

# Fairness in Practice: A Survey on Equity in Urban Mobility

An Yan, Bill Howe  
University of Washington  
{yanan15,billhowe}@uw.edu

## 1 Introduction

More than 54 percent of the world's population lives in urban areas [1]. Predicting dynamic urban activities such as energy consumption, air pollution, public safety, and traffic flows has become a fundamental task for improving the quality of human life. Urban mobility is closely intertwined with these problems, and is therefore a major determinant of quality of life [2], crucial to employment opportunities and access to resources such as education and health care [3].

Evidence suggests that residents of low-income and minority neighborhoods are concentrated away from economic opportunity and public resources [4]. Injustice of transportation services experienced by these residents further reinforces social exclusion as the availability and quality of transportation impact a person's access to opportunities [5, 6, 7, 8]. For example, one study revealed that living in neighborhoods with longer commute times is associated with lower employment rates of younger generations [9]. As a result, transportation equity issues have motivated government agencies to develop extensive multimodal transportation networks[6].

*New mobility* is about emerging transportation modes, including but not limited to car-sharing, bike-sharing, and ride-hailing or Transportation Network Companies (TNCs) [10]. New mobility services provide technology-based, on-demand, and affordable alternatives to traditional means. These services offer a chance to address persistent equity issues in transportation. However, new mobility services also bring new equity concerns. For example, people without internet service, smart phones, or credit cards are not able to get access to the services. Moreover, studies show that algorithms or human beings that distribute app-based mobility services may discriminate against people of color [11].

This paper reviews the methods and findings of mobility equity studies, with a focus on new mobility. The paper is structured as follows: Section 2 presents the background of transportation equity. Section 3 summarizes the main findings from current equity studies for mobility systems, with a brief discussion on future research. Section 4 reviews the commonly used methods for evaluating the equity of mobility service provision and usage and considers strengths and weaknesses. Section 5 discusses the relationship between the transportation equity community and the fairness in machine learning community. Section 6 concludes the paper.

---

*Copyright 2019 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.*

**Bulletin of the IEEE Computer Society Technical Committee on Data Engineering**

---

This work is supported by the National Science Foundation under grant NSF BIGDATA-1740996.

## 2 Equity in Mobility Systems: Background

Automated decision systems powered by machine learning and big data have been widely employed in many applications including credit scoring, criminal justice, online advertising, employment, etc [12, 13, 14, 15]. These systems have been hailed as efficient, objective, and accurate alternatives to human decision makers [16]. However, increasing evidence has shown that data-driven systems contain biases. For example, Google’s image recognition system wrongly identified black users as gorillas [17]. Amazon’s same-day delivery services excluded predominantly black neighborhoods in many cities to a varying degree [18].

Even if the algorithms themselves are well-intentioned, they can replicate and amplify human biases encoded in the data, thus resulting in unequal distribution of impacts across different demographic groups [12, 19, 20]. This effect is due to machine learning algorithms seeking to fit the training data as closely as possible to make accurate predictions. The process of learning also “accurately” captures historical signals of discrimination. In 2017, Caliskan et al. [21] found that an influential language corpus [22] generated by machine learning algorithms accurately reproduced historic biases. The corpus reflects societal stereotypes such as female names are more associated with family while male names are more associated with career. Not only do algorithms pick up discrimination in data, they also magnify them [23]. This effect is often due to the underrepresentation of minority groups in training data, leading to higher error rates for the minorities. One study [24] revealed that a widely-used predictive policing tool, PredPol, would reinforce the bias in the police-recorded, resulting in disproportionate policing of minority communities.

The heightened concerns about automated decision systems concentrate not only on discrimination, but also on a range of related issues, including transparency, privacy, and accountability [25]. These issues often intertwine and conflict with one another in practice. In the context of automatic decision systems, transparency is about the openness and understandability of data and models and accountability is about being responsible for the decisions [26]. Transparency is a critical prerequisite for accountability. In the absence of concrete evidence of intentional discrimination, it is difficult to hold an individual or organization accountable for biased decisions.

In practice, transparency for automatic decision systems is not easily achievable. Burrell [27] summarized three types of barriers to transparency: 1) intrinsic opacity, where some algorithms such as deep learning models are difficult to understand and interpret; 2) illiterate opacity, which says the general public may lack the expertise to understand the algorithms; and 3) intentional opacity, which is often resulted from intellectual property protection of the algorithm developers.

### 2.1 Definitions of transportation equity

Equity in the context of mobility has been studied independently since well before the recent interest in generalized fairness methods for machine learning. These efforts suggest that domain-specific and context-sensitive approaches should be incorporated into any fairness-aware ML system. Equity for mobility is the fair distribution of transportation costs and benefits, among current (and future) members of society [5].

There are mainly two perspectives from which to examine equity: horizontal equity and vertical equity. *Horizontal equity* (also called fairness and egalitarianism) is concerned with providing equal resources to individual or groups considered equal in ability and need, which means the public policies avoid favoring one individual or group over another. Horizontal equity suggests that those who pay more should receive superior services.

*Vertical equity* (also referred to as social justice, environmental justice, and social inclusion) is concerned with allocating resources to individuals or groups that differ in income, social class, mobility need, or ability. It advocates that public policies favor disadvantaged groups by providing discounts or special services, therefore compensating for overall inequities. One way to evaluate vertical equity is *equity of opportunity*, meaning that disadvantaged groups should have adequate access to transportation resources. Equity of opportunity is usually measured by access to services. In contrast, “*equity of outcome*” is usually measured by the actual usage of the

systems across groups. There is a general agreement about the goal of equity of opportunity, but less agreement about equity of outcome [5, 28, 29].

There are other ways to define transportation equity. Social equity indicates the differences between socio-economic groups. Spatial equity refers to the differences in transport services among geographic regions [30]. These different definitions often overlap or conflict with each other. For example, horizontal equity requires the users to pay for what they get, whereas vertical equity prioritizes the needs of disadvantaged groups such as the low-income or ethnic minorities in the form of discounts [31].

## 2.2 Evaluation of mobility equity

Mobility equity research addresses a wide range of issues, including, for example, economic studies on how transportation is subsidized and taxed, and operational studies on how negative impacts of transportation systems are distributed among different groups [6]. Litman [5] proposed four variables to consider when performing any equity evaluation.

