pollutants and found acetaldehyde and formaldehyde had the highest concentration levels for White, high-income populations in Tampa. Buzzelli and Jerrett (Buzzelli & Jerrett, 2007) also found high income and dwelling values in central Toronto are susceptible to exposure. Havard et al. (Havard et al., 2009) found a nonlinear relationship between NO_2 levels and the deprivation index.

7.2   Research gap application of equity measures on emerging transportation technologies
Currently, many transportation equity assessments focus on traditional transportation systems, whereas very few studies have been conducted to analyze the equity performance of emerging transportation systems. This section identifies several research gaps that would be applied to assess equity in emerging transportation:

- **Assessment of other emerging technologies.** As pointed out in Section 6, only the equity performance of some bike- and car-sharing systems have been studied. The equity issue of other emerging

transportation technologies such as electric vehicles, ride-sharing, and autonomous vehicles are missing in the current literature.

- **Proposition of station-free measurements.** The limited studies on bike- and car-sharing simply address whether the provision of bike/car-sharing facilities/services is equal among the population, which is similar to the classic coverage-based approach proposed by Delbosc and Currie (Delbosc & Currie, 2011) to evaluate equity performance for the public transportation system in Melbourne (Delbosc & Currie, 2011). As have pointed out in Section 4.1, this approach is based on the concept of the service area of a station. However, many emerging shared-mobility services such as free-floating bike-sharing, free-floating car-sharing, ride-sharing, and autonomous vehicles are station-free. Thus, the concept of service area may not be directly applicable to such systems, and methodologies that enable the concept of service area are needed to make this approach applicable. Otherwise, other non-coverage-based approaches should be proposed.

- **Consideration of the operation characteristics of emerging transportation services.** Many operational details concerning the service quality of an emerging transportation system are not the same as those in traditional transportation services. For instance, in bike-sharing systems, rebalancing (redistributing bikes across the network using a fleet of vehicles) is a unique operation that can largely affect service quality but has not been studied in assessing equity performance in traditional transportation systems. Thus, although measurements that can reflect the service quality of traditional public transit systems are present, methodologies that can capture more operational details are needed for emerging transportation technologies.

- **Assessment under a multimodal transportation system context.** Different travel modes in a transportation system are interrelated with and affected by each other, especially for many emerging transportation services. For example, apart from providing stand-alone services for short-distance travel, bike-sharing is also an effective solution to the first/last-mile problem for transit. In this context, analyzing a bike-sharing system without considering public transit is not sensible. Also, a multimodal perspective is important, as it provides overall equity performance of a transportation system instead of just one mode, which could be used to compare the overall equity performance before and after the advent of an emerging transportation technology. Hence, developing a methodology that can be applied in a multimodal transportation system context should be an interesting future direction.

- **Integrated assessment with respect to economy, environment, and public health.** Research on emerging transportation technologies mainly focuses on accessibility (more precisely, accessibility to bike- or car-sharing facilities). However, whether these innovative services can bring benefits to different demographic groups in terms of environment and public health has not been frequently discussed, and no integrated assessment takes into account all these three aspects. Research has found

that car-sharing programs produce more point emissions at parking areas, whereas the shift from automobile users to bike-sharing customers can reduce emissions on roads. In this context, traffic emissions and dispersion might need to be reconsidered, as does the corresponding benefit distribution. Such an integrated study has not been carried out for traditional transportation systems either.

- **Disaggregate measures with high-resolution inputs.** Due to the lack of high-resolution data, equity assessment in traditional transportation systems usually adopts aggregate information of both population and cost/benefit measures, e.g., zone-level information. This practice can provide a macroscopic assessment of the equity performance of a transportation system, but results in some errors. Many existing emerging transportation technologies are based on smartphone applications or website toolkits, e.g., Uber, Zipcar, Ofo, and some future technologies are believed to be the same (http://www.next-future-mobility.com/). This operational mode opens up an opportunity to collect individual-level population characteristics. Further, vehicles offering emerging transportation services are usually installed with GPS, which may offer individual-level trajectory information. With this, the proposition of disaggregate assessment approaches might be an interesting topic.

References

Ahmed, A., Sadullah, A. F. M., & Yahya, A. S. (2017). Errors in accident data, its types, causes and methods of rectification-analysis of the literature. *Accident Analysis & Prevention*.

Aldred, R. (2018). Inequalities in self-report road injury risk in Britain: A new analysis of National Travel Survey data, focusing on pedestrian injuries. *Journal of Transport & Health*.

Atkinson, A. B. (1970). On the measurement of inequality. *Journal of economic theory, 2*(3), 244-263.

Boarnet, M. G., Giuliano, G., Hou, Y., & Shin, E. J. (2017). First/last mile transit access as an equity planning issue. *Transportation Research Part A: Policy and Practice, 103*, 296-310.

Boehmer, T. K., Wendel, A. M., Bowers, F., Robb, K., Christopher, E., Broehm, J. E., . . . Ralph, J. (2017). US Transportation and Health Tool: Data for action. *Journal of Transport & Health, 6*, 530-537.

Boyce, J. K., Zwickl, K., & Ash, M. (2016). Measuring environmental inequality. *Ecological Economics, 124*, 114-123.

Brauer, M., Hoek, G., van Vliet, P., Meliefste, K., Fischer, P., Gehring, U., . . . Lewne, M. (2003). Estimating long-term average particulate air pollution concentrations: application of traffic indicators and geographic information systems. *Epidemiology*, 228-239.

Buck, D. (2013). Encouraging equitable access to public bikesharing systems. *ITE Journal, 83*(3), 24-27.

Buzzelli, M., & Jerrett, M. (2007). Geographies of susceptibility and exposure in the city: environmental inequity of traffic-related air pollution in Toronto. *Canadian journal of regional science, 30*(2).

Cesaroni, G., Porta, D., Badaloni, C., Stafoggia, M., Eeftens, M., Meliefste, K., & Forastiere, F. (2012). Nitrogen dioxide levels estimated from land use regression models several years apart and association with mortality in a large cohort study. *Environmental Health, 11*(1), 48.

Currie, G. (2004). Gap analysis of public transport needs: measuring spatial distribution of public transport needs and identifying gaps in the quality of public transport provision. *Transportation Research Record: Journal of the Transportation Research Board*(1895), 137-146.

Currie, G. (2010). Quantifying spatial gaps in public transport supply based on social needs. *Journal of Transport Geography, 18*(1), 31-41.

de Hoogh, K., Korek, M., Vienneau, D., Keuken, M., Kukkonen, J., Nieuwenhuijsen, M. J., . . . Cesaroni, G. (2014). Comparing land use regression and dispersion modelling to assess residential exposure to ambient air pollution for epidemiological studies. *Environment international, 73*, 382-392.

Delbosc, A., & Currie, G. (2011). Using Lorenz curves to assess public transport equity. *Journal of Transport Geography, 19*(6), 1252-1259.

Deng, Y., & Cardin, M.-A. (2018). Integrating operational decisions into the planning of one-way vehicle-sharing systems under uncertainty. *Transportation Research Part C: Emerging Technologies, 86*, 407-424.

Dirgawati, M., Barnes, R., Wheeler, A. J., Arnold, A.-L., McCaul, K. A., Stuart, A. L., . . . Heyworth, J. S. (2015). Development of land use regression models for predicting exposure to NO2 and NOx in metropolitan Perth, Western Australia. *Environmental Modelling & Software, 74*, 258-267.

Eeftens, M., Beelen, R., Fischer, P., Brunekreef, B., Meliefste, K., & Hoek, G. (2011). Stability of measured and modelled spatial contrasts in NO2 over time. *Occupational and environmental medicine*, oem. 2010.061135.

El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., & Loong, C. (2016). The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice, 91*, 302-316.

Forkenbrock, D. J., & Schweitzer, L. A. (1999). Environmental justice in transportation planning. *Journal of the American Planning Association, 65*(1), 96-112.

Fujita, E. M., Campbell, D. E., Zielinska, B., Sagebiel, J. C., Bowen, J. L., Goliff, W. S., . . . Lawson, D. R. (2003). Diurnal and weekday variations in the source contributions of ozone precursors in California's South Coast Air Basin. *Journal of the Air & Waste Management Association, 53*(7), 844-863.

Gannett Fleming Inc. (2010a). *Technical Report 1 — Tampa Bay Regional Planning Model (TBRPM) Version 7.0 Validation Report*. Retrieved from

Gavin, K., Bennett, A., Auchincloss, A. H., & Katenta, A. (2016). A brief study exploring social equity within bicycle share programs. *Transportation Letters, 8*(3), 177-180.

Golub, A., & Martens, K. (2014). Using principles of justice to assess the modal equity of regional transportation plans. *Journal of Transport Geography, 41*, 10-20.

Goodman, A., Wilkinson, P., Stafford, M., & Tonne, C. (2011). Characterising socio-economic inequalities in exposure to air pollution: a comparison of socio-economic markers and scales of measurement. *Health & place, 17*(3), 767-774.

Grengs, J. (2010). Job accessibility and the modal mismatch in Detroit. *Journal of Transport Geography, 18*(1), 42-54.

Griffin, G. P., & Sener, I. N. (2015). *Equity analysis of transit service in large auto-oriented cities in the United States*. Retrieved from

Gurram, S. (2017). Understanding the Linkages between Urban Transportation Design and Population Exposure to Traffic-Related Air Pollution: Application of an Integrated Transportation and Air Pollution Modeling Framework to Tampa, FL.

Gurram, S., Stuart, A. L., & Pinjari, A. R. (2015). Impacts of travel activity and urbanicity on exposures to ambient oxides of nitrogen and on exposure disparities. *Air Quality, Atmosphere & Health, 8*(1), 97-114.

Gurram, S., Stuart, A. L., & Pinjari, A. R. (2018). Agent-based modeling to estimate exposures to urban air pollution from transportation: exposure disparities and impacts of high-resolution data. *Computers, Environment, and Urban Systems*(In review).

Guzman, L. A., Oviedo, D., & Rivera, C. (2017). Assessing equity in transport accessibility to work and study: The Bogotá region. *Journal of Transport Geography, 58*, 236-246.

Habermann, M., Billger, M., & Haeger-Eugensson, M. (2015). Land use regression as method to model air pollution. Previous results for Gothenburg/Sweden. *Procedia Engineering, 115*, 21-28.

Harner, J., Warner, K., Pierce, J., & Huber, T. (2002). Urban environmental justice indices. *The professional geographer, 54*(3), 318-331.

Harper, S., Charters, T. J., & Strumpf, E. C. (2015). Trends in socioeconomic inequalities in motor vehicle accident deaths in the United States, 1995–2010. *American journal of epidemiology, 182*(7), 606-614.

Hatzopoulou, M., Miller, E., & Santos, B. (2007). Integrating vehicle emission modeling with activity-based travel demand modeling: case study of the Greater Toronto, Canada, Area. *Transportation Research Record: Journal of the Transportation Research Board*(2011), 29-39.

Hatzopoulou, M., & Miller, E. J. (2010). Linking an activity-based travel demand model with traffic emission and dispersion models: transport's contribution to air pollution in Toronto. *Transportation Research Part D: Transport and Environment, 15*(6), 315-325.

Havard, S., Deguen, S., Zmirou-Navier, D., Schillinger, C., & Bard, D. (2009). Traffic-related air pollution and socioeconomic status: a spatial autocorrelation study to assess environmental equity on a small-area scale. *Epidemiology, 20*(2), 223-230.

Health Effects Institute Panel. (2010). *Traffic-related air pollution: a critical review of the literature on emissions, exposure, and health effects*. Retrieved from

Iacono, M., Krizek, K. J., & El-Geneidy, A. (2010). Measuring non-motorized accessibility: issues, alternatives, and execution. *Journal of Transport Geography, 18*(1), 133-140.

Jang, S., An, Y., Yi, C., & Lee, S. (2017). Assessing the spatial equity of Seoul's public transportation using the Gini coefficient based on its accessibility. *International Journal of Urban Sciences, 21*(1), 91-107.

Jerrett, M., Arain, A., Kanaroglou, P., Beckerman, B., Potoglou, D., Sahsuvaroglu, T., . . . Giovis, C. (2005). A review and evaluation of intraurban air pollution exposure models. *Journal of exposure science and environmental epidemiology, 15*(2), 185.

Kaplan, S., Popoks, D., Prato, C. G., & Ceder, A. A. (2014). Using connectivity for measuring equity in transit provision. *Journal of Transport Geography, 37*, 82-92.

