

# Do infrastructure deserts exist? Measuring and mapping infrastructure equity: A case study in Dallas, Texas, USA

Zheng Li<sup>a,\*</sup>, Xinlei Wang<sup>b</sup>, Jessie Zarazaga<sup>a,c</sup>, Janille Smith-Colin<sup>a</sup>, Barbara Minsker<sup>a,c</sup>

<sup>a</sup> Department of Civil Environmental Engineering, Southern Methodist University, United States of America

<sup>b</sup> Department of Statistical Science, Southern Methodist University, United States of America

<sup>c</sup> Hunt Institute for Engineering and Humanity, Southern Methodist University, United States of America

## ARTICLE INFO

### Keywords:

Infrastructure equity

Neighborhood infrastructure

Statistical model

Data-driven

## ABSTRACT

Neighborhood infrastructure, such as sidewalks, medical facilities, public transit, community gathering places, and tree canopy, provides essential support for safe, healthy, and resilient communities. This paper examines the presence and condition of neighborhood infrastructure and shows that “infrastructure deserts” exist, which are low-income neighborhoods with significantly more deficient infrastructure. A generalized data-driven framework is developed and implemented at the street-level for 12 types of neighborhood infrastructure in Dallas, Texas. The results show significant infrastructure inequities across income levels for most types of infrastructure. Statistical inference predicts (with 95 % confidence) that low-income neighborhoods are 2.0 to 3.5 times more likely to have highly deficient infrastructure (8 or more deficient infrastructure types) than high-income areas and 1.4 to 2.4 times more likely to have highly deficient infrastructure than middle-income neighborhoods. This paper addresses the methodological challenge of considering multiple infrastructure types and provides a guide for infrastructure investment prioritization.

## 1. Introduction

This paper develops and implements a generalizable data-driven framework for assessing the condition and equity of neighborhood infrastructure. Neighborhood infrastructure is a system of relatively small-scale physical structures and service facilities that play an essential role in improving residents' lives, health, safety, and social justice (MacDonnell et al., 2004; Reduction & Recovery(GFDRR), 2017). Understanding neighborhood infrastructure is necessary to support community wellbeing and prioritize future infrastructure investments. While efforts have been made to study the condition and impact of individual types of neighborhood infrastructure, the focus and scope of such efforts remain relatively singular (limited to one or a few infrastructure types) and fail to treat infrastructure as a diverse, multi-component system. For example, leisure walking experiences are affected by multiple infrastructure features such as sidewalk condition, crosswalk presence at intersections, street pavement condition, and street tree cover. So, to truly understand the overall impact of infrastructure on a neighborhood, multiple infrastructure types must be assessed through an integrated approach.

The forms of neighborhood infrastructure can be physical structures

(such as sidewalks, crosswalks, pedestrian trails, street lights, street tree canopy) or facilities (such as hospitals) located or operated outside of the neighborhood as community services. These facilities provide human development support (such as health clinics, financial facilities), public services (such as transportation, schools, libraries, internet), and shared space for social gatherings and recreational activities (such as parks, trails, community centers, noise walls). Neighborhood infrastructure types can be quite diverse and vary from community to community depending on geophysical, socio-cultural, and economic factors. Therefore, the estimation of impacts, changes, and future development of neighborhood infrastructure not only requires a thorough understanding of infrastructure conditions but also the community's socio-economic and cultural settings.

Previous studies have shown the importance of neighborhood infrastructure for human health, community growth, and community safety. For example, neighborhood infrastructure, particularly sidewalks, streets, and access to local destinations such as grocery stores, parks, and recreation facilities, have impacts on obesity (Booth et al., 2005); related chronic health outcomes (Roux & Mair, 2010); health behaviors (Gordon-Larsen, 2006; Leventhal & Brooks-Gunn, 2000; McCormack & Shiell, 2011); mental health outcomes (Leventhal &

\* Corresponding author.

E-mail address: [zli1@smu.edu](mailto:zli1@smu.edu) (Z. Li).

<https://doi.org/10.1016/j.cities.2022.103927>

Received 11 December 2021; Received in revised form 13 June 2022; Accepted 5 August 2022

Available online 19 August 2022

0264-2751/© 2022 Elsevier Ltd. All rights reserved.

Brooks-Gunn, 2000; Ross, 2000; Roux & Mair, 2010); and social well-being outcomes (Handy et al., 2002; McCormack & Shiell, 2011; Rogers et al., 2013). Pedestrian-friendly streets, open green spaces, and well-maintained neighborhood infrastructure (such as sidewalks, crosswalks, healthcare, food stores, and community centers) not only promote healthy activities such as walking and bicycling (Handy et al., 2002) but also enhance social interactions (Gunn et al., 2017), social cohesion and social capital (Rogers et al., 2013). These factors facilitate the organic growth of community attitudes toward healthy and active lifestyles (Gunn et al., 2017; Srinivasan et al., 2003; Ulmer et al., 2016). Furthermore, studies have shown the positive influence of well-established neighborhood infrastructure (such as sidewalks, crosswalks, street lighting) on perceived and actual safety from crime or traffic-related events (Farrington & Welsh, 2002; Hong & Chen, 2014; Hwang et al., 2017). For example, well-designed crosswalks and sidewalks help reduce pedestrian-vehicle crashes (Hwang et al., 2017).

Conversely, the lack of quality neighborhood infrastructure can increase community susceptibility to natural disasters and chronic economic crises (Aldrich & Meyer, 2015). Lack of neighborhood infrastructure has also been considered a critical indicator of social injustice based on three primary allocation principles (Deutsch, 1975): equity, equality, and need. Equity is an issue of distributive justice and calls for fairness where service are provided based on rectifying the existing patterns of service distribution (Lucy, 1981). Equality means that everyone receives the same public service (Lucy, 1981), which can lead to more harmonious social relationships (Deutsch, 1975) but may not meet individual needs. The concept of need is consistent with the idea that those needing more service should receive more rather than less (Lucy, 1981). Each of these three principles operates in a specific domain. For neighborhood infrastructure, including public utilities, parks, and facilities, equality is often impossible to achieve in the sense of equal access because of the variation in community development and terrain. The need is also likely tied to population and distributed geographically (Lucy, 1981). Therefore, analyzing the equitable distribution of infrastructure services is a better approach for understanding the present condition and future investments. Thus we primarily focus on the equity aspect of neighborhood infrastructure distribution in this work.

One way of studying infrastructure equity, as suggested by the U.S. Department of Transportation, is to compare the infrastructure characteristics or conditions in neighborhoods with high concentrations of socially vulnerable populations (such as low-income households, minorities, and car-free households) compared to those in adjacent neighborhoods or to regional averages (Forkenbrock et al., 2004). Following this guideline, many researchers have evaluated infrastructure conditions and discovered infrastructure inequities across many individual types of neighborhood infrastructure. Studies have shown economic and ethnic disparities in walkability, street trees, public transportation, parks, pedestrian crosswalks, and trails (Aman & Smith-Colin, 2020; Borowski et al., 2018; Hirsch et al., 2017; Wolch et al., 2005; Xiao et al., 2017). Grocery stores and farmers' markets have also been widely studied in the realm of "food deserts," as areas showing substantial inequities across different socio-economic and racial groups (Hilmers et al., 2012). While these studies show the importance of individual infrastructure types and their impacts on communities, the presence and impact of multiple deficient types of physical infrastructure on a community is not yet known. Hence, a systematic condition assessment is needed at the community level to evaluate multiple neighborhood infrastructure types and support examination of their aggregated impacts on the community.