- Type of equity: horizontal equity or vertical equity
- Impact (costs and benefits) categories: public facilities and services, user costs and benefits (e.g., taxes and fares), service quality (e.g., public transportation service quality including frequency, speed, safety, reliability, comfort, etc.), external impacts (e.g., air pollution), economic impacts (e.g., access to employment), and regulation and enforcement (e.g., parking regulations)
- Measurement unit: per capita, per unit of travel (e.g., per trip), or per dollar.
- Categorization of people: demographics (e.g., age, household type, race), income class, ability, location, mode (e.g., pedestrians, public transit), industry (e.g., freight, public transit), and trip type (e.g., commutes)

This paper focuses on the equity of new mobility systems service provision and usage across different social-economic, demographic, or geographic groups.

## 3 Findings from Equity Research in Mobility Systems

We describe findings in the literature across 1) public transportation, and 2) new mobility services.

**Public transportation** Transportation equity has long been a major concern of governmental agencies, researchers, and the general public [5, 6, 7]. Despite the tremendous investment in transportation system development and progress in transportation equity research, there are still many long-standing equity issues resulted from unequal distributions of transport resources across different socioeconomic groups and spatial regions [32, 33, 8, 34]. A number of studies have found out that an uneven urban development has resulted in a lack of public transport supply for disadvantaged groups. For example, Vasconcellos [35] pointed out in Brazil, road systems are developed in a radial pattern. Low-income residents usually settled in fringes of the city with irregular pavements or hilly areas that are subject to landslides. The urban centers with good public services are mostly occupied by the high-income people. Similarly, Ricciardi [8] found that there is an unequal spatial distribution of public transport resources in two Australian cities. Their analysis showed that 70% of Perth's population shares one third of the public transit supply. Moreover, three socially disadvantaged groups — the elderly, low-income, and no-car households have less access to public transport services compared to the overall population. Some studies also showed that the economic burden and negative climate impacts of transportation systems is disproportionately imposed on disadvantaged people [33, 30, 36]. In recognizing these issues, many cities now have incorporated social equity into urban transportation planning. However, one study found that

social equity goals are often not translated into clear and actionable items and there is a lack of appropriate methods for assessing their achievements [37, 5]. Current literature on equity in public transport suggests that disadvantaged groups as a whole experience inequitable access to public transport services but suffer from significant negative impacts from the transportation systems.

### 3.1 New mobility

**Bikeshare** A number of researchers have studied equity in bikeshare systems. Several studies found that bikeshare stations were typically located in urban centers with high population density [38, 39], and there was a lack of stations in low-income areas. In an assessment of bikeshare systems in seven US cities, Ursaki et al. found significant differences in the race, education level, and income of population inside and outside bike share service areas in four cities [40]. Other studies also indicated that in North America, advantaged groups tend to have more access to docked bikeshare than disadvantaged groups [41]. Recently, free-floating (dockless) bike share systems have been introduced in several major cities in China and the United States [42, 43, 44]. Free-floating bikeshare systems may have different equity landscape from docked systems. There are no stations in the city, therefore there are no fixed service areas. In this way, access to bikes are transient and largely dependent on the placement of individual bikes, which is driven by user demand and companies' bike rebalancing strategies. As free-floating bikeshare systems are fairly new, the impact on equity are unclear. In examining access equity of dockless bikes in Seattle, Mooney et al. found out that more college-educated and higher-income residents have access to more bikes. They also found out that bike demand is highly correlated with rebalancing destinations [43], suggesting that the companies themselves are accountable for equity issues that arise.

Equal access to bikeshare does not imply equity of actual usage. Several studies found inequalities in the usage of bikeshare systems [45, 46]. For example, Daddio et al. [45] found a negative association between station-level usage with non-white population in Washington, D.C. The disparities in use partially stem from the inequalities of access, but there are many other factors that inhibit bikeshare use among disadvantaged groups. McNeil et al. found out that the biggest barrier to bikeshare is traffic safety, regardless of race or income [47]. Lower-income people of color have more concerns about costs of membership and more misconceptions about bikeshare than higher-income white people. Another study [48] found that credit card requirement, lack of computer access, annual subscription fee, and lack of bike lanes etc. are reported by low-income residents as barriers to bikeshare. Shaheen et al. [6] identified five types of barriers to use bikeshare including spatial, temporal, economic, physiological, and social barriers, and provided policy recommendations. Overall, current literature suggests that disparities exist in the access and use of bikeshare systems.

**Ride-hailing** Ride-hailing can potentially redefine car access, mitigating the mobility divide resulted from car ownership [49]. But the equity of ride-hailing services remains unclear. Several studies found that the service quality in terms of waiting times is not necessarily associated with the average income or minority fraction of pickup locations [50, 51]. A recent study [49] found that users in low-income neighborhoods actually use Lyft more frequently than users in high-income neighborhoods in Los Angeles. The findings of this study suggest that Lyft may provide automobile alternatives to neighborhoods with less access to cars. These findings contradict the conclusions from other studies, which suggest that TNCs provide poor services to low-income neighborhoods [52].

Another thread of research examined the discrimination in TNCs. Ge et al. [11] found out that TNC drivers discriminate against African American riders, resulting in longer waiting times and higher trip cancellation rate in Boston and Seattle. Similarly, Brown [49] found that black riders experienced four percent higher trip cancellation rates and longer waiting times than white riders in Los Angeles. Middleton [53] examined rider-to-ride discrimination in ridesharing. Results showed that white respondents in majority white counties are more likely to hold discriminatory attitudes towards riders of other races or class. A few studies investigated the relationship between TNCs and public transit. For example, Jin et al. [54] studied whether Uber contributes to the horizontal equity of transportation system. Their results implied that Uber has insignificant improvement

over transportation equity in New York City. In short, the extent to which ride-hailing forestall or exacerbate inequalities in transportation is not well understood.

## 4 Methods of Evaluating Transportation Equity

A variety of research methods including survey research [47], interviews and focus group [48], content analysis [37], correlational research [50, 51, 45, 46, 38], experimental research [11, 49], and equity metrics [28, 54, 39], have been employed to evaluate transportation equity. These methods differ in their focuses and features, but can be used together to complement each other. Statistical tests (that routinely used in correlational research) and equity metrics are two key techniques for discrimination discovery in both transportation and machine learning research. Experimental research allows the identification of causal relationships between variables. This section focuses on correlational research, experimental research, and equity metrics.