Kawabata, M. (2009). Spatiotemporal dimensions of modal accessibility disparity in Boston and San Francisco. *Environment and Planning A, 41*(1), 183-198.

Kraemer, J. D., & Benton, C. S. (2015). Disparities in road crash mortality among pedestrians using wheelchairs in the USA: results of a capture–recapture analysis. *BMJ open, 5*(11), e008396.

Kravetz, D., & Noland, R. (2012). Spatial analysis of income disparities in pedestrian safety in northern New Jersey: is there an environmental justice issue? *Transportation Research Record: Journal of the Transportation Research Board*(2320), 10-17.

Lee, J., & Madanat, S. (2017). Optimal design of electric vehicle public charging system in an urban network for Greenhouse Gas Emission and cost minimization. *Transportation Research Part C: Emerging Technologies, 85*, 494-508.

Levin, M. W. (2017). Congestion-aware system optimal route choice for shared autonomous vehicles. *Transportation Research Part C: Emerging Technologies, 82*, 229-247.

Levy, J. I., Chemerynski, S. M., & Tuchmann, J. L. (2006). Incorporating concepts of inequality and inequity into health benefits analysis. *International journal for equity in health, 5*(1), 2.

Levy, J. I., Greco, S. L., Melly, S. J., & Mukhi, N. (2009). Evaluating efficiency-equality tradeoffs for mobile source control strategies in an urban area. *Risk Analysis, 29*(1), 34-47.

Li, X., Ma, J., Cui, J., Ghiasi, A., & Zhou, F. (2016). Design framework of large-scale one-way electric vehicle sharing systems: A continuum approximation model. *Transportation Research Part B: Methodological, 88*, 21-45.

Litman, T. (2002). Evaluating transportation equity. *World Transport Policy & Practice, 8*(2), 50-65.

Ma, J., Li, X., Zhou, F., & Hao, W. (2017). Designing optimal autonomous vehicle sharing and reservation systems: A linear programming approach. *Transportation Research Part C: Emerging Technologies, 84*, 124-141.

Ma, J., Li, X., Zhou, F., Hu, J., & Park, B. B. (2017). Parsimonious shooting heuristic for trajectory design of connected automated traffic part II: computational issues and optimization. *Transportation Research Part B: Methodological, 95*, 421-441.

Martin, E., Shaheen, S. A., & Lidicker, J. (2010). Impact of carsharing on household vehicle holdings: Results from North American shared-use vehicle survey. *Transportation Research Record, 2143*(1), 150-158.

Mishra, S., Welch, T. F., & Jha, M. K. (2012). Performance indicators for public transit connectivity in multi-modal transportation networks. *Transportation Research Part A: Policy and Practice, 46*(7), 1066-1085.

Morency, P., Gauvin, L., Plante, C., Fournier, M., & Morency, C. (2012). Neighborhood social inequalities in road traffic injuries: the influence of traffic volume and road design. *American journal of public health, 102*(6), 1112-1119.

Mortazavi, S. A. H., & Akbarzadeh, M. (2017). A Framework for Measuring the Spatial Equity in the Distribution of Public Transportation Benefits. *Journal of Public Transportation, 20*(1), 3.

Noland, R. B., Klein, N. J., & Tulach, N. K. (2013). Do lower income areas have more pedestrian casualties? *Accident Analysis & Prevention, 59*, 337-345.

Pal, A., & Zhang, Y. (2017). Free-floating bike sharing: solving real-life large-scale static rebalancing problems. *Transportation Research Part C: Emerging Technologies, 80*, 92-116.

Pettit, P. (1974). A theory of justice? *Theory and Decision, 4*(3-4), 311-324.

Ricciardi, A. M., Xia, J. C., & Currie, G. (2015). Exploring public transport equity between separate disadvantaged cohorts: a case study in Perth, Australia. *Journal of Transport Geography, 43*, 111-122.

Rijnders, E., Janssen, N., Van Vliet, P., & Brunekreef, B. (2001). Personal and outdoor nitrogen dioxide concentrations in relation to degree of urbanization and traffic density. *Environmental Health Perspectives, 109*(Suppl 3), 411.

Rowangould, G. M. (2013). A census of the US near-roadway population: Public health and environmental justice considerations. *Transportation Research Part D: Transport and Environment, 25*, 59-67.

Rowangould, G. M. (2015). A new approach for evaluating regional exposure to particulate matter emissions from motor vehicles. *Transportation Research Part D: Transport and Environment, 34*, 307-317.

Ryan, P. H., & LeMasters, G. K. (2007). A review of land-use regression models for characterizing intraurban air pollution exposure. *Inhalation toxicology, 19*(sup1), 127-133.

Sánchez-Cantalejo, C., Ocana-Riola, R., & Fernández-Ajuria, A. (2008). Deprivation index for small areas in Spain. *Social Indicators Research, 89*(2), 259-273.

Saviskas, S., & Sohn, P. (2015). *Bikeshare and equity in Berkeley, CA.* Paper presented at the 94th annual meeting of the Transportation Research Board, Washington, DC.

Shaheen, S., Guzman, S., & Zhang, H. (2012). Bikesharing across the globe. *City cycling, 183*.

Shellooe, S. D. (2013). Wheels When Who Wants Them: Assessing Social Equity and Access Implications of Carsharing in NYC.

Sider, T., Hatzopoulou, M., Eluru, N., Goulet-Langlois, G., & Manaugh, K. (2015). Smog and socioeconomics: an evaluation of equity in traffic-related air pollution generation and exposure. *Environment and Planning B: Planning and Design, 42*(5), 870-887.

Singleton, P. A., & Clifton, K. J. (2017). Considering health in US metropolitan long-range transportation plans: A review of guidance statements and performance measures. *Transport Policy, 57*, 79-89.

Steinbach, R., Green, J., Edwards, P., & Grundy, C. (2010). 'Race'or place? Explaining ethnic variations in childhood pedestrian injury rates in London. *Health & place, 16*(1), 34-42.

Steinbach, R., Green, J., Kenward, M. G., & Edwards, P. (2016). Is ethnic density associated with risk of child pedestrian injury? A comparison of inter-census changes in ethnic populations and injury rates. *Ethnicity & health, 21*(1), 1-19.

Stroh, E., Oudin, A., Gustafsson, S., Pilesjö, P., Harrie, L., Strömberg, U., & Jakobsson, K. (2005). Are associations between socio-economic characteristics and exposure to air pollution a question of study area size? An example from Scania, Sweden. *International Journal of Health Geographics, 4*(1), 30.

Stuart, A. L., Mudhasakul, S., & Sriwatanapongse, W. (2009). The social distribution of neighborhood-scale air pollution and monitoring protection. *Journal of the Air & Waste Management Association, 59*(5), 591-602.

Stuart, A. L., & Zeager, M. (2011). An inequality study of ambient nitrogen dioxide and traffic levels near elementary schools in the Tampa area. *Journal of environmental management, 92*(8), 1923-1930.

Tian, N., Goovaerts, P., Zhan, F. B., & Wilson, J. G. (2010). Identification of racial disparities in breast cancer mortality: does scale matter? *International Journal of Health Geographics, 9*(1), 35.

Tian, N., Xue, J., & Barzyk, T. M. (2013). Evaluating socioeconomic and racial differences in traffic-related metrics in the United States using a GIS approach. *Journal of exposure science and environmental epidemiology, 23*(2), 215.

U.S. EPA. (2010a). *Emissions Modeling Clearinghouse Biogen- ic Emission Sources*. Retrieved from

U.S. EPA. (2010b). *Motor Vehicle Emission Simulator (MOVES) User Guide for MOVES2010a*. Retrieved from

Ursaki, J., & Aultman-Hall, L. (2016). *Quantifying the equity of bikeshare access in US cities.* Paper presented at the 95th Annual Meeting of the Transportation Research Board, Washington, DC.

Wang, R., Henderson, S. B., Sbihi, H., Allen, R. W., & Brauer, M. (2013). Temporal stability of land use regression models for traffic-related air pollution. *Atmospheric Environment, 64*, 312-319.

Welch, T. F., & Mishra, S. (2013). A measure of equity for public transit connectivity. *Journal of Transport Geography, 33*, 29-41.

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., & Goodman, A. (2014). Health effects of the London bicycle sharing system: health impact modelling study. *Bmj, 348*, g425.

Yu, C.-Y. (2014). Environmental supports for walking/biking and traffic safety: income and ethnicity disparities. *Preventive medicine, 67*, 12-16.

Yu, H., & Stuart, A. L. (2013). Spatiotemporal distributions of ambient oxides of nitrogen, with implications for exposure inequality and urban design. *Journal of the Air & Waste Management Association, 63*(8), 943-955.

Yu, H., & Stuart, A. L. (2016). Exposure and inequality for select urban air pollutants in the Tampa Bay area. *Science of the Total Environment, 551*, 474-483.

Yu, H., & Stuart, A. L. (2017). Impacts of compact growth and electric vehicles on future air quality and urban exposures may be mixed. *Science of the Total Environment, 576*, 148-158.

Zhang, Y., & Lin, G. (2013). Disparity surveillance of nonfatal motor vehicle crash injuries. *Traffic injury prevention, 14*(7), 697-702.

Part 2

Exploring the equity performance of bike-sharing systems with disaggregated data: A story of southern Tampa

1. Introduction

The very first bike-sharing system appeared in Amsterdam in 1965 but collapsed quickly due to vehicle damage and theft. The next generation, the coin-deposit system, was launched in Farsø and Grenå, Denmark in 1991 but was not warmly embraced as the theft issue was still unsolved. The third generation, also known as the IT-based bike-sharing system, did not appear until 1996. These systems adopt advanced IT technologies (e.g., smart cards, digital docking systems) and usually come with densely deployed infrastructures, and consequentially won great popularity several years after its first appearance in England. The latest generation, i.e., the free-floating bike-sharing system, incorporates more sophisticated technologies (e.g., GPS bike tracking, smartphone applications, redistribution innovation) and thus further promotes the adoption of bike-sharing systems (Shaheen et al., 2010). Nowadays, bike-sharing has become one of the most fast-growing transportation modes all over the world (Schmidt, 2018). As of the end of 2016, the number of cities that were operating a bike-sharing system had increased to around 1000 all over the world (Wikipedia, 2018), with China owning the largest bike fleet. In the United States, the number of shared bikes had grown from 42,500 at the end of 2016 to around 100,000 by the end of 2017 (NACTO, 2018), together with 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 the great success, bike-sharing systems are shown to bring significant benefits to individuals and society as a whole. By either providing stand-alone service or working as a solution to the first/last mile problem in public transit, bike-sharing systems can reduce our dependence on private automobiles and bolster public transit usage, therefore reducing the fossil fuel consumption and tailpipe emission (Zhang and Mi, 2018). Being an active transportation mode, bike-sharing induces more physical activities from individuals, which then brings positive health impacts overall (Woodcock et al., 2014). Further, not as intuitive as its environmental and public health benefits, the promotion of bike-sharing also contributes significantly to the economic development through various ways such as saving travel time, creating job

opportunities, reducing household transportation expenses and booming the tourism industry (Castro, 2011).

Against the proliferation of the bike-sharing system and all its positive impacts, however, more and more people have come to question its equity impacts, specifically, whether benefits brought by the bike-sharing system are distributed among the society in a fair and reasonable manner, especially for disadvantaged population groups. Indeed, surveys have shown that equity impacts are a real problem in some bike-sharing systems. For example, in Washington D.C., black people account for around 50% of the population but only 4% of the Capital Bikeshare membership in 2016 (Benjamin, 2017). In light of this issue, many operators and administrators of bike-sharing programs have initiated efforts to overcome the user barriers and address the inequality issues. A survey on 20 ongoing or planned bike-sharing programs in the U.S. (Buck, 2013) found that, to lower access barriers, many bike-sharing programs had implemented or were intended to implement some countermeasures, e.g. stations in diverse neighborhoods, income-based discount programs (NACTO, 2018), and community outreach campaigns (Mcneil, 2015), to name a few. Meanwhile, research funding has also been awarded to explore the answer to this question. Yet to date related studies are still very limited.