Previous studies have assessed neighborhood infrastructure using a variety of measures. Inspecting infrastructure surface exteriors and identifying defects is the most common method to assess infrastructure such as streets and sidewalks (Frackelton et al., 2013; Shah et al., 2013). Researchers have also used proximity as a metric for assessing coverage of infrastructure services such as parks (Cutts et al., 2009; Nicholls,

2001; Wolch et al., 2005; Xiao et al., 2017), healthcare (Wang & Luo, 2005), public transportation system (bus stops, rail stations) (El-Geneidy et al., 2010; O'Sullivan & Morrall, 1996), and fresh food supplies (Gordon et al., 2011; Ploeg et al., 2009). In addition, measures derived from field audits, secondary data sources, and satellite imagery/videos have allowed the assessment of ground-based or hard-to-measure infrastructure such as neighborhood street walkability (Ewing et al., 2006; Frackelton et al., 2013; Parks & Schofer, 2006; Su et al., 2019), street tree canopy (Maco & McPherson, 2002; Roberts et al., 2019) internet speed (Perzynski et al., 2017), and street condition (Engineering-USACE, 2014; Shah et al., 2013). However, the integrated physical assessment of multiple neighborhood infrastructure types faces methodological challenges and limitations in effective implementation (Rollings et al., 2015). Due to the difficulty of gathering neighborhood-scale data of multiple infrastructure types on a large scale, most previous studies have primarily focused on either small-scale studies of one or several infrastructure types or large-scale studies at city or regional scale of only a single infrastructure type.

This study fills the gap between these two scales by providing quantitative infrastructure condition assessment at city scale and systematically combining multiple infrastructure types to show overall infrastructure condition using a data-driven framework. The results obtained from the framework are overlain with aspects of the neighborhood's socio-economic data (for this case study, income level) to evaluate infrastructure equity.

## 2. Materials and methods

Fig. 1 shows the generalized data-driven framework developed in this work to assess multiple neighborhood infrastructure types at the neighborhood level quantitatively and to explore infrastructure inequity in urban settings. The framework consists of three primary components: 1) compute neighborhood infrastructure deficiency by aggregating the presence and condition of each infrastructure type from street to neighborhood level; 2) compare infrastructure deficiency across income levels to identify the existence of infrastructure deserts; and 3) identify infrastructure inequity using statistical models. Each component is discussed in more detail in the following subsections.

The framework given in Fig. 1 has several benefits: 1) integrating street-level condition assessment with neighborhood-level socio-economic characteristics such as income; 2) enabling the addition of new infrastructure types while still maintaining robustness; and 3) providing a highly generalized approach that can be applied to other cities or regions with available data. The first step of the framework is to determine a proper spatial representation of neighborhood. Ideally, the chosen representation should naturally represent the residency boundary as a neighborhood. To aid in generalization, the framework is applied with a consistent spatial representation of neighborhoods as Census block groups for two main reasons. First, the size of a block group (typically ranges from 500 to 1000 housing units) and the cartographic representation approximates the overall size and geometry of a neighborhood. Secondly, by definition, census block groups share compatible population densities. Finally, the Census block group is the smallest administrative boundary for which the Census Bureau freely publishes sample data (Bureau, 2020). Therefore, the Census block group seamlessly aligns with the U.S. Census Bureau's socio-economic attributes.

The infrastructure condition data comes from both publicly accessible databases and computer-assisted programs/algorithms, as presented in Table 1. For instance, the noise walls data were obtained by inspecting Google Street View images along the major highways in Dallas and labeling the presence of recognized noise walls manually. Crosswalk data were collected at intersections of all residential streets and within school zones using a deep learning object recognition algorithm called YoLo3 (Redmon & Farhadi, 2018) and high-resolution Google satellite imagery. Finally, it is worth mentioning that internet service data are published at the Census tract level (Commission, 2019).

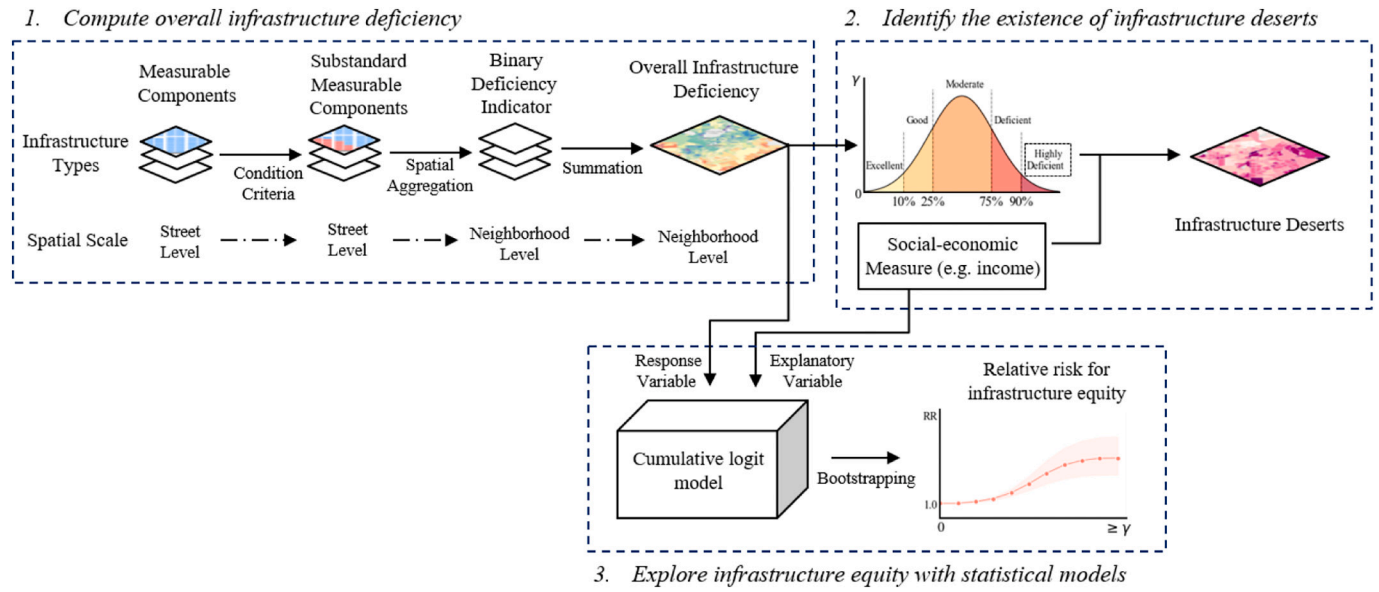


Fig. 1. Overview of the neighborhood infrastructure assessment framework.

To be spatially consistent with other infrastructure types, internet data were assumed to be uniform within Census tract and then allocated into Census block groups. Please see Appendix A.1 for more detailed implementation, descriptions, and data sources for each infrastructure type.

### 2.1. Compute overall infrastructure deficiency

The first step of the framework (shown in Step 1 of Fig. 1) examines each infrastructure type's condition and computes the overall neighborhood infrastructure deficiency. At the street level, metrics for measuring infrastructure condition vary across different infrastructure types and may vary in different cities. A neighborhood-level binary deficiency indicator ( $\theta$ ) is used to aggregate from street level measures within the neighborhood to represent deficiency of individual infrastructure type. To compute the binary infrastructure deficiency indicator, any quantifiable infrastructure components are identified within a neighborhood. Depending on the infrastructure type, multiple types of elements could be measured in the neighborhood.

This study focuses on two primary metrics for assessing deficiencies: 1) direct or secondary condition attributes associated with the available condition dataset (e.g., pavement and sidewalk cracking or damage); and 2) physical presence or service area provided by the infrastructure. For the first metric, it is straightforward to apply the criteria following existing guidelines and determining substandard measurable components without additional spatial transformations. For example, for pavement conditions, a numerical attribute called Pavement Condition Index (PCI) is often used to represent pavement segments' surface conditions, and empirical guidelines have suggested cutoffs for adequate pavement condition ratings (e.g., the pavement is classified as poor if its PCI is <55) (Engineering-USACE, 2014).

For the second type of metric, proximity calculations are needed to establish service areas before substandard criteria can be applied. A service area is computed around each infrastructure facility (such as hospitals, parks, food stores) that describes which residents in surrounding housing units can reach the facility within a predefined travel range or duration along the road network (Talen & Anselin, 1998). Thus, the residential households living within the service area are assumed to have access to the facilities and to receive related services. Residential households outside of service areas are assumed not to receive such services. The choice of predefined travel distances depends on the type of infrastructure facility. For instance, a 10-minute walking distance is

frequently used as the walkable distance to adjacent parks and pedestrian & bicycle trails (Kaczynski et al., 2009; Rollings et al., 2015). The distance of 1-mile or less is widely accepted as the threshold for determining sufficient access to fresh food supplies in food desert studies (Jiao et al., 2012; Ploeg et al., 2009).