### 4.1 Correlational research

Correlational research aims to explore the relationship between two or more natural occurring variables. It determines which variables are related and how they are related (e.g., positive or negative) [55]. The two main steps involved in correlational research are measurement and data analysis. Researchers collect and measure variables from a variety of settings, but do not control over or manipulate them. Data analysis (e.g., statistical analyses, GIS methods, visualization) is applied to describe the relationships between variables. Correlational research does not establish a causal relationship between variables, but allows researchers to examine the associations among many variables at the same time.

Many equity studies employ correlational research to discover associations between transportation services provision (or usage) and sensitive attributes (i.e., percentage of minority in a neighborhood). Statistical methods (e.g., regression, t-test) are often used to discover statistical relationship. GIS methods (e.g., buffer analysis) are usually employed at the same time for generating variables for statistical tests, analyzing spatial distribution, and visualizing results. Three examples of correlational research are presented below.

**Example 1: Quantifying the equity of bikeshare access in US cities** Ursaki and Aultman-Hall [40] examined the access equity of docked bikeshare systems in seven US cities by comparing the socioeconomic characters of areas within and outside bikeshare service areas. A service area is defined as a 500m buffer around a bike station. The equitable situation for a city is that the characteristics (e.g., percent white) of population inside the service areas are not different from the population outside service areas.

The authors obtained docking station locations data from both the open data portals and the operators directly. Socioeconomic data including population density, race, education level, income, and age was obtained from ACS at census block group (CBG) level. Then the socioeconomic variables inside and outside service areas per CBG was calculated for each city. Student's t-tests were performed to assess statistical significance. Their results showed that the low-income, the elderly, and the minority have less access to bikeshare. For example, in Chicago, the percentage of African Americans inside and outside service areas is 18.7% and 41.9%, respectively.

This example examines seven cities in one study, providing a more holistic view of the equity of bikeshare access compared to studies that focus on only one city. Nevertheless, this study has several limitations. First, equity analysis is only conducted at city level. Although the socioeconomic variables inside and outside service areas were calculated at CBG level, the authors did not discuss the spatial heterogeneity within each city. Second, docking station placement is only one of the factors that influence access equity. This study did not consider other important factors such as the supply of bikes at each station over time. Lastly, the Student's t-tests may give misleading results in this study, because the spatial dependencies among neighboring CBG violate the independence assumption required by the test.

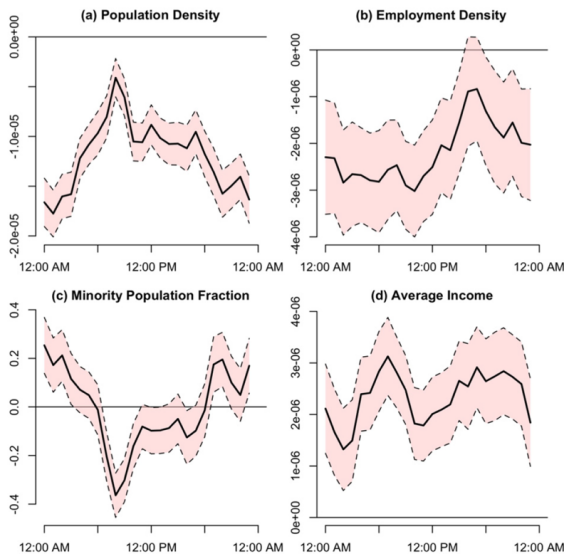


Figure 1: Coefficient estimates and 95% confidence interval of spatial error model for (a) population density, (b) employment density, (c) minority population fraction, and (d) average income [50].

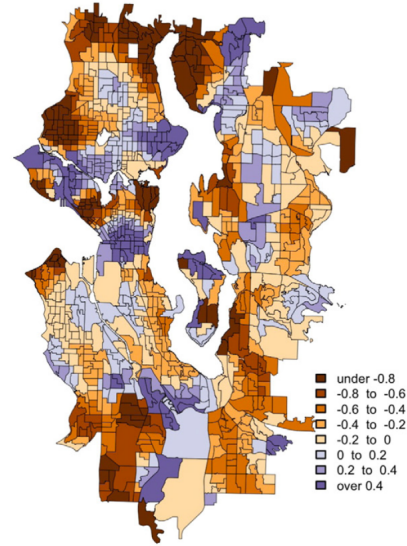


Figure 2: Coefficients for minority fraction from geographically weighted regression. Purple indicates a positive association between expected waiting time and minority fraction; gold indicates a negative association [50].

**Example 2: Transportation network company wait times in Greater Seattle, and relationship to socioeconomic indicators** Hughes and MacKenzie [50] investigated the relationships between wait times for UberX and socioeconomic indicators at census block group (CBG) level in Greater Seattle area. They obtained wait times by making UberX requests through Uber API using quasi-randomly selected locations throughout Greater Seattle. They collected about 1 million data points over a two-month period in 2015. Socioeconomic data including population density, employment density, average income, and minority population fraction was collected from the American Community Survey 5-year estimates (ACS).

They first fitted a regression model with mean waiting times in a CBG as dependent variable and socioeconomic attributes as independent variables. Using a Moran index test, they found significant spatial dependencies among waiting times in different CBG. Subsequently, they developed a spatial error model for each hour of the day to incorporate spatial effect into regression. Results showed that after adjusting the other covariates, higher population density and employment density were associated with shorter waiting time, but that the fraction of minorities in a block group did not significantly associated with waiting times, and that the relationship between these two variables varied between positive and negative throughout the day (Figure 1). In addition, higher average income is associated with longer wait times, suggesting that low-income areas enjoy better services. Geographically weighted regression (GWR) [56] was used to inform different impacts of each socioeconomic variable on different regions. GWR results showed that the relationship between the fraction of minority and wait times is mostly negative. They concluded that “white and wealthy” areas do not necessarily enjoy a better TNC service in terms of wait times.

The strength of this study is that it examined both spatial and temporal variations of the effects of different variables (e.g., minority fraction) on TNC waiting times. An interesting addition to this study is to include factors describing the urban form, such as road network into analysis into analysis. For example, the authors found out that higher income is associated with longer waiting time. It is possible that areas with dense road networks tend to experience shorter waiting times, and high-income individuals tend to live in areas with sparse road networks. If this is case, it implies that current urban infrastructure may contribute to the inequalities of new mobility

services. For the same reason, it is unclear if the relationships found in this study will generalize to other cities, of which urban forms (e.g., road network, crime rate, employment density, etc.) differ from Seattle.