Despite substantial efforts in practice and a handful of pioneering studies on bike-sharing equity, there is still not yet a comprehensive framework to evaluate the equity performance of bike-sharing systems. Thus, this paper proposes a methodological framework for quantitatively accessing the equity performance of bike-sharing systems with disaggregated data, using southern Tampa as a case study. Different from previous studies, this framework considers disaggregated individual data and the accessibility that individuals obtained from a bike-sharing system. In other words, we study how accessibility from a bike-sharing system is distributed among individuals in society. Following this idea, a full synthetic population, not a small sample or aggregated zonal level data, in southern Tampa is utilized for the analysis. With disaggregated data, the proposed method unveils important messages that might be absorbed by existing methods with aggregated data and thus avoids misleading our understanding of equity. Further, to measure the benefits bought by bike-sharing systems, we propose an individual bike-sharing accessibility model that incorporates the unique operational characteristics of bike-sharing (i.e., walking-cycling-walking) and trip chaining in an individual's daily travel itineraries. The consideration of these factors makes the model avoid an overestimation of the bike-sharing accessibility and therefore allows us a better understanding of equity impacts. Experimental results verify the validity and necessity of the proposed framework and also draw some interesting managerial insights that can assist the bike-sharing operator in determining their future expansion plan for southern Tampa.

The remainder of this paper is organized as follows. Section 2 provides related studies and the unique contributions of this study. Section 3 introduces the study context and the datasets used in this study. The

general methodological framework is discussed in detailed in Section 4. Section 5 presents the experiment results to validate the proposed framework and draw some managerial insights. Finally, Section 6 concludes the paper and briefly discusses some potential future research directions.

2. Literature review

Equity has been a classical topic in transportation studies whose history dates back to the Civil Rights Act of 1964, with abundant research performed on transportation equity assessment (Welch and Mishra, 2013). According to Litman (2002), transportation equity can be divided into three categories: horizontal equity, vertical equity with regard to income and social class, and vertical equity with regard to mobility need and ability. Horizontal equity is the most frequently studied perspective; it requires each individual or group to be treated with the same distribution of costs or benefits and to bear costs proportionate to the benefits they receive (Litman, 2002). For vertical equity with regard to income and social class, it is more equitable if policies favor economically- and socially-disadvantaged groups (Pettit, 1974). Vertical equity with regard to mobility need and ability requires the needs of individuals or groups with impaired mobility are satisfied (Litman, 2002). In this study, we consider both horizontal equity and vertical equity with regard to income and social class. For the convenience of illustration, hereafter we call this latter type simply vertical equity.

Following the above definitions, different methods have been proposed for analyzing the equity performance of a transportation system. Generally speaking, horizontal equity is measured in terms of geographic areas or population aggregated to a specific geospatial scale (due to the lack of individual-level data). Popular methods for horizontal equity analysis include applications of the Lorenz curve and Gini index (Delbosc and Currie, 2011; Guzman et al., 2017; Kaplan et al., 2014; Lucas et al., 2016; Welch and Mishra, 2013), Atkinson index (Levy et al., 2009), geographic mapping analysis (Kaplan et al., 2014), etc. Regarding vertical equity, the analysis usually makes intergroup comparisons of costs and benefits to different socioeconomic groups categorized by income level, education level, race and/or ethnicity, etc. Frequently adopted methods include distribution comparison with basic descriptive statistics (Boarnet et al., 2017), environmental justice index (Harner et al., 2002), subgroup inequality index (Stuart et al., 2009; Yu and Stuart, 2016, 2013), ANOVA test (El-Geneidy et al., 2016; Sider et al., 2015), regression models (Goodman et al., 2011; Harper et al., 2015) and many more. In all these approaches, data are typically aggregated to a specific geospatial scale due to the lack of individual-level data; scale units have included traffic analysis zones (Mishra et al., 2012), census tracts (Boarnet et al., 2017), municipalities (Oswald Beiler and Mohammed, 2016). However, 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, the method proposed by Bills and Walker (2017) does not consider the situation where the benefit distribution is highly skewed, which could render the proposed individual difference density comparison difficult.

Further, despite the extensive studies on transportation equity, only a few have investigated bike-sharing. For example, to evaluate the equity of people's accessibility to bike-sharing stations, Ursaki and Aultman-Hall (2016) used a Student's t-test to compare social and economic characteristics of census block groups that are within and outside the service areas for eight U.S. bike-sharing programs. Likewise, Gavin et al. (2016) compared bike-share membership survey data with census residence characteristics within bicycle service areas for three cities and concluded that users are more likely to be residents who are male, young, white, affluent, and educated. Although there are vast disparities of accessibility to bike-sharing, Saviskas and Sohn (2015) surveyed populations in Berkeley and concluded that low-income and high-income people had similar levels of interest in using bike-sharing. Though these studies offer us simple and useful methods to study equity impacts of bike-sharing programs, they fail to consider how individuals' accessibility may change because of the inception of the bike-sharing systems.

Thus, to analyze equity impacts of bike-sharing, an individual bike-sharing accessibility model is necessary, which leads us to related studies in transit accessibility modeling. Previous methods modeled transit equity from two aspects. One is the coverage-based approach, which treats transit stations as travelers' destinations and quantifies travelers' accessibility to the transit system as the proportion of areas or population that can be served by the public transit system in the geographic unit of analysis (Currie, 2010; El-Geneidy et al., 2010; Murray, 2001). These measures can offer a simple and intuitive metric to evaluate the structure of a transit network, but cannot capture its spatial-temporal connectivity and fail to consider traveler's travel demand (Nassir et al., 2016). To address these issues, the other aspect, the reachability-based approach, considers the travelers' O (origin) – D (destination) pairs and models the transit accessibility as a decreasing function of the travel impedances with estimated travel time (Kawabata and Shen, 2006; Liu and Zhu, 2004; Moniruzzaman and Páez, 2012; O'Sullivan et al., 2000), 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. In contrast to a large body of literature on public transit accessibility modeling, studies modeling bike-sharing accessibility are more limited.

As bike-sharing systems are becoming increasingly popular, there is an imperative need for a comprehensive equity assessment framework for bike-sharing accessibility. This study aims to bridge this gap between the soar of the bike-sharing industry and the lack of a sophisticated equity assessment

methodology. This study makes a number of contributions to the existing literature. First, we propose a bike-sharing equity assessment framework that considers both disaggregated data and the individual accessibility that people obtain from bike-sharing systems. This framework can be used for assessing both horizontal and vertical equity. Second, the individual bike-sharing accessibility model incorporates the unique operational characteristics of bike-sharing (i.e., walking-cycling-walking) and trip chaining in an individual's daily travel itinerary. Finally, the proposed methodological framework is applied to the Coast Bike Share System in southern Tampa, which not only demonstrates the application of the proposed framework but also draws interesting managerial insights.

3. Study context and data collection

This section presents the study context and data collection for this study. An overview of the study area is first presented, followed by a description of the data collection and preparation process.

3.1. Study area

This study area is the southern part of Tampa, which locates in the south of the largest city in the Tampa Bay Area (see Fig. 1 (a), (b)). The area is 57.7 square miles in size with 167,992 people in 2017. Since its inception in late 2014, Coast Bike Share (a for-profit bike-sharing service provider in the Tampa Bay Area) has been running an independent bike-sharing system in Downtown Tampa, the central business district (CBD) in this area (see Fig. 1 (c)), with a total fleet size of around 130 at 42 stations (http://coastbikeshare.com/). Reports reveal that this system has brought significant benefits to the city, for example, improved accessibility, reduced traffic congestion and saved parking space, to name a few. However, the beneficiaries of the Coast Bike Share system are very limited, since, as can be seen from Fig. 1 (c), a large portion of the investigated area is still beyond the service area of the Coast Bike System. This naturally raises the question of whether the benefit distribution of the Coast Bike Share systems is equal among different geographical units in southern Tampa.
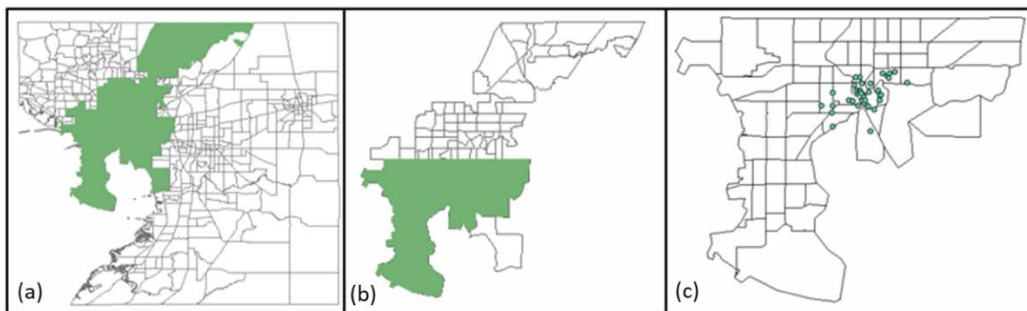

Fig. 1: (a) Location of the City of Tampa; (b) Location of study area of southern Tampa; (c) The Coast Bike Share System in southern Tampa

43

The area is also an excellent testbed for investigating whether the benefit distribution is equal among different sociodemographic groups because of its sociodemographic diversity. According to the US Census Bureau, female accounts for 48% of the total population and the age distribution in the city consists of 18.0% under 18 years, 68% between 18 and 64, and 11.8% over 65. The white, black and Asian racial categories composed 59%, 10% and 4.0% of the population, respectively, 25% of which were Hispanic or Latino origin. Finally, 15.3% of the households live below the poverty line while 31% earn more than $100K per year.

3.2. Data collection and preparation

We use land parcel as the geographic unit of analysis. Consider a set of parcels indexed as $p \in \mathcal{P} := [1,2,\cdots,P]$ and a set of individuals indexed as $i \in \mathcal{I} := [1,2,\cdots,I]$ residing in these parcels. Three datasets are needed for the disaggregated modeling approach proposed in this paper, as follows:

Bike-sharing provision: Let $\mathcal{B} := [1,2,\cdots B]$ be the set of bike-sharing facilities (i.e., bike-sharing stations for station-based systems and potential parking spots for free-floating systems) in the investigated area. With the coordinates of each bike-sharing facility $b \in \mathcal{B}$ provided by the Coast Bike Share and those of the centroids of each parcel, we compute the distances between each bike sharing facility and the centroid of each parcel, denoted as $d_{bp}, \forall b \in \mathcal{B}, p \in \mathcal{P}$. Then, we can compute individuals' willingness to walk to bike-sharing facilities at each parcel, denoted as w_p, with the distance decay function for walking $f(d_{bp})$ as follows:

$$w_p := \max_{b \in \mathcal{B}} f\left(d_{bp}\right), \forall p \in \mathcal{P}. \tag{1}$$

Note that the maximum value among all bike-sharing facilities is adopted because the bike-sharing service is usable to individuals within a parcel as long as one bike-sharing facility is accessible to them. One common example of the distance decay function for walking is that $f(d_{bp}) = \alpha_1 e^{-\alpha_2 d_{bp}}$, where α_1 and α_2 are parameters that should be calibrated with empirical data (Hochmair, 2015).

Individual travel demand: Previous studies show that despite random deviations, individual mobility patterns in urban space show certain regularity (Jiang et al., 2016; Jiang et al., 2017; Schneider et al., 2013), and therefore we model an individual's travel demand as her regular itinerary. Essentially, an individual's daily travel itinerary can be defined as a sequence of consecutive trips indexed as $n \in \mathcal{N}_i := [1,2,\cdots,N_i]$, where N_i is the number of trips that individual i commences over a day. Let p_{in}^- and p_{in}^+ be the origin and destination of individual i's n-th trip, respectively. Then, individual i's itinerary can be defined as $\mathcal{T}_i := \{(p_{in}^-, p_{in}^+), \forall n \in \mathcal{N}_i\}, \forall i \in \mathcal{I}$, where $p_{in}^+ = p_{i(n-1)}^-, \forall n \in \mathcal{N}_i \backslash \{1\}$. This dataset can be generated in a variety of ways such as household travel survey (Jiang et al., 2016; Gurram et al., 2015), inference from cell record

data (Jiang et al., 2016), inference from social media data (Hasan and Ukkusuri, 2014), agent-based simulation (Gurram, 2017) and so on. In this paper, we used the daily activity and travel itineraries of individuals in the study region that were simulated by Gurram (2017) using the Person Day Activity and Travel Simulator (Daysim) developed by Bradley et al. (2010). Readers are referred to Gurram (2017) for the detailed simulation mechanism and process. From the simulation results, we use the travel distance of each trip, denoted as $d_{in}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$, to compute the willingness to cycle for that trip, denoted as c_{in}, with the distance decay function for cycling as follows:

$$c_{in} \coloneqq g(d_{in}), \forall i \in \mathcal{I}, n \in \mathcal{N}_i. \qquad (2)$$

One example of the distance decay function for cycling is that $g(d_{in}) = \alpha_3 e^{-\alpha_4 d_{in}}$, where α_3 and α_4 are calibrated parameters (Hochmair, 2015).