Once all of the infrastructure facility locations and predefined travel distances are obtained, GIS is used to create the service area as spatial polygons. For the service area generation, the street network is used to ensure an accurate estimate of distances along the road network. Next, centroids of parcel data are used to represent and locate residential households. Lastly, all residential households living within and outside the service areas are identified using GIS intersect methods; this allows households outside of any service area to be identified as receiving substandard infrastructure services.

For this analysis, a fraction number ( $\mu$ ) is defined to represent substandard measurable components as a percentage of measurable infrastructure components within the same neighborhood:

$$\mu = \frac{M_{std}}{M} \quad (1-1)$$

where  $M_{std}$  represents substandard measurable infrastructure components within the neighborhood; and  $M$  is the total number of measurable components within the neighborhood. As shown in Eq. (2-2), the binary infrastructure deficiency indicator ( $\theta$ ) equals 1 if at least half (i.e., a majority) of the measurable components in a particular neighborhood are substandard ( $\mu \geq 0.5$ ). Otherwise  $\theta = 0$ .

$$\theta = \begin{cases} 0; & \text{if } \mu \geq 0.5 \\ 1; & \text{if } \mu < 0.5 \end{cases} \quad (1-2)$$

The benefit of using such indicators is to normalize all infrastructure measurements to the same scale of 0 or 1, which allows multiple binary indicators to be combined mathematically in later steps. The above procedure is repeated until  $\theta$  is obtained for all infrastructure types and then the overall infrastructure deficiency ( $\gamma$ ) of each neighborhood is computed as the summation of  $\theta$  (Eq. (1-3)):

$$\gamma = \sum_{i \in I} \theta_i \quad (1-3)$$

where  $I$  includes all the considered infrastructure types. The summation of multiple deficiency indicators into a single metric represents the overall neighborhood condition. Thus,  $\gamma$  ranges from zero to the total

**Table 1**  
Substandard criteria for neighborhood infrastructure types.

| Infrastructure Type          | Assessment Unit        | Substandard Criteria                                                                                                 | Criteria Reference                                    |
|------------------------------|------------------------|----------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------|
| Streets                      | Street Segments        | Pavement Condition Index (PCI) < 55                                                                                  | (Engineering-USACE, 2014)                             |
| Sidewalks                    | Sidewalk Segments      | Any existence of obstruction, damage, or missing segments                                                            | (Frackelton et al., 2013)                             |
| Internet                     | Residential Households | Internet speed <200 kbps in at least one direction                                                                   | (Commission, 2019)                                    |
| Crosswalks                   | Street Intersections   | Missing crosswalks at intersections with traffic lights or school zones                                              | (Zegeer et al., 2005)                                 |
| Noise Wall                   | Residential Households | Within 200 ft of the highway and no noise walls present                                                              | (Rochat & Reiter, 2016)                               |
| Street Tree Canopy           | Street Segments        | The average percentage of street segment covered by tree canopy <25 %                                                | (Maco & McPherson, 2002)                              |
| Public Transportation Access | Residential Households | Not within 5-min walking distance (0.4 km) of the bus stop or 10-min walking distance (0.8 km) of the rail station   | (El-Geneidy et al., 2010; O'Sullivan & Morrall, 1996) |
| Medical Facility Access      | Residential Households | Not within 2 miles (3.2 km) of major hospitals or 1-mile (1.6 km) of walk-in clinics or urgent care                  | (Ashwood et al., 2011; Bersamin et al., 2011)         |
| Bike & Pedestrian Trails     | Residential Households | Not within 10-min walking distance (0.8 km)                                                                          | –                                                     |
| Gathering Places             | Residential Households | Including public parks, libraries, farmer markets and community centers. Not within 10-min walking distance (0.8 km) | (Xiao et al., 2017)                                   |
| Food Access                  | Residential Households | Nearby food stores are not within a 1-mile distance (1.6 km)                                                         | (Gordon et al., 2011)                                 |
| Bank Access                  | Residential Households | Nearby bank branches are not within a 1-mile distance (1.6 km)                                                       | –                                                     |

number of infrastructure types considered. If a neighborhood does not have any deficient infrastructure types,  $\gamma=0$ . Finally, to aid in interpretation, a categorical representation of overall infrastructure deficiency is created based on the percentile of the resulting  $\gamma$ . As such, the resulting overall infrastructure deficiency values are defined as (1) *Excellent* ([0 % ~ 10 %]), (2) *Good* ([10 % ~ 25 %]), (3) *Moderate* ([25 %~75 %]), (4) *Deficient* ([75 %~90 %]) and, (5) *Highly deficient* ([90 %~100 %]).

## 2.2. Identify the existence of infrastructure deserts

The next step is to find neighborhoods that are both economically disadvantaged and significantly lacking in neighborhood infrastructure relative to wealthier neighborhoods. These areas are labeled “infrastructure deserts,” analogous to “food deserts,” which are defined as low-income neighborhoods with insufficient access to healthy food sources (Cummins & Macintyre, 2002; Ploeg et al., 2009; Wrigley, 2002); and “transit deserts,” which are transit-dependent areas that lack adequate public transit service (Aman & Smith-Colin, 2020; Jiao & Dillivan, 2013). The introduction of “infrastructure deserts” presents a more comprehensive and integrated perspective of neighborhood

weakness in physical assets and community services. We agree that including the effects/needs of residents can provide a more comprehensive view of some types of infrastructure. Determining deficiency based on differences between need and supply (Jiao, 2017) is widely used and accepted in the transportation arena, e.g. However, this approach is difficult to implement for neighborhood infrastructure due to the diversity of “infrastructure types”. It is challenging to properly define and measure the demand and need when multiple infrastructure types are considered. In terms of number of residents, our framework uses the number of residential households to approximate the size and the need of residents within the neighborhood when assessing infrastructure types related to facility access, for other types of infrastructure (e.g., pavement condition), assessment is made by street segment, which is more relevant than number of residents. For example, we compute the number of residential households and compare it with the total number of households within the same neighborhood to determine whether infrastructure access is sufficient. The category of *Highly deficient* infrastructure condition from the previous step is chosen as the quantitative representation of neighborhoods as significantly more deficient in infrastructure presence and condition. Such areas that are also low income are identified as infrastructure deserts.

To define neighborhood income category, neighborhoods are classified into three groups (*low*, *middle*, and *high*) using tertiles of annual median household income (Chau et al., 2013; Franco et al., 2008). A few studies use annual median family income as an income variable instead of household income (Franco et al., 2008). In this study, annual median household income since 2013 has more available historical data than annual family income and has been used to interpolate missing income data for some neighborhoods (Bureau, 2020) Our analyses indicate these choices do not significantly bias the resulting spatial patterns.

## 2.3. Explore infrastructure inequity with statistical models

Lastly, to account for any uncertainty within the observed data, statistical models can serve to further explain the relationship between neighborhood infrastructure condition and income, as well as calculate the significance of infrastructure inequity. Since the overall infrastructure deficiency is computed as an ordinal integer according to Eq. (1-3), the cumulative logit model (also called proportional odds model) (Agresti, 2003) is appropriate for this case as it was designed for a response variable that takes values in a set of ordered categories (multiple ordinal responses). This model was initially proposed by Walker and Duncan (Walker & Duncan, 1967) as an extension of the logistic regression model for binary responses.

In this study, the model relates a response variable  $Y$ , consisting of ordered categories (e.g., overall infrastructure deficiency), to a categorical explanatory variable (e.g., neighborhood income characteristics) with  $k + 1$  levels and represented by  $\mathbf{x}$ , a vector of  $k$  dummy variables that represent  $k$  different levels (the remaining level is chosen as the reference level). The model has the following generalized representation:

$$\text{logit}[\text{Pr}(Y \leq j | \mathbf{x})] = \alpha_j + \beta^T \mathbf{x}, j = 1, \dots, J - 1 \quad (2-1)$$

where  $\text{Pr}(Y \leq j | \mathbf{x})$  is the cumulative probability of the event ( $Y \leq j$ ),  $\alpha_j$  are the unknown intercept parameters, and  $\beta^T = (\beta_1, \beta_2, \dots, \beta_k)$  is a vector of regression coefficients used for all response categories.  $J$  is the total number of response categories and *logit*, also known as the log-odds transformation, is the inverse function for the standard logistic cumulative distribution function:

$$\text{logit}(t) = \log \frac{t}{1-t} \quad (2-2)$$

The model assumes the same effects  $\beta$  for each logit. Thus the regression coefficient vector,  $\beta$ , does not depend on  $j$ , implying that the log-odds ratio is proportional to the difference between two  $\mathbf{x}$  values and



shares the same proportionality constant regardless of  $j$ . This is also called the proportional odds assumption. The validity of this assumption can be checked based on a  $\chi^2$  score test (Peterson & Harrell, 1988).