Example 1 and Example 2 both examined the equity of service provision in terms of a single indicator (service area coverage and waiting times). These two examples sought to evaluate *equity of opportunity*. While waiting times and docking station locations are important, they do not fully imply the disparities in actual use. For example, individuals without a smartphone cannot use shared bikes even if the docking station is located close to them. The following example approaches this problem from another perspective, namely, focusing on evaluating the *equity of outcome*.

**Example 3: Inequalities in usage of a public bicycle sharing** Ogilvie and Goodman [38] explored the correlation between the usage of a bikeshare system in London and socioeconomic attributes. The dataset they use is the anonymized user registration data of a bikeshare system. They examined two dependent variables separately: mean number of trips made by a registered user per month (continuous) and whether a registered user has ever made any trip (binary). They constructed a series of independent variables from the registration data, including gender, place of residence, income deprivation (English Indices of Deprivation) at the level of the Lower Super Output Area (LSOA, a base unit of UK census data), non-White percentage of residential LSOA, distance from residence to nearest bike station, number of stations within 250m of residence, month of registration, etc.

The authors employed linear regression to examine the relationship between “mean number of trips per month” and independent variables, and logistic regression to examine the relationship between “ever made any trip” and independent variables. Spatial autocorrelation was accounted for using maximum likelihood estimation. Regression results showed that female users made fewer trips than males per month and users in more deprived areas are less likely to live close to a bike station. After adjusting for the distance from residential area to station, those in more deprived areas made more trips than those in the least deprived areas. They concluded that disparities exist in usage of the system across population, and the system has potentials to fulfill unmet need if services expand to more deprived areas.

This study examined the equity of individual-level bikeshare usage. Although the authors found that female users tend to have fewer trips than male users, they cannot determine the cause of this observed relationship. It could be that females tend to have fewer bike trips at night or to regions with high crime rates due to safety concerns. After adjusting for crime rate or time of day, the association between bike usage and gender may change. This brings up another limitation of this research resulted from the use of automatically collected data from bikeshare system. The authors did not have control over the data collection process, so what they could study is also limited. Moreover, constrained by the data availability, the authors had to use area-level socioeconomic variables derived from the postcode of registration debit or credit card. It is thus unclear whether the conclusions would still hold if individual-level variables were available. The temporal scale of the data (7 months) limits the possibility to explore seasonal effects of bike usage.

**Advantages and limitations of correlational research for equity studies** By using correlational research, researchers can examine the relationships between transportation provision (or usage) and a wide range of variables collected from various sources. This is especially true when large amount of automatically collected data (e.g., smart card data, bikeshare trip database) is available. Correlational research is appropriate when researchers are unable to manipulate the variables due to practical or ethical reasons. For example, in equity studies, area-level variables such as average income of a CBG is not controllable.

One limitation of correlational research is that a significant correlation does not allow the researcher to determine a causal relationship, because there could be many factors that the researcher did not study but contribute to the correlation. And these factors could be independent of mobility service provision. Further inquiry is needed to corroborate the findings from a correlational study. Another limitation is that correlational research heavily depends on data availability and data quality, as discussed in Example 3.

## 4.2 Experimental Research

Experimental research enables researchers to identify causal relationship. In an experiment design, the researcher seeks to fully control the environment conditions so that variables of interest can be manipulated, while other variables are controlled (or randomized) across conditions. In this way, the effects of variables of interest can be tested by comparing between two or more conditions. Unlike correlational research, experimental research strictly controls for the impacts of variables not of interest, thus allowing the effects of variables of interest to be measured upon the outcome [55].

**Example: Racial and gender discrimination in transportation network companies** Ge et al. [11] studied the racial and gender discrimination in Transportation Network Companies (TNCs). They undertook two randomized control trails, hailing about 1500 rides in Seattle and Boston and recording service quality indicators. In the Seattle experiment, the treatment is race. Eight RAs (two African American females, two white females, two African American males, and two white males) were hired to request rides. Measures including estimated waiting times, acceptance time (time between trip request and acceptance), actual waiting times (time between acceptance and arrival), trip cancellation rate, trip duration, costs, and ratings were recorded by screenshots for each trip. To control for variables not of interest, the authors adopted a number of strategies. The RAs are undergraduate students, avoiding confounding factors such as age. They were given the identical smartphones using the same carriers, and received the same data collection instructions. The RAs were instructed to minimize their interactions with the driver, preventing the introduction of factors that influence ratings and travel time. Specific routes were developed to control for pick-up locations and travel duration. These routes were randomly assigned to RAs. RAs were also instructed to travel after evening rush hours from Mondays to Thursdays. Ordinary least squares regression (OLS) results showed that acceptance time is longer for African American riders than white riders for both UberX and Lyft.

In the Boston experiment, the authors adopted a within-group design. They hired eight RAs with a range of ethnic backgrounds summoning UberX or Lyft rides in Boston, each requesting rides under a “white-sounding” name and a “distinctively black name”. This change in design aims to control the differences in data collection practices among RAs. In this case, the treatment is whether the rider has a black sounding name. Other aspects of experiment design are similar to those of the Seattle experiment. OLS results showed that riders with African American-sounding names experienced more frequent trip cancellations, and that African American males have higher cancellation rates than white males. Further analysis revealed that trip cancellations concentrated in pick-up locations with low population density. They concluded that racial discrimination exists in TNC services in Seattle and Boston.

**Advantages and limitations of experimental research for equity studies** Experimental research allows for drawing causal conclusions. This is because experiments are conducted in controlled conditions and researchers can claim that the changes in outcomes are caused by the variable of interest.

Experimental research has notable limitations. First, the requirement that controlling all variables that might influence the outcomes is sometimes not realistic. This is especially true for experiments conducted in a natural environment. For example, in Ge et al.’s Seattle experiment [11], there are variances in the data collection practices (e.g., the time lag between taking screenshots and sending requests) among RAs. This influences the measurement of outcome variables. Second, compared to automatically collected data or survey data, experiments are often not able to produce large amount of data. Data collection in experiments is often expensive and labor-intensive. Finally, although experimental research can determine causal effects (e.g., racial discrimination exists in TNC services in Seattle), it cannot unveil the reasons why the outcome occurred (e.g., why TNC drivers discriminate against certain races). Further investigation through other research methods (e.g., interviews) is needed to understand the phenomenon.