Individual geographical/sociodemographic attributes: For each individual $i \in \mathcal{I}$, geographical and sociodemographic attributes are needed. In this paper, we consider two geographical attributes: the parcel and traffic analysis zone (TAZ) in which individuals reside; and five sociodemographic attributes: age group (0-18, 18-45, 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, upper income with an annual household income above $75,000), race (white, black, Asian, other), and ethnicity (Hispanic and non-Hispanic). Please note that larger geographic units such as census tract are not considered since with the existing bike-sharing system, the zonal bike-sharing accessibility cannot show significant differences in such a large analysis unit. Because real individual-level sociodemographic data are not available due to privacy reasons, we used data on hypothetical individuals to represent the population in the study area; these data were generated by Gurram (2017) using an iterative proportional fitting approach (Beckman et al., 1996) based on the 2010 census data (US Census Bureau). Interested readers can refer to Gurram (2017) for the detailed information. We will show that, in the following section, with these individual attributes, we can aggregate the individual-level measures into different geographic or sociodemographic group-level measures, which then enables the equity analysis on different levels as needed.

4. Methodology

This section proposes a new approach to evaluating the equity performance of bike-sharing systems with disaggregated data, i.e. individual-level data. A tour-based individual bike-sharing accessibility modelling method is first presented. Based on this method, we will then discuss how to analyze the equity performances of the bike-sharing systems with the disaggregated data.

4.1. Tour-based individual bike-sharing accessibility modeling

This subsection mathematically formulates the individual bike-sharing accessibility.

4.1.1. Bike-sharing accessibility modeling for a single trip

We first model the bike-sharing accessibility for a single trip $\{p_{in}^-, p_{in}^+\}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$. As shown in Fig. 2, such a trip is essentially comprised of three consecutive steps: (i) Walking to pick up a bicycle at a bike-sharing facility $b \in \mathcal{B}$ near her origin p_{in}^-; (ii) Cycling from b to another bike-sharing facility $b' \in B$ near her destination p_{in}^+; (iii) Returning the bicycle at b' and then walking to p_{in}^+. As mentioned before, existing analyses on the equity dimension of bike-sharing systems usually assume that only the population residing within the service area of a bike-sharing facility enjoys the accessibility to bike-sharing (e.g., Gavin et al., 2016). Regardless of its intuition and simplicity, this method cannot precisely capture how accessibility changes with the distance to a bike-sharing facility. In light of these issues, we propose a measure that takes into account all three steps in a trip.
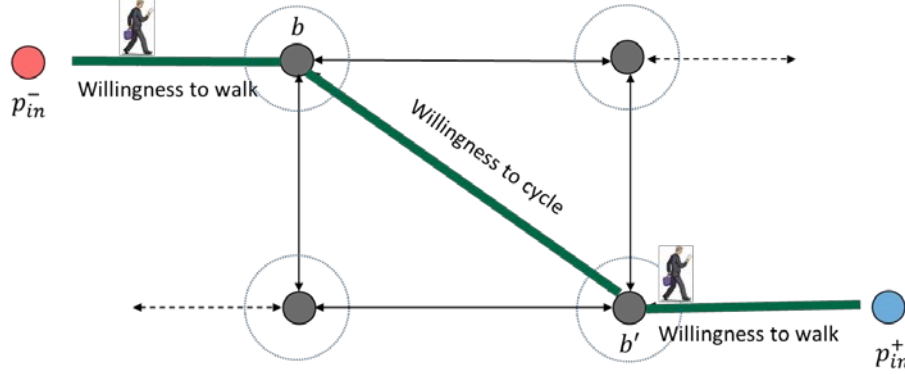

Fig. 2: The walking-cycling-walking process of a bike-sharing trip

For the walking process, we use an individual's accessibility to bike-sharing facilities at her origin and destination, $w_{p_{in}^-}$ and $w_{p_{in}^+}$, respectively, as her willingness to walk to the bike-sharing facilities. These two measures can be easily obtained with an enumeration process over the set of parcels. For the cycling process, since if individual i can access p_{in}^+ from p_{in}^- with shared bicycles is dependent on her willingness to cycle between these two parcels, the cycling accessibility can be easily obtained through Eq. (2). As the trip is a consecutive process, the bike-sharing accessibility for a single trip $\{p_{in}^-, p_{in}^+\}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$, denoted as $a_{in}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$, can be formulated as the product of the accessibility to bike-sharing facilities at the origin and destination parcels as well as the cycling accessibility between these parcels, i.e.,

$$a_{in} = w_{p_{in}^-} c_{in} w_{p_{in}^+}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i. \tag{3}$$

4.1.2. Tour-based bike-sharing accessibility modeling for an individual

With the bike-sharing accessibility for a single trip $\{p_{in}^-, p_{in}^+\}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$, we are now ready to model the accessibility for the entire travel itinerary $\mathcal{T}_i, \forall i \in \mathcal{I}$. Before modeling, we first use an illustrative example to highlight the need of considering tours in modeling the individual bike-sharing accessibility. In Fig. 3, nodes 1 through 5 represent home, convenience store, work place, restaurant and shopping mall, respectively, and bike-sharing facilities are located at each node. The bike-sharing accessibility of each trip are also 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 0.2, respectively, which are relatively high values. However, these results may not be realistic in practice. The traveler likely drives for trip (1, 2) considering that she has to drive to work (i.e. trip (2, 3)) after this trip. Likewise, trip (5, 1) is likely to be commenced by car since she might have to drive for the previous 2 trips. Therefore, the resulting bike-sharing accessibility is overestimated by the trip-based approach in both situations. To address this drawback, we propose a tour-based approach blow considering a traveler's trip chaining for more realistic evaluations.
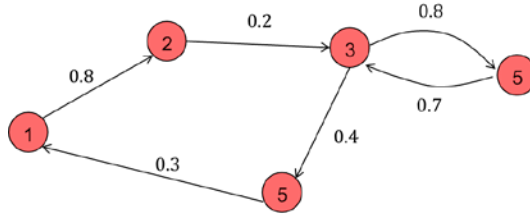

Fig. 3: An illustrative example for the necessity for the tour-based analysis

The first step of the tour-based approach is to break the travel itinerary $\mathcal{T}_i, \forall i \in \mathcal{I}$ into a set of subtours indexed as $m \in \mathcal{M}_i, \forall i \in \mathcal{I}$ with a subtour generation algorithm (Algorithm 1). In this algorithm, we first define the sequence of visited locations of individual $i \in \mathcal{I}$ as $\mathcal{X}_i$ and remove its repeated elements to obtain her set of activity locations $\mathcal{O}_i$. Afterwards, we define and initialize five variables or sets, including individual i's set of tours $\mathcal{M}_i$, the number of times that o has been visited till the current iteration $E_o, \forall o \in \mathcal{O}_i$, set of o's indexes that has been checked till the current iteration $F_o, o \in \mathcal{O}_i$, index of current checking location s, and tour index m. With these, we then iterate sets $\mathcal{X}_i$ and $\mathcal{O}_i$ in an outer and inner loop, respectively, to divide the travel itinerary into multiple subtours as follows: (i) Check if the current checking location x is the same as an activity location o. If yes, we increase the number of times that o has been visited, i.e., E_o, by 1 and add the index of the current checking location, i.e., s, into set F_o; and if not, move on to the next step. (ii) Check if o has been visited more than one time (i.e., $E_o > 1$). If yes, we update the tour index m by 1 and add it to the set of tours $\mathcal{M}_i$. Since F_o records the indexes that o has been checked so far, $\min(F_o) + 1$ and $\max(F_o)$ actually represent indexes of the first and final visited locations in tour m. Thus, we find all the visited locations between $\min(F_o)$ and $\max(F_o)$ (including $\max(F_o)$), i.e.,

$\{n | n \in \mathcal{N}_i, \min(F_o) < n \le \max(F_o)\}$, and add them to set $\mathcal{N}_{im}$. Because tour m has been extracted, we next remove all trips in $\mathcal{N}_{im}$ from $\mathcal{N}_i$. Further, for the last location in tour m, we reinitialize E_o and F_o as 1 and $\max(F_o)$, respectively, because next tour starts from this location. For all other locations in tour m, E_o and F_o are reinitialized as 0 and an empty set, respectively. If not, move on to the next iteration of the inner loop. When all $o \in \mathcal{O}_i$ are visited, update the index of the current checking location s by 1 and then move on to the next iteration of the outer loop.

With this, we can compute the bike-sharing accessibility for each subtour $m \in \mathcal{M}_i$ as the product of the accessibility of all trips in that tour considering the chaining of these trips, i.e.,

$$a_{im} = \prod_{n \in \mathcal{N}_{im}} a_{in}, \forall i \in \mathcal{I}, m \in \mathcal{M}_i. \tag{4}$$

Then, the bike-sharing accessibility of an individual $i \in \mathcal{I}$, denoted as $a_i, \forall i \in \mathcal{I}$, can be formulated as the average of the accessibility of all her subtours, i.e.

$$a_i = \frac{\sum_{m \in \mathcal{M}_i} a_{im}}{|\mathcal{M}_i|}, \forall i \in \mathcal{I}. \tag{5}$$

Algorithm 1. Subtour Generation Algorithm

  Input: $\mathcal{I}, \mathcal{N}_i, \mathcal{T}_i, i \in \mathcal{I}$

1. **for** $i \in \mathcal{I}$

2. $\mathcal{X}_i \leftarrow \{p_{i1}^-\} \cup \{p_{in}^+ | n \in \mathcal{N}_i\}$; // generate the sequence of visited locations

3.   Remove repeated elements from $\mathcal{X}_i$, resulting in $\mathcal{O}_i$ // generate the set of activity locations

4. $\mathcal{M}_i \leftarrow \emptyset$; //initialize the set of tours as an empty set

5. $E_o \leftarrow 0, \forall o \in \mathcal{O}_i$ // initialize the number of visits till the current iteration for o as 0

6. $F_o \leftarrow \emptyset, \forall o \in \mathcal{O}_i$ // initialize set of o's indexes that has been checked till the current iteration as $\emptyset$

7. $s \leftarrow 0; m \leftarrow 0$ //initiate index of the current checking location and tour index as 0

8.   **for** $x \in \mathcal{X}_i$

9.   **for** $o \in \mathcal{O}_i$

10.   **if** $x = o$ // if the current checking location x is the same as activity location o

11. $E_o \leftarrow E_o + 1$ // increase the number of times that o has been visited by 1

12. $F_o \leftarrow F_o \cup s$ // Add the index of the current checking location into set F_o

13.   **end if**

14.   **if** $E_o > 1$ // if o has been visited twice, meaning a tour has been completed

15. $m \leftarrow m + 1; \mathcal{M}_i \leftarrow \mathcal{M}_i \cup \{m\}$; // update the tour index by 1 and add it into set $\mathcal{M}_i$

16. $\mathcal{N}_{im} \leftarrow \{n | n \in \mathcal{N}_i, \min(F_o) < n \le \max(F_o)\}$ // obtain the visited locations in tour m

17. $\mathcal{N}_i \leftarrow \mathcal{N}_i \backslash \mathcal{N}_{im}$ // remove all locations in tour m from the set of visited location $\mathcal{N}_i$

18.	$E_o \leftarrow 1, F_o \leftarrow \max(F_o)$, $o = \{p_{in}^+ \mid n \in \mathcal{N}_i, n = \min(F_o)\}$ // for the last location in	
19.	tour m, update E_o as 1 and F_o as $\max(F_o)$	
20.	$E_o \leftarrow 0, F_o \leftarrow \emptyset$, $o \in \{p_{in}^+ \mid n \in \mathcal{N}_i, \min(F_o) < n < \max(F_o)\}$ // for other locations in	
21.	tour m, update E_o as 0 and F_o as an empty set	
22.	**end if**	
23.	**end for**	
24.	$s \leftarrow s + 1$ // update index of the current checking location	
25.	**end for**	
26. **end for**		

Output: $\mathcal{M}_i, \forall i \in \mathcal{I}, \mathcal{N}_{im}, \forall m \in \mathcal{M}_i$

4.2. Equity analysis

To understand the distribution of bike-sharing accessibility, we integrate the individual accessibility measures, geographic and sociodemographic attributes to perform a few equity analyses from both the horizontal and vertical equity perspectives. Though various approaches to tackling this problem have been proposed in the literature, few of them takes into account disaggregated data and bike-sharing systems simultaneously. Therefore, in this section we discuss how equity analysis can be carried out with the unique disaggregated measures for bike-sharing systems in this paper.