Applying this model with overall infrastructure deficiency as a response variable and income level as a single explanatory variable results in:

$$\text{logit}[Pr(\gamma \leq j|x)] = \alpha_j + \beta_M x_M + \beta_H x_H, j = 1, 2, \dots, J-1 \quad (2-3)$$

where  $\gamma$  is the computed overall infrastructure deficiency with each value of integer representing one category;  $x_M, x_H$  are two dummy variables:  $x_M = 1$  if the income level is *middle*,  $x_M = 0$ ;  $x_H = 1$  if income level is *high*, and  $x_H = 0$ ;  $x_M = x_H = 0$  if the income level is *low*, serving as the reference level.  $J$  is the max number of deficient infrastructure types observed ( $J = 11$  in this case study) and  $\beta_M, \beta_H$  are regression coefficients for the dummy variables of the categorical covariate with three levels (*low, middle, high*).

The fitted model is validated by performing a Likelihood Ratio Test (LRT) (Agresti, 2003) between the fitted model and the same model using a multinomial link. With the null hypothesis that proportional odds assumption holds, a  $p$ -value of  $>0.05$  indicates that the data do not show gross violation of the assumption. A relative risk measure of deficient infrastructure types ( $RR_{xj}$ ) is then computed between different income levels to draw statistical conclusions. In particular, the relative risk of low-income neighborhoods of having “more deficient ( $>j$ )” infrastructure types compared to neighborhoods with income-level denoted as  $x$  is written as:

$$RR_{xj} = \frac{Pr(\gamma > j|low\ income)}{Pr(\gamma > j|x)} \\ = \frac{1 - Pr(\gamma \leq j|low\ income)}{1 - Pr(\gamma \leq j|x)} = \frac{1 + e^{\alpha_j + \beta_x}}{1 + e^{\alpha_j}} \\ j = 1, \dots, J-1, x = \{M, H\} \quad (2-4)$$

Relative risk offers adequate measures to compare overall infrastructure condition across different neighborhood income levels. Given neighborhood income  $x$  and the number of deficient infrastructure types  $j$ , if the relative risk value ( $RR_{xj}$ ) is larger than one; then low-income neighborhoods show a higher risk of having  $j$  or more deficient infrastructure types than neighborhoods with income level  $x$ , also showing evidence of infrastructure inequity.

To obtain the confidence intervals for relative risk at each  $j$ , a bootstrapping method (James et al., 2013) is used with 20,000 iterations to compute the upper (97.5 %) and lower (2.5 %) confidence level of the relative risk estimates. All of the statistical computations described herein are executed with the statistical software R. The cumulative logit model is fit using the function *polr* from package MASS (Ripley et al., 2013). All coefficients were exported and visualized using Python.

## 2.4. Case study – Dallas, Texas

Dallas, TX, USA, was selected as the case study for this first infrastructure assessment framework. It is a pertinent choice because of its ongoing strong economic development, significant infrastructure problems and issues, and plans for redevelopment activities in the future to address infrastructure issues, as well as its history of racial and wealth segregation by neighborhood. Dallas is one of the Rockefeller Foundation's 100 Resilient Cities; its resilience strategy was released in 2018 and includes equity and neighborhood infrastructure investment as core goals. Dallas has the highest level of income inequality in the United States (U.S.) (Florida, 2017; Taylor & Fry, 2012) and one of the highest rates of increase in urban heat among major US cities (Habeeb et al., 2015). Furthermore, Dallas County has the 4th highest number of pedestrian fatalities among U.S. counties in 2016 (Retting et al., 2017). The city also rated significantly lower than the national average in street

and infrastructure maintenance, according to a community survey in 2018. These statistics highlight the existing neighborhood infrastructure issues and make Dallas an ideal location to study neighborhood-scale infrastructure equity.

To assess neighborhood infrastructure in Dallas, a total of 12 neighborhood infrastructure types with available and newly derived data were considered (pavement, sidewalk, crosswalk, noise wall, public transportation, trails, medical facilities, food stores, community gathering places, bank service, street tree canopy, and internet service). Measurable data for each infrastructure type were identified based on multiple types of data (tabular data, spatial lines, or spatial points) and related references and infrastructure design guidelines as shown in Table 1. The table describes measurable components for each infrastructure type and corresponding evaluating criteria. For noise walls, only households near highways can be affected by the presence or absence of noise walls. Hence, the evaluation is restricted only to residential households within 200 ft of major highways (Rochat & Reiter, 2016).

Lastly, annual median household income of Census block groups in the Dallas region were obtained from the 2018 U.S. Census table – B19013. For block groups with missing income records, the average between historical information at the block group level (linear regression using the past five years' income records, as available) and current-year income at the tract level is used to fill in missing data. This method offers a better estimation for missing income data because it accounts for currency inflation over the years and impacts of nearby neighborhoods within the same Census tract. After filling in missing income records, a total of 790 neighborhoods across Dallas had complete income and infrastructure condition data. The neighborhood income was then categorized as *low-income* (347 neighborhoods), *middle-income* (205 neighborhoods), and *high-income* (238 neighborhoods) using tertiles across whole Dallas county. The cutoffs between income levels were \$44,100 for the 33rd percentile and \$70,200 for the 66th percentile.

## 3. Results

### 3.1. Individual and overall infrastructure condition

Applying the framework (described in detail below) to the Dallas dataset required the following steps. First, individual infrastructure types were rated within each neighborhood using the substandard infrastructure criteria (Table 1). Second, overall neighborhood infrastructure deficiency ( $\gamma$ ) was then calculated by summing the binary deficiency indicators across all 12 infrastructure types. Categories of infrastructure deficiency were then allocated to each neighborhood based on percentiles to the nearest integer value of  $\gamma$  could have: *Excellent* ( $\gamma \leq 3$ ) (0 % ~ 10 %), *Good* ( $\gamma = 4$ ) (10 % ~ 25 %), *Moderate* ( $5 \leq \gamma \leq 6$ ) (25 % ~ 75 %), *Deficient* ( $\gamma = 7$ ) (75 % ~ 90 %), and *Highly deficient* ( $\gamma \geq 8$ ) (90 % ~ 100 %). Following the definition of infrastructure deserts, low-income neighborhoods with *Highly deficient* infrastructure ( $\gamma \geq 8$ ) were then identified across the city. Finally, a cumulative logit model was fit between overall infrastructure deficiency and income level to compute the relative risk of *low-income* neighborhoods versus wealthier neighborhoods.

Fig. 2 shows the percentage of neighborhoods with deficiencies for each individual type of infrastructure by income level. This distribution of deficient infrastructure exhibits three different patterns by infrastructure type: 1) For crosswalks, internet service, medical facilities, noise walls, and food access, the share of neighborhoods with deficient infrastructure is much higher in low-income neighborhoods than others, showing a decreasing trend with increasing income; 2) For pavement, sidewalks, community gathering spaces, trail access, and street tree canopy, the share of deficient infrastructure does not show much difference across the three income groups; 3) For public transit, an increasing trend exists with deficient infrastructure versus income level.