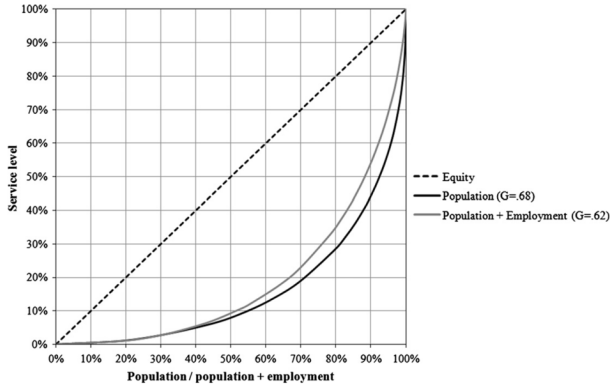


Figure 3: Use Lorenz curves to compare the equity of public transport service level to demand (population and employment) [28].

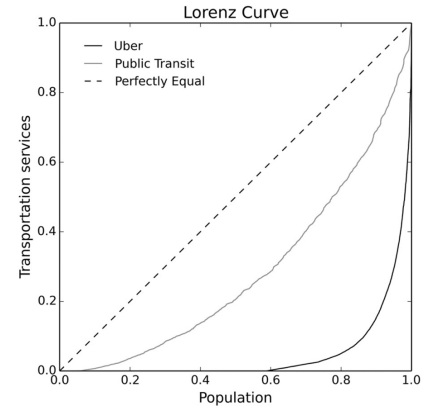


Figure 4: Use Lorenz curves to compare the equity of public transport and Uber service level to population [54].

### 4.3 Equity Metrics

Metrics that measure the distribution of some mobility system impacts (e.g., service level) have been widely adopted in transportation equity evaluation. Unlike statistical tests which focus on the discovery of discrimination or inequalities, metrics directly gauge the degree of equity by a single value. Equity metrics used in transportation equity research differ but overlap with those used in fairness in machine learning research. The similarities between them partially arise from the fact that both fields borrowed ideas from other domains such as social welfare and economics. For example, Gini coefficient (or Gini index), initially proposed to represent cumulative income and wealth distribution across a population, is one of the most popular equity metrics used in transportation to gauge the equity of transportation resource allocation [8]. However, Gini index has not yet received much attention in machine learning community [57]. Perhaps this is because fairness in machine learning research has primarily concentrated on classification problems that used in credit scoring, profiling of potential suspects, hiring, etc., for which other metrics are more appropriate. On the other hand, there are a few metrics, such as the “80% Rule” [58], were used in both communities [59]. The following examples introduce the use of Gini index and the “80% Rule” in transportation equity.

**Example 1: Using Lorenz curves to assess public transport equity in Melbourne** Delbosc and Currie [28] proposed to use Gini index as an equity metric of public transit service provision. A Lorenz curve is a graphical representation of Gini index. The figure below (see Figure 3) illustrates an example of a Lorenz curve representing the cumulative income across a population. The perfect equitable income distribution is plotted as the dashed line (line of equity) and an inequitable distribution of income is represented by the solid curve (Lorenz curve). A point on the solid curve can be interpreted as X percent (e.g., 70%) of population shares about Y percent (e.g., 25%) of the total income of the population. Gini index is the ratio of the area between the line of equity and the Lorenz curve (A), divided by the total area under the line of equity (A+B).

Delbosc and Currie applied a gini index to compare the equity of public transport service level to a proxy of demand (population and employment) in Melbourne. Service level of a census tract is expressed as a composite index taking into account bus, tram, and rail service areas and frequency. Using the service level index and the population of each census tract, the authors generated the first Lorenz curve (black solid curve) as shown in Figure 3. The Gini index is 0.68 for overall population in Melbourne. This can be interpreted as 70% of the population shares 19% of the public transport services. A second Lorenz curve (a grey solid curve) was calculated, taking into account the employment density. The Gini index for total population and employment is 0.62, appearing

Borough	Weekday		Weekend	
	Public transit	Public transit + Uber	Public transit	Public transit + Uber
Bronx	0.5803	0.5802	0.6079	0.6073
Brooklyn	0.5738	0.5725	0.5852	0.5783
Manhattan	0.6324	0.6276	0.6625	0.6352
Queens	0.7425	0.7424	0.7359	0.7332
Staten Island	0.4336	0.4335	0.4951	0.4949
Whole city	0.6653	0.6321	0.6429	0.6349

Table 2: Gini coefficient without and with Uber [54]

more equitable than the first curve. Nevertheless, these two curves suggest that inequalities exist in public transit, with only a small portion of the population enjoying the majority of transit services.

In this example, the Gini index serves as a measure of horizontal equity, that is, providing equal resources to those equal in need. The need for transportation supply of each census tract is approximated by population and employment density. So the perfect equitable distribution is that every unit of population and jobs shares the same transport resources. The need of special demographic groups (vertical equity) is not considered.

**Example 2: Using Lorenz curves to access the equity of Uber and public transit in New York City** Most recently, Jin et al. [54] employed Lorenz curves and Gini index to study the equity of Uber services in New York City (see Figure 4). They calculated service level for Uber and public transit using a similar approach as Delbosch and Currie [28]. Their results suggested that Uber is less equitable than public transit: 20% of population shares the 95% of Uber services.

They further compared Gini indexes of different boroughs for public transit and public transit + Uber (see Table 2). The results (see Table 2) shows that with Uber, the Gini index of the whole city reduced by about 0.03 on weekdays and by about 0.008 on weekends. This implies that Uber has insignificant impact on the transportation equity of New York City.

This study exemplifies how Gini index can be used to compare transportation equity across regions and across modes. This is possible because Gini index has several desirable features: it does not depend on the size of the population, the overall transit supply level, or the geographic units. For example, Gini index can be used to examine the equity of a neighborhood, a city, or a country. And it enables the comparison of equity between a city with high level of transit supply and one with low supply.