4.2.1. Horizontal equity

As mentioned previously, horizontal equity can be analyzed from both a geographic and (grouped) population perspective. For the convenience of the illustration, hereafter we name the equity analysis from these two perspectives as spatial equity and population equity, respectively. In general, it is hard to answer the question of whether the benefit/cost distribution is equal among the entire population due to the lack of individual-level data. Thus, previous studies usually use aggregated population data to investigate the population equity. Nevertheless, in this study, we can analyze the population equity with individual-level data. An easy way to reach this end is the application of the Lorenz curve and Gini index (Delbosc and Currie, 2011). Lorenz curves, a graphical analysis tool from economics (Lorenz, 1905), describe the cumulative distribution of accessibility across the population and thus can offer us an intuition on the distribution of the bike-sharing accessibility among the population. In contrast, to obtain an overall quantitative assessment of the population equity, Gini index is necessary. It is a value ranging from 0 to 1, with 1 indicating the most skewed distribution of the bike-sharing accessibility and 0 the most even distribution. Note that the value of Gini index just offers a quantitative description of how concentrated resources are distributed. To evaluate whether the distribution is equitable or not, the planning agencies

objectives must be taken into account. Please refer to Delbosc and Currie (2011) for the mathematical formulation of the Gini index.

Apart from the population equity, the disaggregated measures can also be applied for spatial equity. With the individual geographic information, we can aggregate the individual measures into different zonal-level measures (i.e. parcels and TAZs in this paper), based on which the equity analysis is carried out. In the following we use parcels as an example to illustrate the aggregation process to compute each parcel's accessibility, denoted as $a_p, \forall p \in \mathcal{P}$, using the individual bike-sharing accessibility $a_i, \forall i \in \mathcal{I}$. Since the population varies across parcels, we sum and normalize the accessibility of all individuals within a parcel as its accessibility indicator. More specifically, let $p_i = 1$ if individual $i \in \mathcal{I}$ resides in parcel $p \in \mathcal{P}$ and 0 otherwise. Then the accessibility in parcel $p \in \mathcal{P}$ can be formulated as

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

(6)

Similar aggregations can be applied to larger geographic units such as TAZs as well. With these aggregated accessibility indicators, we then used Kernel Density Tool in Arcgis Toolbox to visualize the distribution of bike-sharing accessibility within the studied area, which has been frequently used for hotspot identification such as crash hotspot (Thakali et al., 2015). Although the geographic mapping analysis offers a quite simple and intuitive way to explore the spatial distribution of the bike-sharing accessibility, it fails to present an overall quantitative metric. Therefore, just as what we do for the population equality analysis, we compute the Gini index to obtain an overall degree of inequality.

4.2.2. Vertical equity

The individual-level accessibility measures can also be used to study the vertical equity. Following previous studies using aggregated data (El-Geneidy et al., 2016), we first conduct some distribution comparisons among different demographic groups using several statistics. A series of ANOVA (analysis of variance) tests are then conducted to investigate whether there are significant differences between the means of the bike-sharing accessibility among different population subgroups.

The above analyses can offer a coarse answer to the vertical equity issue. Nevertheless, important information may be masked by simply using simple summary statistics to describe a distribution. Further, the above analyses cannot quantitatively capture the whole picture describing the disparity with bike-sharing accessibility levels. Thus, following Stuart et al. (2009), we use the subgroup inequality index that is a ratio of subgroup population fractions, specifically the ratio of the fraction of the population of an area with a given benefit/cost level that is a particular population subgroup to the fraction that subgroup comprises of the total population of the whole study area. To calculate this index, the population is divided

into R groups indexed as $r \in \mathcal{R} := [1, 2, \cdots R]$ based on a specific sociodemographic attribute (e.g. race) and the accessibility are divided into A levels indexed as $\bar{a} \in \mathcal{A} := [0, 1, \cdots A]$. Let $r_i = 1$ if individual i belongs to population subgroup $r \in \mathcal{R}$ and otherwise $r_i = 0$. Further, let $\bar{a}_i = 1$ if individual i's accessibility is no less than $\bar{a} \in \mathcal{A}$ and otherwise $\bar{a}_i = 0$. With these two binary variables, we define the fraction of population that belongs to subgroup $r \in \mathcal{R}$ as

$$Y_r := \frac{\sum_{i \in \mathcal{I}} r_i}{I}, \forall r \in \mathcal{R}. \tag{7}$$

Likewise, the fraction of population that belongs to subgroup r with an accessibility level above $\bar{a} \in \mathcal{A}$ can be defined as

$$Y_{r\bar{a}} := \frac{\sum_{i \in \mathcal{I}} r_i \bar{a}_i}{\sum_{i \in \mathcal{I}} \bar{a}_i}, \forall r \in \mathcal{R}. \tag{8}$$

Then, the ratio of $Y_{r\bar{a}}$ to Y_r can be used to describe to what extent members in population subgroup r are disproportionally distributed among the population with a bike-sharing accessibility above level $\bar{a}$. Yet the value of this ratio ranges from 0 to infinity, which causes difficulty in interpreting the results. To address this issue, a log transformation is used to formulate the subgroup inequality index as follows.

$$F_{r\bar{a}} = \log\left(\frac{Y_{r\bar{a}}}{Y_r}\right) \tag{9}$$

where $F_{r\bar{a}}$ quantifies the degree to which members in subgroup $r \in \mathcal{R}$ are disproportionally distributed among the population with bike-sharing accessibility above level $\bar{a} \in \mathcal{A}$. With this formulation, the subgroup inequality index turns out to be easily interpreted. A negative $F_{r\bar{a}}$ indicates that members in subgroup r are disproportionally lowly distributed among the population with bike-sharing accessibility above level $\bar{a}$ while a positive $F_{r\bar{a}}$ just indicates the opposite trend. Finally, an index value of 0 reveals that members in subgroup r are not disproportionally distributed among the population with bike-sharing accessibility above level $\bar{a} \in \mathcal{A}$.

5. Results and analysis

This section presents the experimental results. Section 5.1 presents the horizontal equity analysis results, answering the question that how the bike-sharing accessibility is distributed among the population and the geographic space in southern Tampa, regardless of individual attributes. Section 5.2 compares the proposed method against several benchmark measures adapted from the existing literature, to highlight the necessity and importance of the proposed methodology. Finally, Section 5.3 presents the vertical equity analysis

results, offering an answer to the question of how the bike-sharing accessibility is distributed among different sociodemographic groups.

5.1. Horizontal equity analysis

First, we measure the horizontal equity using the Lorenz curve and Gini index from both the population and geographic perspectives, as shown in Fig. 4. As can be seen from Fig. 4 (a), the distribution of bike-sharing accessibility is highly skewed among the population in southern Tampa, with over 90% of the population having no bike-sharing accessibility at all and around 2% of the population enjoying 50% of the bike-sharing accessibility. This high skewness is also reflected by the large value of the Gini index, i.e., 0.964. Further, Fig. 4 (b) reveals similar results from the view of spatial equity when we use parcels as our geographic unit of analysis, i.e., an extremely skewed Lorenz curve and a Gini index of 0.960. These results indicate that more than 90% of the parcels in southern Tampa does not have bike-sharing accessibility while 50% of the bike-sharing accessibility is concentrated in around 2% of the parcels. Interestingly, the distribution is less skewed when we adopt TAZ as our geographic unit of analysis, with a Gini index of 0.854, indicating that around 75% of the TAZs does not have bike-sharing accessibility. Overall speaking, these analyses show that the improved accessibility benefits thanks to the bike-sharing program in southern Tampa are concentrated in only a small portion of the population and parcels. This result, however, is not very surprising considering the existing deployment scheme of the Coast Bike System in southern Tampa. As we have mentioned in the study context, Coast Bike only deployed stations in downtown Tampa so actually a large portion of the studied and thus the population is not served by the bike-sharing system.
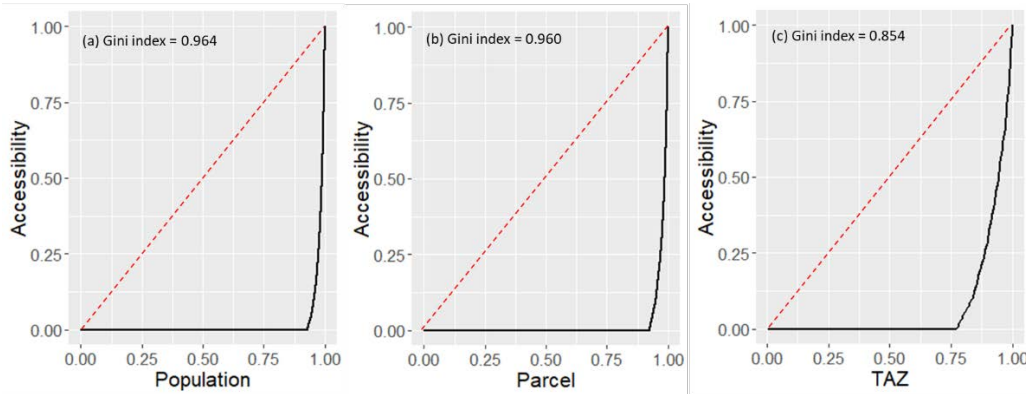

Fig. 4: Lorenz curves and Gini indexes at: (a) the population level; (b) parcel level; (c) TAZ level

Knowing the bike-sharing accessibility is not evenly distributed in southern Tampa, next, we ask the question of where there are inequalities. Put another way, which regions receive higher bike-sharing accessibility and which receive lower? This question can be answered through the geographic mapping

analysis on both the parcel and TAZ levels, as reported in Fig. 5. Not surprisingly, most of southern Tampa's geographic areas have no bike-sharing accessibility, which is consistent with the results from the quantitative method. Areas with bike-sharing accessibility are either within or on the edge of downtown Tampa, where bike-sharing stations are deployed. This observation supports our previous argument that the deployment scheme results in a skewed distribution of the bike-sharing accessibility. Besides, an interesting finding from Fig. 5 (a) is that for the entire area of southern Tampa, the bike-sharing accessibility seems relatively low, even for the downtown area. This may be attributed to, first, the low density of bike-sharing stations. Previous studies (Du and Cheng, 2018) show that distance between bike-sharing stations and traveler's origins/destinations plays an important role in determining the adoption of a bike-sharing system. Though stations have been deployed in southern Tampa, they are still not dense enough so that travelers' willingness to walk to these stations are relatively low. Second, the majority of the residents' daily travel demands in southern Tampa are long-distance travels that biking cannot cover. This may lead to a lower value on individuals' willingness to cycle, and thus the bike-sharing accessibility. We will provide further evidence for these two arguments in the next section. Finally, we want to note that TAZ-level analysis shows a similar trend as that from the parcel-level analysis but with much lower values of the bike-sharing accessibility. However, we can hardly find any significant difference of the bike-sharing accessibility from these two maps, which highlights the importance of the quantitative method.
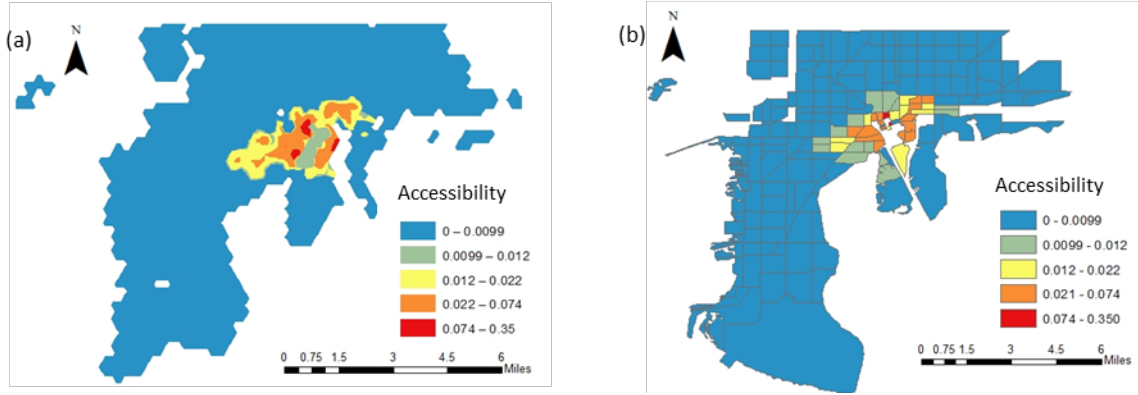

Fig. 5: Geographic mapping analysis with our approach at: (a) parcel level; (b) TAZ level

5.2. Verification of the proposed methodology

In this section, we benchmark the proposed method against three methods (adapted) from the literature to illustrate the necessity and importance of the proposed method.