Some of the results are consistent with previous findings, which show

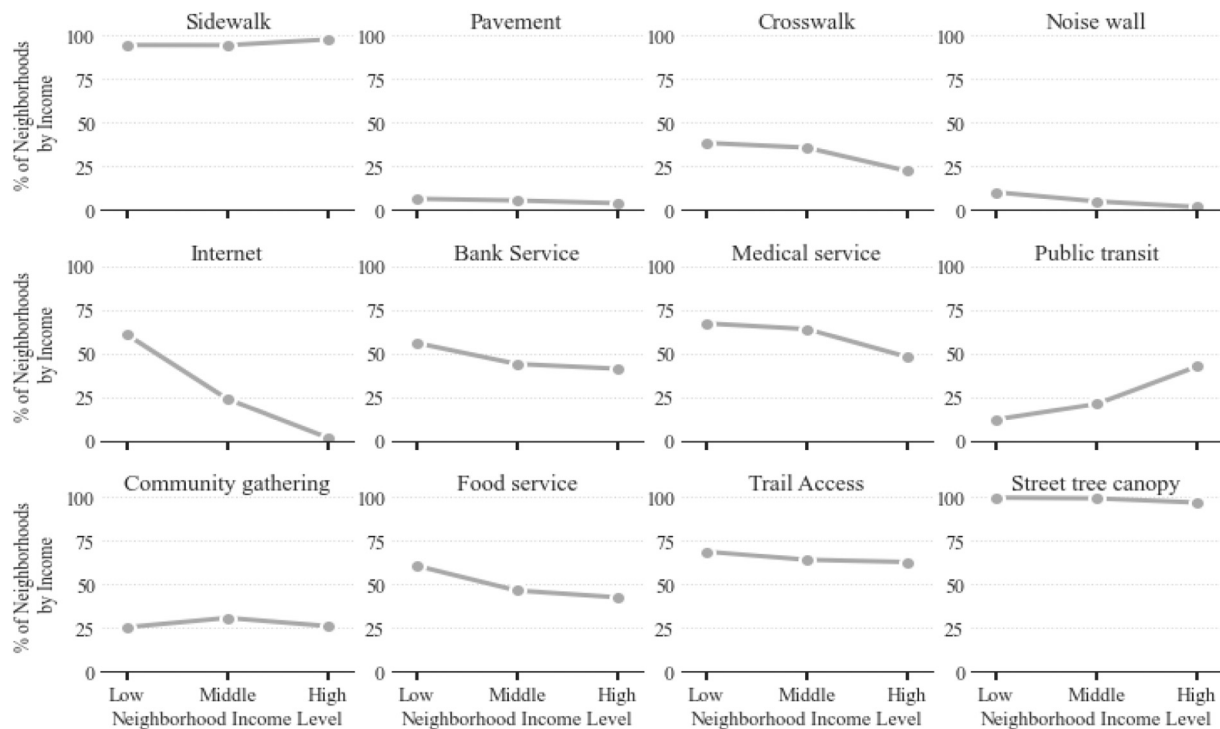


Fig. 2. Percentage of deficient infrastructure by neighborhood income level and infrastructure type.

inequities across community's socio-economic status for individual infrastructure, such as crosswalks (Bereitschaft, 2017; Su et al., 2019), internet service (Perzynski et al., 2017), and food service (Chau et al., 2013). However, high-income neighborhoods experience more deficiency than low-income neighborhoods for public transit and, to some extent, sidewalks. This finding is not consistent with the literature (Borowski et al., 2018) and may be due to higher percentage of vehicle ownership in high-income neighborhoods.

These types of mixed relationships between infrastructure types and neighborhoods' socio-economic status introduce challenges to studying infrastructure equity by individual infrastructure type. This illustrates the need to consider multiple infrastructure types simultaneously and to develop a multi-infrastructure framework with an overall infrastructure deficiency metric.

Fig. 3 (a) shows a histogram of overall infrastructure deficiency as a percentage of Census block groups, the spatial unit defining

neighborhoods in this study. Infrastructure deficiency categories are also represented by color, with dashed lines showing the boundaries between each category. Fig. 3 (b) shows the distributions of overall infrastructure deficiency by income level; the y-axis represents the percentage of neighborhoods with the same income level. The results show that the overall infrastructure deficiency ranges from 1 to 11, meaning that all neighborhoods have at least one deficient infrastructure type and none are deficient in all infrastructure types (12 types in total). The majority of neighborhoods have between 4 and 7 deficient infrastructure types.

Overall, 14 % of the neighborhoods are classified as *Excellent* for their overall infrastructure condition, while 13 % of neighborhoods are *Highly deficient*. Fig. 3 (b) shows that middle-income neighborhoods have a similar distribution to high-income neighborhoods, except that there are more high-income neighborhoods with very few infrastructure deficits (*Excellent*). However, low-income neighborhoods clearly exhibit

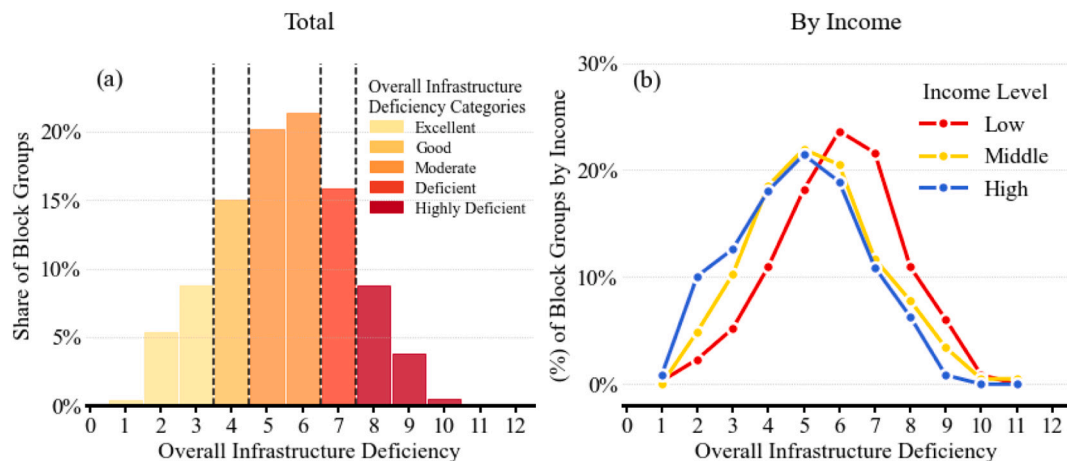


Fig. 3. Overall infrastructure deficiency. (a) Histogram of overall infrastructure deficiency as a percentage of block groups, (b) Histogram of overall infrastructure deficiency as a percentage of block groups by income level.

higher overall infrastructure deficiency than other neighborhoods, as the distribution is horizontally shifted toward *Highly deficient* (8 or more deficient infrastructure types). This pattern reveals evidence of inequitable infrastructure provision between low-income neighborhoods and others.

### 3.2. Infrastructure deserts

Fig. 4 shows the map of infrastructure deserts (low-income neighborhoods with *Highly deficient* infrastructure ( $\gamma \geq 8$ )) in Dallas. A total of 62 neighborhoods were identified as infrastructure deserts. The infrastructure deserts also show a clear spatial pattern where more low-income neighborhoods in the south (24 %, or as 54 out of 221 low-income neighborhoods) are identified as deserts than low-income neighborhoods in the north (6 %, or as 8 out of 126 low-income neighborhoods).

As further comparison of infrastructure deserts versus other areas, Fig. 5 shows individual deficient infrastructure types as a percentage of neighborhoods citywide versus within infrastructure deserts. It suggests that more than half of the neighborhoods citywide have inadequate street tree canopy, sidewalk, noise wall, trail access, medical facility access, and food access. However, infrastructure deserts have substantially more of these deficiencies. Furthermore, more than half of infrastructure deserts have deficient crosswalks and access to bank services, internet services, and gathering places. Overall, street tree canopy and sidewalks are the most widespread deficient infrastructure types.

### 3.3. Relative risk and infrastructure inequity

The estimated parameters for the fitted cumulative logit model (Eq. (2-3)) are shown in Table 2. The positive coefficients for ( $\beta_M$ ,  $\beta_H$ ) indicate a tendency for overall infrastructure deficiency to become less deficient for middle-income and high-income neighborhoods compared to low-income neighborhoods. The estimated coefficient for the middle-income neighborhoods ( $\beta_M$ ) is 0.714, and the estimated coefficient for high-income neighborhoods ( $\beta_H$ ) is 1.124. Both coefficients show that middle-income and high-income neighborhoods have the tendency of less deficient infrastructure types than low-income neighborhoods.

Meanwhile, values show the tendency of overall infrastructure deficiency toward less deficient to be stronger for high-income neighborhoods than middle-income neighborhoods in comparison to low-income neighborhoods. To test the model assumption of proportional odds with these parameters, a likelihood ratio test (16 degrees of freedom) was performed between the fitted model and the same model with a multinomial link. With the null hypothesis that proportional odds assumption holds, a  $p$ -value of 0.678 was computed, which indicates that the data do not show gross violation of the assumption.