One limitation of Gini index is its heavy reliance on data. As Jin et al. noted, the main reason to choose New York City as study area is data availability. Beyond availability, all data sources have limitations (e.g., census data is not up-to-date) that would be calculated into Gini index. Another limitation lies in the way the service level is calculated. Studies that employed Gini indexes tend to use different methods to calculate service level [30]. It is unclear whether changing the service level indicator will significantly affect Gini index. These two limitations suggest that Gini indexes should be interpreted with caution.

**Example 3: Evaluation of the equity of bikeshare system accessibility** Meng [39] applied the “80% Rule” to evaluate access equity of a bikeshare system in Chicago. The “80% Rule” was advocated by the US Equal Employment Opportunity Commission to detect disparate impacts in employee selection procedures. The 80% Rule states that if the selection rate for minorities is less than 80% of the rate of non-minorities, the procedure is deemed to be discriminatory [58]. Similar to Ursaki and Aultman-Hall (see Example 1 of Section 4.1), the analysis is based on the locations of docking stations. The author created a 0.25-mile buffer around each station as service area, and calculated the demographic characteristics (i.e., race, gender, education, language proficiency, and income level) of population inside each service area. For each station, the equity metric based on the 80% Rule is calculated as follows:

$$Ratio = \frac{Number\ of\ minorities / Total\ number\ of\ minorities}{Number\ of\ non - minorities / Total\ number\ of\ non - minorities} \quad (19)$$

The results show that more than 33% of the stations have ratios below 0.8 for all demographic characteristics (except for gender) under examination.

There are several limitations of this study. First, instead of providing a city-level ratio, the authors computed station-level ratios and examined equity using the percentage of stations that violate the 80% Rule. This approach is problematic when the stations are not equally distributed. It is possible that a majority of the stations are all located in a small portion of the city and they tend to have similar ratios. Second, docking station placement cannot sufficiently represent access to bikeshare, as discussed in Example 1 of Section 4.1. Despite these weaknesses, this study serves as a typical example of using fairness metrics to evaluate vertical equity in new mobility systems.

**Advantages and limitations of equity metrics** Equity metric provides a single measure of equity, making it possible to track trends over time and conduct comparative studies between cities. It is easier for non-experts to interpret compared to statistical tests, therefore suited for conveying evaluation results to broader audience.

However, the reliability of metrics heavily depends on the quality of data sources. Moreover, different metrics often reflect competing goals. For example, Gini index measures horizontal equity, emphasizing individuals with equal ability or need gets equal resources. The 80% Rule shares a similar spirit of group fairness [59], which advocates for equal resource distribution across difference demographic groups. The choice of metrics may significantly affect evaluation results, so the use of multiple metrics is important.

#### 4.4 Other methods

Apart from the three research methods described above, surveys, interviews, and focus groups have been used for transportation equity studies. These methods can be used to develop a deeper understanding of why inequalities exist based on the opinions, attitudes, and experiences from stakeholders of mobility systems. For example, McNeil, Nathan, et al. [47] conducted a survey of residents living in underserved neighborhoods with bikeshare stations. The findings revealed that minority respondents have more barriers, for example, costs of membership, to using shared bikes than non-minorities. This helps to explain why providing adequate spatial access to disadvantaged neighborhoods alone is not enough to address the disparities in actual use.

## 5 Transportation Equity and Fairness in Machine learning

In examining the fairness (equity) definitions from transportation equity community and fair machine learning (FairML) community, we observe that a natural mapping between them can be established, though further effort is needed to create a consistent mapping between concepts in one domain to the other. *Horizontal equity* echoes the spirit of *individual fairness* (similar people should be treated similarly). *Vertical equity* resembles *group fairness* (sensitive attributes should be independent from outcomes). This is true in cities where there is an uneven distribution of transport supply across different socioeconomic groups. Vertical equity encourages compensating for such inequalities by policies favoring disadvantaged groups. This aligns with group fairness that the level of transportation supply in a city should be the same across different groups. Vertical equity and group fairness are only “roughly” related because by definition, group fairness stresses “independence” between sensitive attributes and outcome, whereas vertical equity does not.

The most commonly used method for evaluating horizontal equity is Gini index. It has not attracted much attention in machine learning community. This may be partially due to the fact that not much attention has been paid to resource allocation problems in fair machine learning research. On the other hand, machine learning

community has developed a few metrics for individual fairness. Individual fairness requires that the “similarity” between a pair of individuals from two demographic groups respectively has to be defined. For example, in making hiring decisions, the algorithm has to possess perfect knowledge of how to compare the “qualification” of two individuals. This is often not realistic in practice and we have to come up with a suitable similarity metric that is best agreed upon among domain experts of a task. Theoretically, individual fairness can be used to evaluate horizontal equity. For example, in a simplified shared bike allocation problem, we use population and employment density as the demand for bikes. Then the differences in demand between two areas  $a$  and  $b$  can be expressed as  $d(a, b)$  according to some similarity function  $d$ . Suppose we have an algorithm assigning bikes to areas, the number of bikes that area  $a$  and  $b$  will get is  $f(a)$  and  $f(b)$ , respectively. A fair allocation satisfying individual fairness requires that for every two pairs of areas in the city:  $D(f(a), f(b)) \leq d(a, b)$ , where  $D$  is another similarity function. The difficulty again, lies in the fact that we do not have perfect knowledge to determine the similarity in demand between two areas.

The majority of transportation equity research focuses on vertical equity. Likewise, more attention has been devoted to group fairness than individual fairness in machine learning community. Transportation equity heavily employs statistical tests for equity analysis, which is appropriate for discovering unfairness. Machine learning uses fairness metrics much more often, because metrics allow researchers to reduce achieving fairness goals to a much simpler problem: minimizing a value that represents unfairness. This is also valid in terms of algorithm design. In fact, some metrics, such as the 80% Rule, have been used in both communities. This connection may open great possibilities for bridging these two domains.

Fair machine learning community focuses almost exclusively on methods, whereas transportation equity concerns more about applications, policies, and interventions. Although fair machine learning approaches hold great promises in optimizing resource allocation in mobility settings, there is a long way to go to design, deploy, and evaluate a fairness-aware data-driven system as a real-world application. At the end of this paper, I hope to highlight the urgency of convergence of these two fields. Ultimately, researchers with knowledge in both fields, practitioners, policy-makers, and citizens should work together towards a common goal: a fair and effective transportation system for all citizens.