5.2.1. The necessity to consider "walking-cycling-walking" in a bike-sharing trip

We start with comparing 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 is to illustrate the necessity to consider all three consecutive steps (i.e., walking-cycling-walking) in a bike-sharing trip. 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 usually quantified as the portion of the area or population that is within the service area of the transit system. In bike-sharing systems, nevertheless, we are not just concerned about if the population in an area can be served by a bike-sharing system, but should also consider individuals' willingness to walk to the bike-sharing facilities. Therefore, we define the coverage-based 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 indeed 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). Since the distance decay function is used in the proposed measure, to make the comparison fair, we formulate the reachability-based measure with the distance decay function as $a_{in} = c_{in}, \forall i \in \mathcal{I}, n \in \mathcal{N}_i$. The rest of operations remain the same as those introduced in the methodology section. The geographic mapping analysis results from these two benchmark methods are shown in Fig. 6 and Fig. 7, respectively. The corresponding Gini indexes are summarized in Table 1.
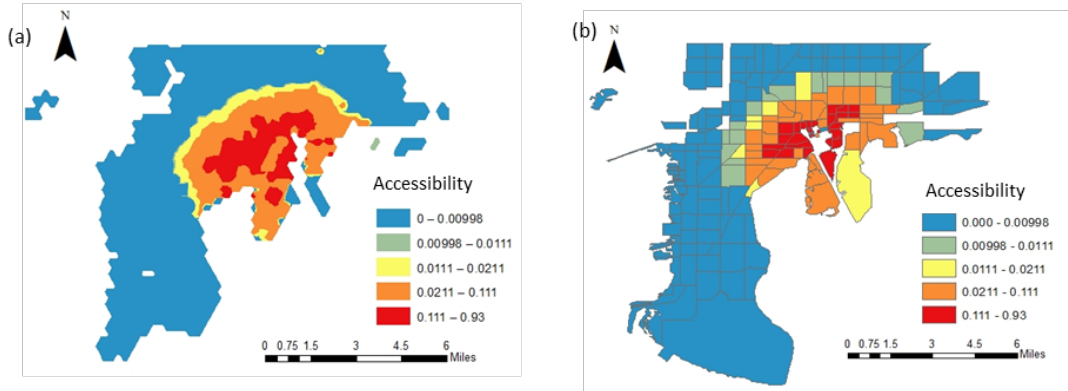

Fig. 6: Geographic mapping analysis with the coverage-based approach at: (a) parcel level; (b) TAZ level
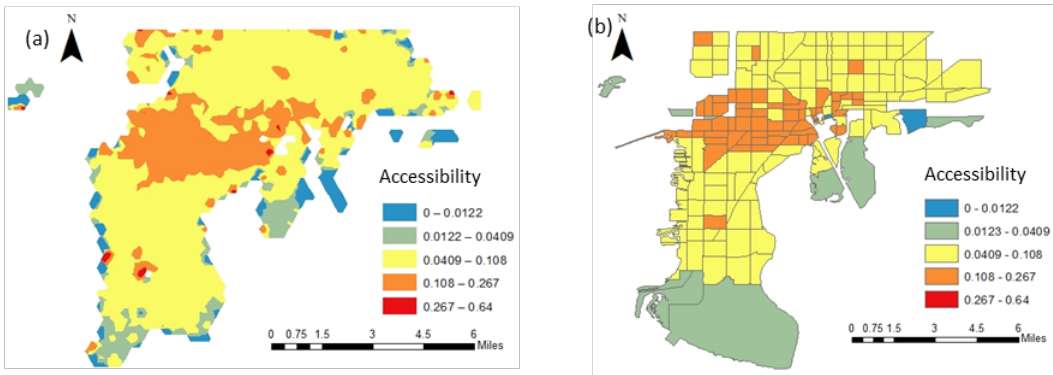

54

Table 1. Gini indexes from our approach and the benchmark approaches

Approach	Population-level	Parcel-level	TAZ-level
Our approach	0.964	0.960	0.854
Coverage-based	0.908	0.888	0.763
Reachability-based	0.620	0.518	0.206
Trip-based	0.883	0.858	0.748

Fig. 6 shows that the coverage-based approach generates a similar distribution of the bike-sharing accessibility across the geographic space as our approach does. However, the resulting bike-sharing accessibility is much higher than that from our approach (especially for areas shaded with red). This is because, the coverage-based approach only considers how an individual can access to bike-sharing facilities in their origin and thus many long-distance trips that cannot be completed with shared bicycles are thought of enjoying high accessibility. Also, trips whose destination without bike-sharing facilities (which means that individuals cannot return the shared bikes and thus would not travel by bikes) are thought of enjoying high accessibility. Therefore, we believe this indicates that ignoring the cycling phase and the walking phase at the destination may contribute to overestimating the bike-sharing accessibility, substantially. A direct consequence of the overestimation is a slightly larger proportion of the population or geographic units that enjoys bike-sharing accessibility in southern Tampa. More spatial units, i.e., either parcels or TAZs, are shaded with non-blue colors (Fig. 6) and the Gini indexes computed for both the parcel and TAZ levels are slightly lower (Table 1). Also, the Gini index from the population perspective is lowered, indicating now only 80% of the residents have no bike-sharing accessibility. These results reveal that only considering the walking phase in the origin might lead to an underestimation to the concentration of the bike-sharing accessibility. Note that this approach also reflects that the coverage of the Coast Bike System is not very high in southern Tampa since the accessibility of most areas is still less than 0.01 even if only the walking phase in the origin is considered. This actually indicates a low density of the bike-sharing station, which demonstrates our argument in the previous section that the low density of bike-sharing facilities is partially responsible for the studied area's overall low bike-sharing accessibility.

From Fig. 7, we can see that the reachability-based approach gives us a completely different distribution of the bike-sharing accessibility in southern Tampa. With this approach, only a small portion of geographic units (less than 20% for parcels and almost 0% for TAZs) are observed to have no bike-sharing accessibility. Further, geographic units with bike-sharing accessibility are not clustered in downtown Tampa any more but scattered across almost the entire studied area, with downtown Tampa and areas to the west of the downtown enjoying relatively high bike-sharing accessibility. As a result, the inequality indexes are all

significantly reduced, implying that the distribution of the bike-sharing accessibility is not as concentrated as that shown by our approach. For example, with a Gini index of 0.206 at the TAZ level, the bike-sharing accessibility can almost be regarded as evenly distributed among all traffic analysis zones in southern Tampa. These results, however, seem to be a bit distant from the reality. In essence, the reachability-based approach measures an individuals' willingness to commence their trips with shared bikes. However, without considering the walking phases in a bike-sharing trips, many short-distance trips that actually cannot be served by the bike-sharing systems due to the unavailability of bike-sharing facilities are still thought to have relatively high bike-sharing accessibility. For instance, the bike-sharing accessibility for parcels near the middle-west edge of the studied area should be much lower than that for parcels in downtown Tampa, because no bike-sharing stations can be found in their communities. However, Fig. 7 show no significant difference between these two areas (both are orange shaded). Thus, only considering individuals' willingness to cycle may result in a misleading conclusion on the distribution of bike-sharing accessibility. However, we note that these results do provide useful information regarding the trip distance in southern Tampa. 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 short-distance trips so their willingness to cycle is relatively low. This observation supports our statement in the previous section that the dominance of long-distance trips is partially responsible for the low bike-sharing accessibility in southern Tampa.

5.2.2. The necessity to carry out tour-based analysis

To illustrate the necessity to carry out the tour-based analysis when modeling the bike-sharing accessibility for an individual, we compare the proposed method with a method using trip-based analysis in this subsection. In the trip-based analysis, the individual bike-sharing accessibility is revised as $a_i = \frac{\sum_{n \in \mathcal{N}_i} a_{in}}{N_i}, \forall i \in \mathcal{I}$ and all other operations remain the same as those introduced in the methodology section. Results from the geographic mapping are shown in Fig. 8 and the corresponding Gini indexes are summarized in Table 1.
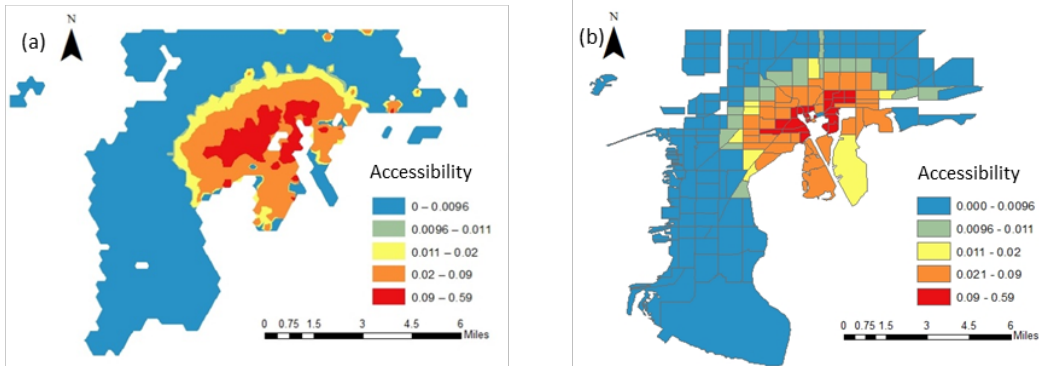
As we can see from Fig. 8, the trip-based analysis also generates a similar distribution of the bike-sharing accessibility as the tour-based analysis does. As expected, this approach overestimates the bike-sharing accessibility since it does not consider the interdependence between the mode choices of several consecutive trips (i.e. a tour). This way, short-distance trips that are not expected (because of the existence of long-distance trips in the trip chain) to be served by the Coast Bike System are thought of with high bike-sharing accessibility. Similar to what we have observed from the coverage-based approach, a natural consequence of such an overestimation is more areas with relatively high bike-sharing accessibility, lower Gini indexes (see Table 1) and thus seemingly less skewed distribution of the bike-sharing accessibility among both the population and geographic units.

5.2.3. The necessity to incorporate disaggregated data in equity analysis

Finally, we want to make a note on the importance of incorporating disaggregated data for equity analysis. As noted above, one of the benefits of using disaggregated data is that they enable us to analyze the horizontal equity from the population perspective with individual-level data. Our results show that such a seemingly simple methodological change is non-trivial. We can see from Fig. 5 to Fig. 8 that as the geographic unit of analysis increases, the bike-sharing accessibility turns out to be lowered. Further, Table 1 presents that the Gini indexes get smaller as the unit of analysis increases from individual to TAZ. These observations indicate that data aggregation tends to absorb the disparities of the bike-sharing accessibility among different individuals. Moreover, the higher the aggregation level (i.e., the larger units of analysis we are using), the more disparities will be absorbed. As a result, the accessibility distribution seems to be less skewed, which can mislead our understanding of horizontal inequality. Therefore, it is better to use disaggregated data, when available, for horizontal equity analysis.

5.3. Vertical equity analysis

This section investigates the distribution of the bike-sharing accessibility based on the individual demographic attributes. Table 2 presents summary statistics of the bike-sharing accessibility by different sociodemographic attributes. Table 3 reports result from the ANOVA test and Fig. 9 plots the subgroup inequality indexes versus the accessibility level.

From Table 2, we find the average population bike-sharing accessibility is extremely low (i.e., 0.0027) in southern Tampa, which is attributable to the small service area of the Coast Bike Share System. This point can also be justified from the observation that the 3rd quartile population bike-sharing accessibility is smaller than the average, indicating that 75% of the entire population is without bike-sharing accessibility.

This result is also consistent with our findings from the horizontal equity analysis. Moving forward to the population subgroups, we find that the bike-sharing accessibility in each sociodemographic group also follows an extremely left-skewed distribution, with the majority of the members having no bike-sharing accessibility (the 3rd quartile for all subgroups are 0). Also, the summary statistics (e.g. mean) of all population subgroups are relatively small values that seem to show almost no difference from each other. Thus, with a quick look at these results, there seems to be no significant difference regarding the bike-sharing accessibility distribution among different sociodemographic groups.