Fig. 6 shows the resulting relative risks (Eq. (2-4)) between: (A low-income and high-income neighborhoods and (B) low-income and middle-income neighborhoods. The  $x$ -axis denotes the overall infrastructure deficiency, with the value being equal or greater than that of the displayed label. The  $y$ -axis represents the value of relative risk estimates, with mean results plotted as lines and 95 % confidence levels denoted by the shaded regions. As indicated in Fig. 6, the positive values of relative risk for both comparisons suggest that low-income neighborhoods show a greater risk of having more deficient infrastructure than middle and high-income neighborhoods. Furthermore, the relative risk mean and confidence intervals increase for both comparisons as overall infrastructure deficiency increases. Low-income neighborhoods are 2.0 to 3.5 times more likely to have *Highly deficient* infrastructure ( $\gamma \geq 8$ ) than high-income neighborhoods; and 1.4 to 2.4 times more likely to have *Highly deficient* infrastructure ( $\gamma \geq 8$ ) than middle-income neighborhoods. Such substantial differences suggest significant infrastructure inequities across income levels.

### 3.4. Robustness of statistical model

To further confirm the association between neighborhood income and overall infrastructure deficiency, the model was refit using continuous (log) income instead of categorical income levels (Table 2). The use of log income helps linearize the exponentially growing trends and reduce bias compared to linear income (Ermini & Hendry, 2008). Table 3 shows the estimates of model parameters. As the log income increases, the positive estimated coefficient shows that the overall infrastructure deficiency has a trend to be “less” deficient, which corroborates the previous findings using categorical income data. Fig. 7

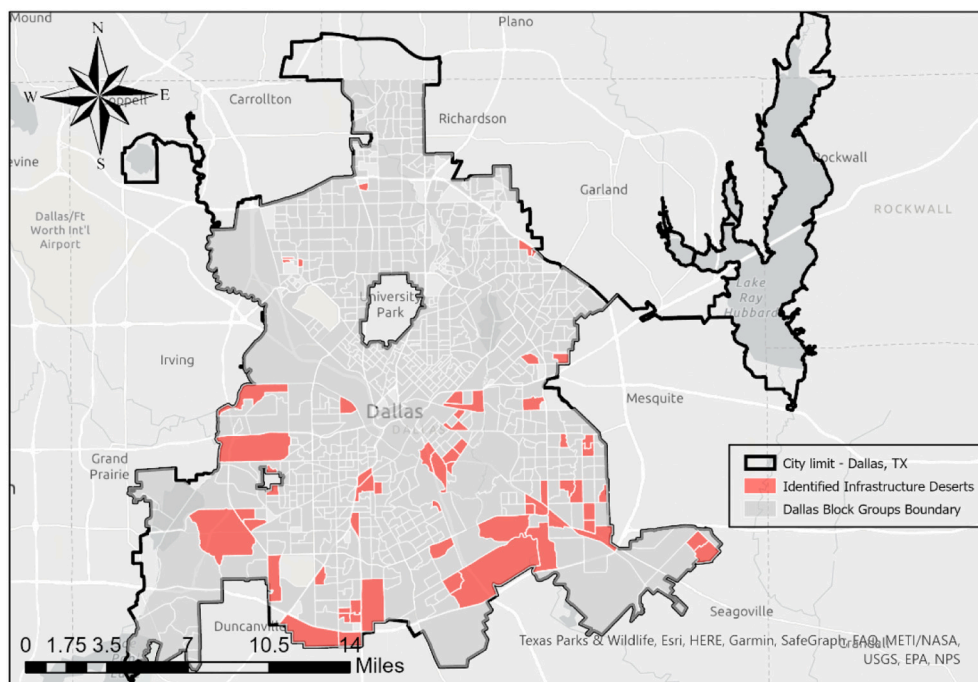


Fig. 4. Identified infrastructure deserts in Dallas, Texas.

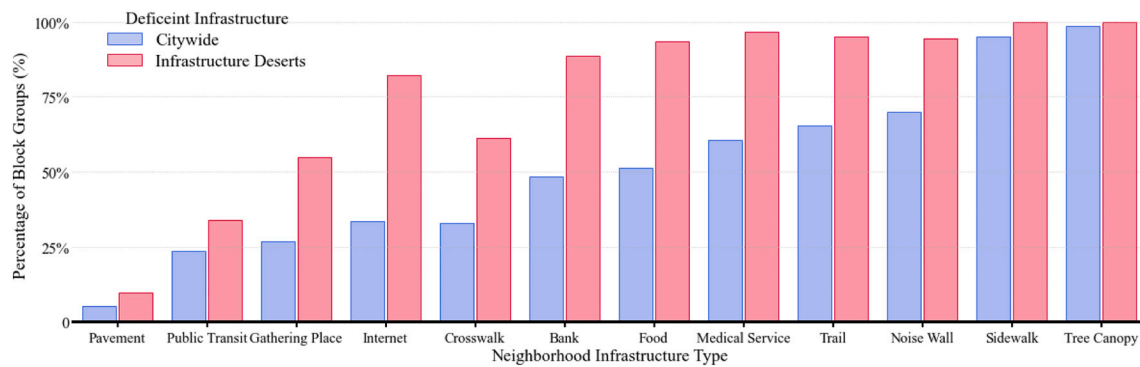


Fig. 5. Deficient infrastructure as a percentage of block groups by infrastructure type.

Table 2

Estimated coefficients of the cumulative logit model. We validated the model assumption (proportional odds) by performing a likelihood ratio test (16 degrees of freedom) between the fitted model and the same model except using a multinomial link. With the null hypothesis that proportional odds assumption holds, the p-value of 0.678 indicates that the data do not show gross violation of the assumption.

|                           |               | Value    | Std. Error | t value  |
|---------------------------|---------------|----------|------------|----------|
| Coefficients              | $\beta_M$     | 0.714    | 0.157      | 4.558    |
|                           | $\beta_H$     | 1.124    | 0.153      | 7.364    |
| Intercepts ( $\alpha_j$ ) | $\alpha_1$    | -6.196   | 0.582      | -10.641  |
|                           | $\alpha_2$    | -3.430   | 0.178      | -19.235  |
|                           | $\alpha_3$    | -2.380   | 0.134      | -17.796  |
|                           | $\alpha_4$    | -1.436   | 0.113      | -12.713  |
|                           | $\alpha_5$    | -0.534   | 0.103      | -5.209   |
|                           | $\alpha_6$    | 0.427    | 0.102      | 4.197    |
|                           | $\alpha_7$    | 1.450    | 0.121      | 12.022   |
|                           | $\alpha_8$    | 2.651    | 0.182      | 14.585   |
|                           | $\alpha_9$    | 4.645    | 0.452      | 10.283   |
|                           | $\alpha_{10}$ | 6.261    | 1.002      | 6.248    |
| Residual Deviance         |               | 3088.505 | AIC        | 3112.505 |

shows the predicted probability of overall infrastructure deficiency by different income percentiles (5th, 25th, 50th, 75th, 95th). Note that the probability curve shifts to the direction of “more” deficient with decreased neighborhood income, again showing a tendency to have more deficient infrastructure types for lower-income neighborhoods. This trend agrees with the earlier findings that lower-income neighborhoods have a significantly higher risk of greater infrastructure deficiency than other neighborhoods and shows the model’s robustness using either continuous or categorical income data.

#### 4. Discussion and conclusions

Given a wide variety of physical attributes within a neighborhood and their inter-dependent interactions, assessing neighborhood infrastructure condition can be highly challenging. The task involves integrating a substantial set of neighborhood infrastructure condition indicators that are multidimensional and heavily data dependent. To our knowledge, there are a lack of approaches or frameworks in the existing neighborhood infrastructure-related literature that consider the diversity of neighborhood infrastructure and study multiple types of infrastructure combined. This paper contributes a novel approach to assessing neighborhood infrastructure condition by systematically measuring multiple infrastructure types and statistically analyzing infrastructure equity across neighborhood income characteristics.