## 6 Conclusion

This paper summarized the findings and methods of equity studies in mobility systems, with a focus on new mobility systems. For new mobility services, it is generally agreed that disparities exist in the access and use of docked bikeshare system, but the equity implications of ride-hailing are still unclear. Further research is needed to understand how to deliver a more equitable new mobility system to serve the need of different groups. Many research methods have been employed in transportation equity studies. Different methods vary in their objectives, strengths and weaknesses. Correlational research can exploit a wide range of data sources and discover associations among many factors, but it cannot determine causal relationships. Equity metrics enable comparative studies among cities and assessment of changes over time, but their reliability is highly dependent on data. Experimental research can produce reliable findings, but is expensive and difficult to control all extraneous variables. The choice of research methods depends on research goals, and multiple methods can be used together to complement each other.

Given the similarities in objectives, concepts, and methods between transportation equity community and fairness in machine learning community, bridging these two domains together holds promise to enable multi-disciplinary breakthroughs.

## References

- [1] Chao Zhang, Keyang Zhang, Quan Yuan, Haoruo Peng, Yu Zheng, Tim Hanratty, Shaowen Wang, and Jiawei Han. Regions, periods, activities: Uncovering urban dynamics via cross-modal representation learning. In *Proceedings of the 26th International Conference on World Wide Web*, pages 361–370. International World Wide Web Conferences Steering Committee, 2017.
- [2] Jason Shafrin, Jeff Sullivan, Dana P Goldman, and Thomas M Gill. The association between observed mobility and quality of life in the near elderly. *PloS one*, 12(8):e0182920, 2017.
- [3] Ya-Fen Chan, Shou-En Lu, Bill Howe, Hendrik Tieben, Theresa Hoeft, and Jürgen Unützer. Screening and follow-up monitoring for substance use in primary care: an exploration of rural–urban variations. *Journal of general internal medicine*, 31(2):215–222, 2016.
- [4] Qi Wang, Nolan Edward Phillips, Mario L Small, and Robert J Sampson. Urban mobility and neighborhood isolation in america’s 50 largest cities. *Proceedings of the National Academy of Sciences*, 115(30):7735–7740, 2018.
- [5] Todd Litman. *Evaluating transportation equity*. Victoria Transport Policy Institute, 2018.
- [6] Susan Shaheen, Corwin Bell, Adam Cohen, and Balaji Yelchuru. Travel behavior: Shared mobility and transportation equity. Technical report, U.S. Department of Transportation, 2017.
- [7] Chelsey Palmateer and David M Levinson. Justice, exclusion, and equity: An analysis of 48 us metropolitan areas. Technical report, University of Minnesota: Nexus Research Group, 2017.
- [8] Anthony Michael Ricciardi, Jianhong Cecilia Xia, and Graham Currie. Exploring public transport equity between separate disadvantaged cohorts: a case study in perth, australia. *Journal of transport geography*, 43:111–122, 2015.
- [9] Raj Chetty, Nathaniel Hendren, Frina Lin, Jeremy Majerovitz, and Benjamin Scuderi. Childhood environment and gender gaps in adulthood. *American Economic Review*, 106(5):282–88, 2016.
- [10] Todd Goldman and Roger Gorham. Sustainable urban transport: Four innovative directions. *Technology in society*, 28(1-2):261–273, 2006.
- [11] Yanbo Ge, Christopher R Knittel, Don MacKenzie, and Stephen Zoepf. Racial and gender discrimination in transportation network companies. Technical report, National Bureau of Economic Research, 2016.
- [12] Indre Zliobaite. A survey on measuring indirect discrimination in machine learning. *CoRR*, abs/1511.00148, 2015.
- [13] Kevin Petrasic, Benjamin Saul, James Greig, Matthew Bornfreund, and Katherine Lamberth. Algorithms and bias: What lenders need to know. *White & Case*, 2017.
- [14] Claire Cain Miller. When algorithms discriminate. *The New York Times*, 9, 2015.
- [15] Cynthia Rudin. Predictive policing using machine learning to detect patterns of crime. *Wired Magazine*, August, 2013.
- [16] Solon Barocas, Moritz Hardt, and Arvind Narayanan. Fairness in machine learning. *NIPS Tutorial*, 2017.
- [17] Jessica Guynn. Google photos labeled black people ‘gorillas’. *USA Today*, 1, 2015.
- [18] David Ingold and Spencer Soper. Amazon doesn’t consider the race of its customers. should it. *Bloomberg*, 2016.
- [19] Bruno Lepri, Jacopo Staiano, David Sangokoya, Emmanuel Letouzé, and Nuria Oliver. The tyranny of data? the bright and dark sides of data-driven decision-making for social good. In *Transparent data mining for big and small data*, pages 3–24. Springer, 2017.
- [20] Solon Barocas and Andrew D Selbst. Big data’s disparate impact. *Calif. L. Rev.*, 104:671, 2016.
- [21] Aylin Caliskan, Joanna J Bryson, and Arvind Narayanan. Semantics derived automatically from language corpora contain human-like biases. *Science*, 356(6334):183–186, 2017.
- [22] Jeffrey Pennington, Richard Socher, and Christopher Manning. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*, pages 1532–1543, 2014.
- [23] Alexandra Chouldechova and Aaron Roth. The frontiers of fairness in machine learning. *CoRR*, abs/1810.08810, 2018.