Table 2. Summary statistics of the bike-sharing accessibility for the population and different subgroups

Population groups		minimum	1st quartile	median	mean	3rd quartile	maximum	Std. dev.
the entire population		0	0	0	0.0027	0	0.41	0.0159
race	white	0	0	0	0.0027	0	0.41	0.0162
	black	0	0	0	0.0024	0	0.40	0.0147
	Asian	0	0	0	0.0030	0	0.21	0.0127
	other	0	0	0	0.0024	0	0.41	0.0150
ethnicity	Hispanic	0	0	0	0.0024	0	0.41	0.0148
	Non-Hispanic	0	0	0	0.0028	0	0.41	0.0161
gender	male	0	0	0	0.0028	0	0.41	0.0166
	female	0	0	0	0.0026	0	0.41	0.0152
income	below poverty	0	0	0	0.0034	0	0.41	0.0185
	middle income	0	0	0	0.0020	0	0.38	0.0132
	upper income	0	0	0	0.0031	0	0.41	0.0171
age	0-18	0	0	0	0.0018	0	0.35	0.0117
	18-45	0	0	0	0.0035	0	0.41	0.0184
	45-65	0	0	0	0.0022	0	0.41	0.0147
	> 65	0	0	0	0.0029	0	0.38	0.0167

However, the ANOVA tests for differences in means tell us a different story. As can be seen from Table 3, the P-values from all tests are less than 0.002, indicating that we have 99.8% confidence to reject the null hypothesis that the mean bike-sharing accessibility is the same across different population subgroups. Thus, for all five sociodemographic attributes considered in this study, there are statistically significant differences among the population subgroups in terms of the average bike-sharing accessibility. To be more specific, among different racial the ethnic categories, Asian (by 11.1%) and non-Hispanic (by 3.7%) receive higher bike-sharing accessibility than the population mean while all of the black, other races, Hispanic subgroups receive 11.1% less bike-sharing accessibility than the population mean. In terms of gender, the average bike-sharing accessibility of male is better than the population mean by 3.7%. Among income
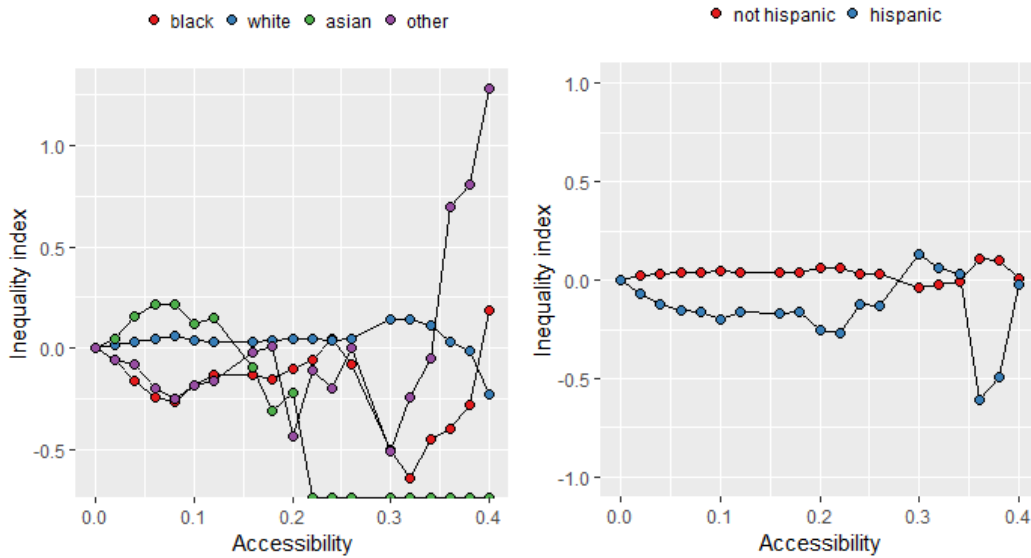
categories, the below poverty and upper income are better off in the average bike-sharing accessibility than the population mean by 25.9% and 14.8%, respectively. Finally, regarding different age groups, both adults aged between 18 and 45 and senior citizens over 65 enjoy higher bike-sharing accessibility than the population on average by 29.6% and 7.4%, respectively. Yet, the average bike-sharing accessibility are lower than the population mean by 33.3% and 18.5%, respectively, for people aged under 18 and between 45 and 65. In overall, these results offer us a coarse understanding on the distribution of bike-sharing accessibility among different population subgroups.

Table 3. Result from the ANOVA tests

Sociodemographic attributes	ANOVA test					
	Source	SS	Df	MS	F	P-value
Race	Intergroup	0.0038	3	0.0013	5.51	0.0017
	Intragroup	39.56	156424	0.0003		
Ethnicity	Intergroup	0.0031	1	0.0031	12.21	0.0005
	Intragroup	39.56	156426	0.0003		
Gender	Intergroup	0.0028	1	0.0028	11.17	0.0008
	Intragroup	39.56	156126	0.0003		
Income	Intergroup	0.0535	2	0.0267	105.7	< 2E-16
	Intragroup	39.51	156125	0.0003		
Age	Intergroup	0.0729	3	0.0243	96.04	< 2E-16
	Intragroup	39.49	156124	0.0003		

The subgroup inequality index can unveil more detailed information on how different levels of the bike-sharing accessibility are distributed among different population groups, which cannot be obtained from the ANOVA test. For instance, from Fig. 9 (a), we can observe slightly but disproportionally high bike-sharing accessibility (with subgroup inequality index values greater than 0) for the white group across almost all accessibility levels, while the means (Table 2) reveals no difference between this group and the population. For Asian people that are found to be disproportionally highly-represented based on comparing means, the index values show that they are actually disproportionally highly-represented when the accessibility is lower than 0.15, but the situation is different when the accessibility is over 0.15. Indeed, the inequality index value for Asian people reaches minus infinity when the accessibility is greater than 0.2, indicating that they do not have bike-sharing accessibility higher than 0.2 at all. Also not captured by the comparing means, the black subgroup receives disproportionally lower bike-sharing accessibility at most accessibility levels but they are extremely highly represented when the accessibility level is greater than 0.35 (with the inequality index greater than 0.75). For different ethnic and gender categories, the distributions change less substantially with the accessibility level.

From the above analysis, we can see that traditional aggregate methods can mask disparities of the bike-sharing accessibility distribution at different accessibility levels, which might cause misleading understandings of the vertical equity issue. Similar findings can also be found in other sociodemographic attributes we consider. We can see from Fig. 9 (b) and (c) that the bike-sharing accessibility is skewed towards the non-Hispanic and male groups at most levels, which is consistent with findings from comparing means. However, some minor deviations still can be observed in both figures. In this situation, though not precise, the disparities predicted by comparing means largely hold. Regarding the income categories (see Fig. 9 (d)), the subgroup inequality index indicates that the below-poverty group consistently receives above average bike-sharing accessibility as accessibility level increases while the story of the middle class goes oppositely. This finding is consistent with comparing the means. What we cannot learn from comparing means is that the disparity among each group increases with the accessibility level. Finally, Fig. 9 (e) also gives us the same result as comparing means; i.e. subgroups "0 ~18 yrs" and "45 ~65 yrs" are lowly-represented while the other two are highly-represented. However, this result does not consistently hold at different accessibility levels since the values of the index of group inequality indexes fluctuate dramatically with the accessibility level.
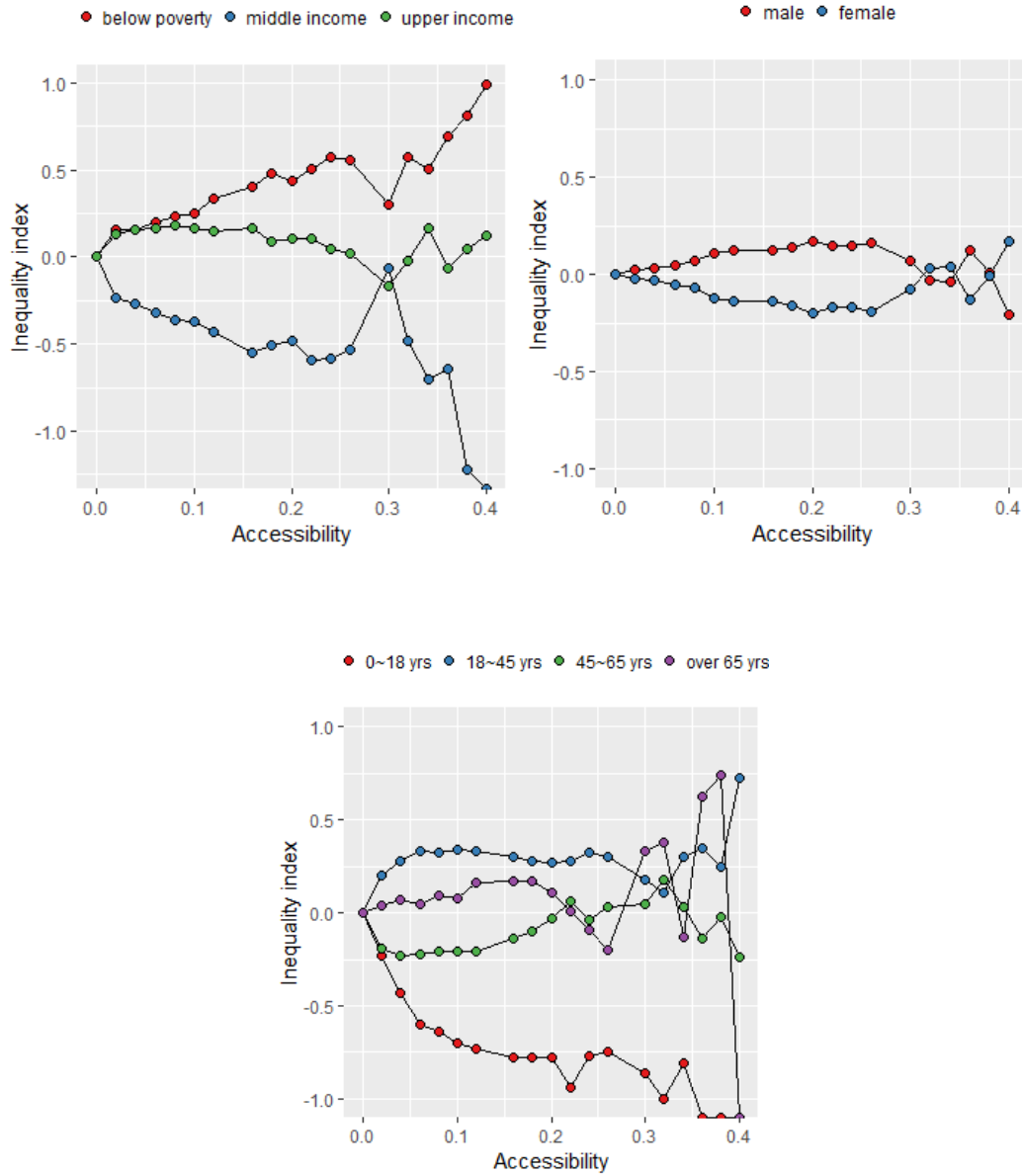
Fig. 9: Subgroup inequality index versus cumulative accessibility level (Dots at the bottom of the figure represents minus infinity)

6. Conclusions

This paper closes the research gap in the literature by developing a comprehensive equity assessment framework on analyzing how the accessibility from a bike-sharing system is distributed in the society with disaggregated data. With the individual travel demand dataset and bike-sharing provision dataset, the framework first models the individual bike-sharing accessibility by taking into account 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 obtained individual accessibility indicators and sociodemographic dataset to carry out a series of equity analyses from both the horizontal and vertical perspective. Apart from traditional analysis methods such as geographic mapping analysis, Gini index, distributional comparison and ANOVA test, a subgroup inequality index is applied to measure the vertical equity quantitatively. The proposed methodological framework is applied to the Coast Bike Share System in southern Tampa. The main findings are summarized as follows.

1. From the horizontal perspective, the distribution of bike-sharing accessibility is highly skewed among both the population and the geographic space in southern Tampa, with both Gini indexes higher than 0.95. Geographic mapping analysis reveals that the accessibility is concentrated in areas within and around downtown Tampa.

2. From the vertical perspective, the bike-sharing accessibility is not evenly distributed among different sociodemographic groups. Overall, the bike-sharing accessibility is higher for whites, Asians, non-Hispanic, male, middle and upper income classes, and people aged between 18 and 45 and over 65. However, the distributions change substantially with the accessibility level for some individual attributes, such as race, income level and age.