Table 3

Cumulative logit model parameters using continuous income.

|                           |                        | Value    | Std. Error | t value  |
|---------------------------|------------------------|----------|------------|----------|
| Coefficients              | Log Income ( $\beta$ ) | 0.670    | 0.105      | 6.404    |
|                           |                        |          |            |          |
| Intercepts ( $\alpha_j$ ) | $\alpha_1$             | -12.941  | 1.292      | -10.014  |
|                           | $\alpha_2$             | -10.163  | 1.163      | -8.741   |
|                           | $\alpha_3$             | -9.117   | 1.153      | -7.906   |
|                           | $\alpha_4$             | -8.182   | 1.146      | -7.141   |
|                           | $\alpha_5$             | -7.294   | 1.139      | -6.406   |
|                           | $\alpha_6$             | -6.349   | 1.132      | -5.609   |
|                           | $\alpha_7$             | -5.340   | 1.130      | -4.725   |
|                           | $\alpha_8$             | -4.145   | 1.136      | -3.648   |
|                           | $\alpha_9$             | -0.215   | 1.207      | -1.780   |
|                           | $\alpha_{10}$          | -0.532   | 1.502      | -0.354   |
| Residual Deviance         |                        | 3104.692 | AIC        | 3126.692 |

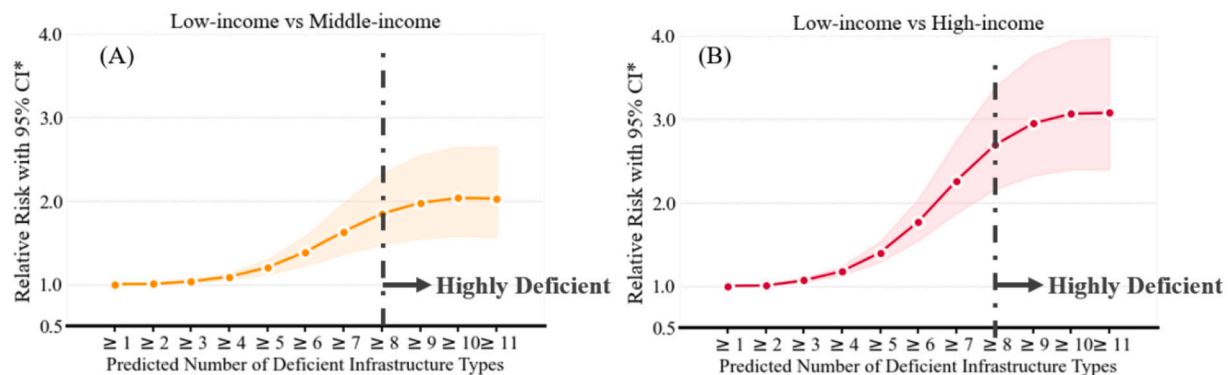


Fig. 6. Relative risk: Computed relative risk is shown as circles, and shaded regions denote its upper (97.5 %) and lower (2.5 %) confidence limits. 95 % confidence interval of both cases were obtained using bootstrapping after 20,000 simulations. (a) The relative risk of overall infrastructure deficiency between low-income and high-income areas, (b) Relative risk of overall infrastructure deficiency between low-income and middle-income areas.



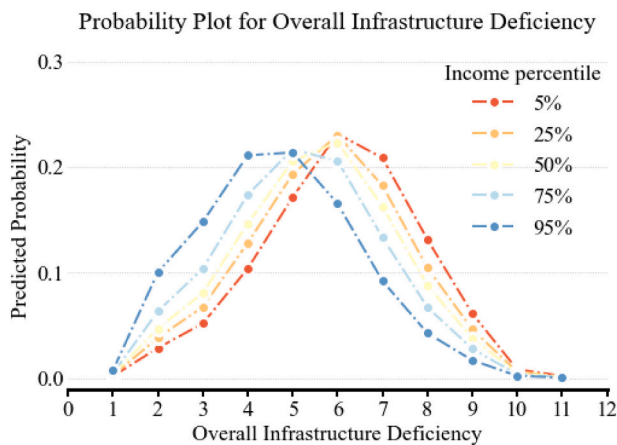


Fig. 7. Predicted probability of overall infrastructure deficiency using continuous income.

A critical strength of this study is the systematic and street-level assessment of multiple neighborhood infrastructure types. The introduction of binary infrastructure indicators and overall infrastructure deficiency effectively integrates multiple infrastructure types and provides a straightforward and intuitive neighborhood-level representation of infrastructure issues. The calculation of overall infrastructure deficiency allows the framework to easily be expanded to other types of infrastructure data (e.g., housing condition).

Unlike previous studies that are focus on single infrastructure types, our assessment results combine multiple infrastructure types and deliver more comprehensive insights into the community's overall infrastructure condition, which provides strategic benefits to stakeholders to plan and ensure equitable development among communities. Furthermore, this study introduces the new concept of “infrastructure deserts” – low-income areas with substantially higher infrastructure deficiency – and through the case study in Dallas, TX shows the presence of infrastructure deserts and infrastructure inequity throughout multiple low-income areas.

The statistical analyses also show that the observed infrastructure inequities between low-income and higher-income neighborhoods are statistically significant. To address these types of infrastructure inequities, long-term investments are needed to improve infrastructure in low-income areas. Investment prioritization based on asset condition and economic impacts (Quadros & Nassi, 2015) is one popular approach for infrastructure management that could be used to foster healthier and more equitable communities. The existence of infrastructure inequity not only identifies vulnerable neighborhoods with inadequate infrastructure resources, but also draws attention to highly deficient areas citywide to optimize investments.

Bond programs, separate from the city's annual operating budget, focus on improving capital funding for the City of Dallas assets including neighborhood infrastructure such as facilities, streets, libraries, and parks (City of Dallas, 2017 [most recent Bond program]). To address the deficiencies identified in this study, significant investments will need to be made in neighborhood infrastructure in future Bond programs. Our results also have other policy implications for infrastructure investments. A notable finding from this study is that sidewalk and street tree canopy deficits are more widespread across the city than any other neighborhood infrastructure type. According to Dallas' “50–50 sidewalk replacement program” (Dallas City Hall, 2016), homeowners are responsible for keeping sidewalks in safe condition and 50 % of the reconstruction cost is shared by the city for homeowners to fix deteriorating sidewalks. The widespread presence of deficient sidewalks suggests that further actions are needed to improve overall sidewalk conditions, such as fully-funded sidewalk replacements in future bond programs.

In addition, the extensive insufficient street tree canopy suggests the need to plant significantly more trees, which have been shown to be effective for reducing urban heat (Tan et al., 2016), particularly given increasing urban warming under climate change (Aniello et al., 1995). Several studies such as the Urban Forest Management Plan (Texas Trees Foundation, 2019) and Urban Heat Management Study (Texas Tree Foundation, 2017) have developed plans for Dallas to increase tree cover and reduce urban heat, but major investments are needed to implement these plans. In particular, street trees are needed to provide shade across paved areas, which reflect the most heat, as well as to shade residents as they walk in the neighborhood (e.g., while accessing public transit or parks).

With careful planning, these types of infrastructure investments could simultaneously improve multiple types of infrastructure (e.g., street trees that are coupled with sidewalks and new community gathering spaces, such as “complete streets” initiatives (LaPlante & McCann, 2008)). Furthermore, implementers of the recently signed Infrastructure Investment and Jobs Act (IIJA), namely the “infrastructure bill” (Brian Naylor & Deirdre Walsh, 2021; Emily Cochrane, 2021), can also benefit from our results. Under IIJA, state governments are required to find ways to identify disinvested communities relevant to each provision's implementation (Carlos Martin et al., 2021; Diana Lonescu, 2022). The discovery of infrastructure inequity and overall infrastructure measures in our study explicitly address these gaps and would be beneficial to IIJA's investment and project prioritization.