- [24] Kristian Lum and William Isaac. To predict and serve? *Significance*, 13(5):14–19, 2016.
- [25] Meg Young, Luke Rodriguez, Emily Keller, Feiyang Sun, Boyang Sa, Jan Whittington, and Bill Howe. Beyond open vs. closed: Balancing individual privacy and public accountability in data sharing. In *Proceedings of the Conference on Fairness, Accountability, and Transparency*, pages 191–200. ACM, 2019.
- [26] Bruno Lepri, Nuria Oliver, Emmanuel Letouzé, Alex Pentland, and Patrick Vinck. Fair, transparent, and accountable algorithmic decision-making processes. *Philosophy & Technology*, 31(4):611–627, 2018.
- [27] Jenna Burrell. How the machine ‘thinks’: Understanding opacity in machine learning algorithms. *Big Data & Society*, 3(1):2053951715622512, 2016.
- [28] Alexa Delbosc and Graham Currie. Using lorenz curves to assess public transport equity. *Journal of Transport Geography*, 19(6):1252–1259, 2011.
- [29] Karen Lucas, Bert Van Wee, and Kees Maat. A method to evaluate equitable accessibility: combining ethical theories and accessibility-based approaches. *Transportation*, 43(3):473–490, 2016.
- [30] Tao Feng and Harry JP Timmermans. Trade-offs between mobility and equity maximization under environmental capacity constraints: A case study of an integrated multi-objective model. *Transportation Research Part C: Emerging Technologies*, 43:267–279, 2014.
- [31] Nikolas Thomopoulos, Susan Grant-Muller, and MR Tight. Incorporating equity considerations in transport infrastructure evaluation: Current practice and a proposed methodology. *Evaluation and program planning*, 32(4):351–359, 2009.
- [32] Steven Raphael and Lorien Rice. Car ownership, employment, and earnings. *Journal of Urban Economics*, 52(1):109–130, 2002.
- [33] Robert D Bullard. Addressing urban transportation equity in the united states. *Fordham Urb. LJ*, 31:1183, 2003.
- [34] Ahmed El-Geneidy, David Levinson, Ehab Diab, Genevieve Boisjoly, David Verbich, and Charis Loong. The cost of equity: Assessing transit accessibility and social disparity using total travel cost. *Transportation Research Part A: Policy and Practice*, 91:302–316, 2016.
- [35] Eduardo Alcantara Vasconcellos. Urban transport policies in brazil: The creation of a discriminatory mobility system. *Journal of transport geography*, 67, 2018.
- [36] Stuart Cohen and Clarrissa Cabansagan. A framework for equity in new mobility, 2017.
- [37] Kevin Manaugh, Madhav G Badami, and Ahmed M El-Geneidy. Integrating social equity into urban transportation planning: A critical evaluation of equity objectives and measures in transportation plans in north america. *Transport policy*, 37:167–176, 2015.
- [38] Flora Ogilvie and Anna Goodman. Inequalities in usage of a public bicycle sharing scheme: socio-demographic predictors of uptake and usage of the london (uk) cycle hire scheme. *Preventive medicine*, 55(1):40–45, 2012.
- [39] Chao Meng. Evaluation of the equity of bikeshare system accessibility: A case study of chicago. Technical report, Georgia Institute of Technology, 2018.
- [40] Julia Ursaki, Lisa Aultman-Hall, et al. Quantifying the equity of bikeshare access in us cities. Technical report, University of Vermont. Transportation Research Center, 2015.
- [41] C Scott Smith, Jun-Seok Oh, and Cheyenne Lei. Exploring the equity dimensions of us bicycle sharing systems. Technical report, Western Michigan University. Transportation Research Center for Livable, 2015.
- [42] Xuefeng Li, Yong Zhang, Li Sun, and Qiyang Liu. Free-floating bike sharing in jiangsu: Users’ behaviors and influencing factors. *Energies*, 11(7):1664, 2018.
- [43] Stephen J Mooney, Kate Hosford, Bill Howe, An Yan, Meghan Winters, Alon Bassok, and Jana A Hirsch. Freedom from the station: Spatial equity in access to dockless bike share. *Journal of Transport Geography*, 74:91–96, 2019.
- [44] Chengcheng Xu, Junyi Ji, and Pan Liu. The station-free sharing bike demand forecasting with a deep learning approach and large-scale datasets. *Transportation research part C: emerging technologies*, 95:47–60, 2018.

- [45] David William Daddio and N Mcdonald. Maximizing bicycle sharing: an empirical analysis of capital bikeshare usage. *University of North Carolina at Chapel Hill*, 8, 2012.
- [46] R Alexander Rixey. Station-level forecasting of bikesharing ridership: station network effects in three us systems. *Transportation Research Record*, 2387(1):46–55, 2013.
- [47] Nathan McNeil, Jennifer Dill, John MacArthur, Joseph Broach, and Steven Howland. Breaking barriers to bike share: Insights from residents of traditionally underserved neighborhoods. *Transportation Research and Education Center (TREC)*, 2017.
- [48] SS Kretman, DC Johnson, and WP Smith. Bringing bike share to a low-income community: Lessons learned through community engagement, minneapolis, minnesota. *Preventing Chronic Disease*, 2011.
- [49] Anne Elizabeth Brown. *Ridehail revolution: Ridehail travel and equity in Los Angeles*. PhD thesis, UCLA, 2018.
- [50] Ryan Hughes and Don MacKenzie. Transportation network company wait times in greater seattle, and relationship to socioeconomic indicators. *Journal of Transport Geography*, 56:36–44, 2016.
- [51] Mingshu Wang and Lan Mu. Spatial disparities of uber accessibility: An exploratory analysis in atlanta, usa. *Computers, Environment and Urban Systems*, 67:169–175, 2018.
- [52] Jennifer Stark and Nicholas Diakopoulos. Uber seems to offer better service in areas with more white people. that raises some tough questions. *The Washington Post*, 2016.
- [53] Scott Scott Russell Middleton. *Discrimination, regulation, and design in ridehailing*. PhD thesis, Massachusetts Institute of Technology, 2018.
- [54] Scarlett T Jin, Hui Kong, and Daniel Z Sui. Uber, public transit, and urban transportation equity: A case study in new york city. *The Professional Geographer*, pages 1–16, 2019.
- [55] Jonathan Lazar, Jinjuan Heidi Feng, and Harry Hochheiser. *Research methods in human-computer interaction*. Morgan Kaufmann, 2017.
- [56] A Stewart Fotheringham, Chris Brunson, and Martin Charlton. *Geographically weighted regression: the analysis of spatially varying relationships*. John Wiley & Sons, 2003.
- [57] Till Speicher, Hoda Heidari, Nina Grgic-Hlaca, Krishna P Gummadi, Adish Singla, Adrian Weller, and Muhammad Bilal Zafar. A unified approach to quantifying algorithmic unfairness: Measuring individual & group unfairness via inequality indices. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pages 2239–2248. ACM, 2018.
- [58] The U.S. EEOC. Uniform guidelines on employee selection procedures. *March 2, 1979*, 43:111–122, 1979.
- [59] Michael Feldman, Sorelle A Friedler, John Moeller, Carlos Scheidegger, and Suresh Venkatasubramanian. Certifying and removing disparate impact. In *Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 259–268. ACM, 2015.