3. The bike-sharing accessibility in southern Tampa is relatively low due to its low density and the large portion of long-distance travel. By considering the "walking-cycling-walking" process in a bike-sharing trip and the trip chaining in individuals' travel itinerary, the proposed method avoids overestimating the bike-sharing accessibility. This finding demonstrates the necessity and importance of the proposed tour-based modeling approach.

4. The disaggregated data enable us to analyze the horizontal and vertical equity at the individual level, which unveils many important messages that might be absorbed with existing methods using aggregated data. Indeed, aggregated data (e.g., mean) may dilute the disparities among individuals, which might mislead our understanding of the equity issue 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 in which this work can be extended. Bettering understanding of how bike-sharing systems interact with other transportation modes in a multimodal transportation system can paint a more complete and realistic picture of the transportation equity in a city. As transportation is increasingly regarded as a service in modern society, taking into account factors that might affect the service quality of the bike-sharing systems such as the number of bikes at stations, repositioning activities is necessary. Another interesting avenue is to offer an assessment on the accessibility, environmental and public health benefits as a whole.

Acknowledgements

References

Beckman, R.J., Baggerly, K.A., Mckay, M.D., 1996. CREATING SYNTHETIC BASELINE POPULATIONS. Transpn Res.-A 30.

Benjamin, S., 2017. What keeps bike share white [WWW Document]. URL https://www.citylab.com/equity/2017/07/what-keeps-bike-share-white/533412/

Bills, T.S., Walker, J.L., 2017. Looking beyond the mean for equity analysis: Examining distributional impacts of transportation improvements. Transp. Policy 54, 61–69. https://doi.org/10.1016/j.tranpol.2016.08.003

Boarnet, M.G., Giuliano, G., Hou, Y., Shin, E.J., 2017. First/last mile transit access as an equity planning issue. Transp. Res. Part A Policy Pract. 103, 296–310. https://doi.org/10.1016/j.tra.2017.06.011

Bradley, M., Bowman, J.L., Griesenbeck, B., 2010. SACSIM: An applied activity-based model system with fine-level spatial and temporal resolution. J. Choice Model. 3, 5–31. https://doi.org/10.1016/S1755-5345(13)70027-7

Buck, D., 2013. Encouraging equitable access to public bikesharing systems. ITE J. (Institute Transp. Eng. 83, 24–27.

Castro, A., 2011. The contribution of bike-sharing to sustainable mobility in Europe 1–236.

Church, R.L., Noronha, V., Lei, T., Corrigan, W., Burbidge, S., Marston, J., 2005. Spatial and Temporal Utility Modeling to Increase Transit Ridership. Transportation (Amst).

Currie, G., 2010. Quantifying spatial gaps in public transport supply based on social needs. J. Transp. Geogr. 18, 31–41. https://doi.org/10.1016/j.jtrangeo.2008.12.002

Delbosc, A., Currie, G., 2011. Using Lorenz curves to assess public transport equity. J. Transp. Geogr. 19, 1252–1259. https://doi.org/10.1016/j.jtrangeo.2011.02.008

Du, M., Cheng, L., 2018. Better understanding the characteristics and influential factors of different travel patterns in free-floating bike sharing: Evidence from Nanjing, China. Sustain. 10. https://doi.org/10.3390/su10041244

El-Geneidy, A., Levinson, D., Diab, E., Boisjoly, G., Verbich, D., Loong, C., 2016. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. Transp. Res. Part A Policy Pract. 91, 302–316. https://doi.org/10.1016/j.tra.2016.07.003

El-Geneidy, A.M., Tétreault, P.R., Surprenant-Legault, J., 2010. Pedestrian access to transit: Identifying redundancies and gaps using a variable service area analysis. Proc. 89th Annu. Meet. Transp. Res. Board 1–18.

Gavin, K., Bennett, A., Auchincloss, A.H., Katenta, A., 2016. A brief study exploring social equity within bicycle share programs. Transp. Lett. 8, 177–180. https://doi.org/10.1080/19427867.2015.1126065

Goodman, A., Wilkinson, P., Stafford, M., Tonne, C., 2011. Characterising socio-economic inequalities in exposure to air pollution: A comparison of socio-economic markers and scales of measurement. Heal. Place 17, 767–774. https://doi.org/10.1016/j.healthplace.2011.02.002

Gurram, S., Stuart, A. L., & Pinjari, A. R. (2015). Impacts of travel activity and urbanicity on exposures to ambient oxides of nitrogen and on exposure disparities. Air Quality, Atmosphere & Health, 8(1), 97-114.

Gurram, S., 2017. Understanding the Linkages between Urban Transportation Design and Population Exposure to Traffic-Related Air Pollution : Application of an Integrated Transportation and Air Pollution Modeling Framework to Tampa , FL.

Guzman, L.A., Oviedo, D., Rivera, C., 2017. Assessing equity in transport accessibility to work and study: The Bogotá region. J. Transp. Geogr. 58, 236–246. https://doi.org/10.1016/j.jtrangeo.2016.12.016

Harner, J., Warner, K., Pierce, J., Huber, T., 2002. Urban environmental justice indices. Prof. Geogr. 54, 318–331. https://doi.org/10.1111/0033-0124.00333

Harper, S., Charters, T.J., Strumpf, E.C., 2015. Trends in socioeconomic inequalities in motor vehicle accident deaths in the United States, 1995-2010. Am. J. Epidemiol. 182, 606–614. https://doi.org/10.1093/aje/kwv099

Hasan, S., Ukkusuri, S. V., 2014. Urban activity pattern classification using topic models from online geo-location data. Transp. Res. Part C Emerg. Technol. 44, 363–381. https://doi.org/10.1016/j.trc.2014.04.003

Hochmair, H.H., 2015. Assessment of Bicycle Service Areas around Transit Stations. Int. J. Sustain. Transp. 9, 15–29. https://doi.org/10.1080/15568318.2012.719998

Jiang, S., Ferreira, J., Gonzalez, M.C., 2017. Activity-Based Human Mobility Patterns Inferred from Mobile Phone Data: A Case Study of Singapore. IEEE Trans. Big Data 3, 208–219. https://doi.org/10.1109/TBDATA.2016.2631141

Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., 2016. Correction for Jiang et al., The TimeGeo modeling framework for urban motility without travel surveys. Proc. Natl. Acad. Sci. 113, E7137–E7137. https://doi.org/10.1073/pnas.1616240113

Kaplan, S., Popoks, D., Prato, C.G., Ceder, A., 2014. Using connectivity for measuring equity in transit provision. J. Transp. Geogr. 37, 82–92. https://doi.org/10.1016/j.jtrangeo.2014.04.016

Kawabata, M., Shen, Q., 2006. Job accessibility as an indicator of auto-oriented urban structure: A comparison of Boston and Los Angeles with Tokyo. Environ. Plan. B Plan. Des. 33, 115–130. https://doi.org/10.1068/b31144

Levy, J.I., Greco, S.L., Melly, S.J., Mukhi, N., 2009. Evaluating efficiency-equality tradeoffs for mobile source control strategies in an urban area. Risk Anal. 29, 34–47. https://doi.org/10.1111/j.1539-6924.2008.01119.x

Litman, T., 2002. Evaluating Transportation Equity Guidance For Incorporating Distributional Impacts in Transportation Planning Evaluating Transportation Equity. World Transp. Policy Pract. 8, 50–65. https://doi.org/www.vtpi.org/equity.pdf

Liu, S., Zhu, X., 2004. An integrated GIS approach to accessibility analysis. Trans. GIS 8, 45–62. https://doi.org/10.1111/j.1467-9671.2004.00167.x

Lorenz, M.O., 1905. Methods of measuring the concentration of wealth. Taylor Fr. 9, 209–219.

Lucas, K., van Wee, B., Maat, K., 2016. A method to evaluate equitable accessibility: combining ethical theories and accessibility-based approaches. Transportation (Amst). 43, 473–490. https://doi.org/10.1007/s11116-015-9585-2

Mcneil, N., 2015. Overcoming Equity Barriers in Bike Share.

Mishra, S., Welch, T.F., Jha, M.K., 2012. Performance indicators for public transit connectivity in multi-modal transportation networks. Transp. Res. Part A Policy Pract. 46, 1066–1085. https://doi.org/10.1016/j.tra.2012.04.006

Moniruzzaman, M., Páez, A., 2012. Accessibility to transit, by transit, and mode share: application of a logistic model with spatial filters. J. Transp. Geogr. 24, 198–205. https://doi.org/10.1016/j.jtrangeo.2012.02.006

Murray, A.T., 2001. Strategic analysis of public transport coverage. Socioecon. Plann. Sci. 35, 175–188. https://doi.org/10.1016/S0038-0121(01)00004-0

NACTO, 2018. Bike Share in the U.S.:2017 [WWW Document]. URL https://nacto.org/bike-share-statistics-2017/

Nassir, N., Hickman, M., Malekzadeh, A., Irannezhad, E., 2016. A utility-based travel impedance measure for public transit network accessibility. Transp. Res. Part A Policy Pract. 88, 26–39. https://doi.org/10.1016/j.tra.2016.03.007

O'Sullivan, D., Morrison, A., Shearer, J., 2000. Using desktop GIS for the investigation of accessibility by public transport: An isochrone approach. Int. J. Geogr. Inf. Sci. 14, 85–104. https://doi.org/10.1080/136588100240976

Oswald Beiler, M., Mohammed, M., 2016. Exploring transportation equity: Development and application of a transportation justice framework. Transp. Res. Part D Transp. Environ. 47, 285–298. https://doi.org/10.1016/j.trd.2016.06.007

Pettit, P., 1974. A theory of justice? Theory Decis. 4, 311–324. https://doi.org/10.1007/BF00136652

Ricciardi, A.M., Xia, J.C., Currie, G., 2015. Exploring public transport equity between separate disadvantaged cohorts: A case study in Perth, Australia. J. Transp. Geogr. 43, 111–122. https://doi.org/10.1016/j.jtrangeo.2015.01.011

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

Schneider, C.M., Belik, V., Couronné, T., Smoreda, Z., González, M.C., 2013. Unravelling daily human mobility motifs. J. R. Soc. Interface 10. https://doi.org/10.1098/rsif.2013.0246

Shaheen, S., Guzman, S., Zhang, H., 2010. Bikesharing in Europe, the Americas, and Asia. Transp. Res. Rec. J. Transp. Res. Board 2143, 159–167. https://doi.org/10.3141/2143-20

Sider, T., Hatzopoulou, M., Eluru, N., Goulet-Langlois, G., Manaugh, K., 2015. Smog and socioeconomics: an evaluation of equity in traffic-related air pollution generation and exposure. Environ. Plan. B Plan. Des. 42, 870–887. https://doi.org/10.1068/b130140p

Stuart, A.L., Mudhasakul, S., Sriwatanapongse, W., 2009. The social distribution of neighborhood-scale air pollution and monitoring protection. J. Air Waste Manag. Assoc. 59, 591–602. https://doi.org/10.3155/1047-3289.59.5.591

Thakali, L., Kwon, T.J., Fu, L., 2015. Identification of crash hotspots using kernel density estimation and kriging methods: a comparison. J. Mod. Transp. 23, 93–106. https://doi.org/10.1007/s40534-015-0068-0

Welch, T.F., Mishra, S., 2013. A measure of equity for public transit connectivity. J. Transp. Geogr. 33, 29–41. https://doi.org/10.1016/j.jtrangeo.2013.09.007

Wikipedia, 2018. List of bike sharing systems [WWW Document]. URL https://en.wikipedia.org/wiki/List_of_bicycle-sharing_systems#cite_note-The_Seattle_Times-1

Woodcock, J., Tainio, M., Cheshire, J., O'Brien, O., Goodman, A., 2014. Health effects of the London bicycle sharing system: Health impact modelling study. BMJ 348, 1–14. https://doi.org/10.1136/bmj.g425

Yu, H., Stuart, A.L., 2016. Exposure and inequality for select urban air pollutants in the Tampa Bay area. Sci. Total Environ. 551–552, 474–483. https://doi.org/10.1016/j.scitotenv.2016.01.157

Yu, H., Stuart, A.L., 2013. Spatiotemporal distributions of ambient oxides of nitrogen, with implications for exposure inequality and urban design. J. Air Waste Manage. Assoc. 63, 943–955. https://doi.org/10.1080/10962247.2013.800168

Zhang, Y., Mi, Z., 2018. Environmental benefits of bike sharing: A big data-based analysis. Appl. Energy 220, 296–301. https://doi.org/10.1016/j.apenergy.2018.03.101