#### 4.1. Limitations and future research

This study has several limitations. First, the spatial representation of neighborhoods is challenging and has been addressed in multiple ways (Flowerdew et al., 2008; Rollings et al., 2015). Despite the widespread use of Census tracts or block groups, there are no definitive studies identifying the best spatial boundary to be used among all available options such as Census tracts, block groups, and zip codes (Flowerdew et al., 2008). Past studies have shown that the types of geographic boundaries used to aggregate data can affect variance, standard deviations, correlations, and regression analyses (Flowerdew et al., 2008). A better approach is to use perceived, resident-defined neighborhood boundaries, which may better represent the neighborhood and neighborhood-based measures such as access to destinations, walking routes, or the number of residences. For example, Nextdoor, a hyperlocal social network service for neighborhoods, offers a more reliable and accurate neighborhood geometry using a crowd-sourcing mechanism that allows users to sketch or modify the neighborhood in which they currently live (Abhyanker, 2014). However, despite better geographic representation, resident-defined boundaries can be affected by neighborhood reputation and can introduce bias in neighborhood-based studies. For example, residents might report living in positively perceived neighborhoods but exclude stigmatized areas (Flowerdew et al., 2008). Besides, resident-defined neighborhood boundaries do not have spatial compatibility with socio-economic measures embedded in administrative boundaries. Since this paper aims to examine multiple infrastructure types and explore spatial patterns of infrastructure conditions at the city level, choosing a Census block group as the representation of neighborhoods is currently the best available approach to include socio-economic characteristics at the smallest scale possible. However, this approach may introduce some errors and bias.

A second limitation is uncertainties in the criteria for measuring substandard infrastructure components. Every criterion was designed to be consistent with either prior studies, or published design guidelines. However, access measures developed using spatial measures and Geographic Information System (GIS) procedures may fail to account for the actual quality of and access to infrastructure (e.g., healthcare facilities) (Rollings et al., 2015; Wells & Rollings, 2012). For example, residents may access facilities that are not necessarily nearby their neighborhoods, potentially due to social networks, transportation

availability, or perceptions of crime and safety (Wells & Rollings, 2012). Hence relying only on proximity without considering social aspects of neighborhoods can introduce errors or biases on infrastructure availability. Another issue is the use of cutoff distances to measure infrastructure accessibility. Although cut-offs used in this study come from relevant studies, those cut-offs admittedly are not guaranteed to be optimal in the cases of any other cities. Recently, increase in GIS implementation has led to improvements in measuring the accessibility of activity locations (Aman & Smith-Colin, 2020; Luo & Wang, 2003; Nicholls, 2001). The gravity model-based method (Weibull, 1976) calculates accessibility based on zones as a function of activity opportunity attractiveness and the travel distance between other zones and the individual's resident zones. It is one of the most popular methods to measure accessibility because of the ease of interpretation and robustness of model extensions (El-Geneidy & Levinson, 2006).

Nonetheless, fixed distance approaches, such as those implemented in this study, remain favorable in many infrastructure-related studies due to their simple intuition and easy implementation. However, the choice of "proper" distance is mostly empirical and lacks theoretical justification. For instance, the critical distance used in assessing healthcare services is 2-mile (3.2 km) for major hospitals and 1-mile (1.6 km) for walk-in clinics and urgent care (Ashwood et al., 2011; Bersamin et al., 2011). However, many factors could affect people's accessibility to these destinations, such as travel behaviors, transportation mode, and city development, resulting in different values of suggested critical distances for accessibility assessment (Neutens, 2015). Despite these inevitable uncertainties, the criteria chosen for this case study are sufficient for a comparative assessment of infrastructure equity across multiple infrastructure types.

Finally, a full and complete assessment of neighborhood infrastructure should involve six main categories: connective infrastructure, protective infrastructure, socio-economic structures, water and sanitation lifelines, energy lifelines, and communication lifelines (Reduction & Recovery(GFDRR), 2017). In this study, 12 infrastructure types were considered that included four of the six categories, excluding energy lifelines and water sanitation lifelines. With additional data availability, more infrastructure types such as stormwater drains, water supply and wastewater pipes, street lights, and energy reliability will undoubtedly add to the story of complex, interdependent dynamics among neighborhood infrastructure.

Similarly, when considering multiple infrastructure types, the weighting scheme (currently equally weighted) and deficient infrastructure cutoff (currently 50 %) could be adjusted to better describe neighborhoods' demand and the focus of a city's development plan. The determination of weights and cutoffs may reflect the city's current infrastructure condition and emphasis on infrastructure. We chose to equally weight each infrastructure type to avoid any bias across infrastructure types. Besides, the overall infrastructure deficiency has a more intuitive and straightforward interpretation if each type is equally weighted. Our equal-weight assumption not only simplifies the nuance among multiple infrastructure types but also keeps the generality of the approach to be consistent and interpretable. However, it is recommended to explore the sensitivity of outcomes to these assumptions in future research.

The framework proposed in this study can easily be expanded to include other infrastructure types as data are available and fine-tuned to match the characteristics of other regions, providing the capacity and flexibility to measure conditions of a wide range of infrastructure types

systematically.

Despite the limitations noted above, this study takes the first step to consider neighborhood infrastructure as a diverse, multi-type system and assesses infrastructure conditions with data-driven approaches. Our findings have important policy implications and lessons for cities and developers that are promoting equitable infrastructure. Much progress has been made on this front in Dallas, with the Dallas Sidewalk Replacement Program (Dallas City Hall, 2016), Urban Forest Management Plan (Texas Trees Foundation, 2019), and other initiatives to improve neighborhood infrastructure and equity.

However, as the findings of this study suggest, infrastructure inequities persist across income lines and planners and policymakers should address these issues to close the "infrastructure gap." In addition to prioritized investments in disadvantaged neighborhoods, community engagement is also needed to better understand the impact of the lack of specific infrastructure types on residents and develop smart and effective strategies such as incorporating the effects/needs of residents based on demand and supply (Jiao, 2017) for promoting priorities in neighborhood infrastructure development that better meets neighborhood needs.

#### CRediT authorship contribution statement

**Zheng Li:** Conceptualization, Methodology, Investigation, Visualization, Writing – original draft. **Xinlei Wang:** Methodology, Writing – review & editing. **Jessie Zarazaga:** Conceptualization, Methodology, Investigation, Visualization, Writing – review & editing. **Janille Smith-Colin:** Conceptualization, Methodology, Writing – review & editing. **Barbara Minsker:** Conceptualization, Methodology, Supervision, Writing – review & editing.

#### Declaration of competing interest

The authors declare that they have no competing interests.

#### Acknowledgments

The authors wish to thank faculties, graduate and undergraduate students who provide their assistance in the data collection: Dr. Bardia Heidari, Dr. Dinesh Rajan, Dr. Doric Earle, Dr. Eric Larson, Dr. Jo Guldi, Dr. Kenneth Berry, Dr. May Yuan, Dr. Michael Hites, Dr. Tammy Leonard, Arefeh Safaei, Azucena Milan, Elizabeth Killingsworth, Farbod Tavakoli, Gurveer Kuar, Javad Jomehpour, Jennifer Ebinger, Jerry Hawkins, Jessica MacKinnon, Luigi Marini, Melanie Ferguson, Mohamed Elsaied, Myriam Zakhem, Owen Wilson-Chavez, Raghuram Srinivas, Sylvia George-Williams.

#### Funding

This material is based upon work supported by the United States National Science Foundation under Grant No. 1835877.

#### Data and materials availability

All pseudocode of the method used to conduct infrastructure assessment, a table of descriptive statistic for assessed individual infrastructure type, and dataset sources are available in the Appendix.

## Appendix A

A.1. Table. Dataset information for individual infrastructure type.

|                              | Dataset Source                                                                                                    | Data Year | Notes    |
|------------------------------|-------------------------------------------------------------------------------------------------------------------|-----------|----------|
| Pavement                     | City of Dallas ArcGIS REST Service <sup>1</sup>                                                                   | 2018      | Polyline |
| Crosswalk                    | Object detection using Google Satellite images on residential intersections                                       | 2019      | Point    |
| Noise Wall                   | Annotated dataset using Google StreetView images along major state highways                                       | 2019      | Point    |
| Internet Service             | Federal Communication Commission <sup>2</sup> Broadband width map at tract level                                  | 2016      | Point    |
| Bank Access                  | Bank branches locations from NCTCOG <sup>3</sup> data center                                                      | 2019      | Point    |
| Medical Facility Access      | Major hospitals are from NCTCOG data center and urgent care or clinics are from Yelp <sup>4</sup> search listings | 2018      | Point    |
| Public Transportation Access | Bus stops, rail stations locations from City of Dallas ArcGIS REST Service                                        | 2018      | Point    |
| Gathering Place Access       | Public parks, libraries, farmer markets and community centers extracted from NCTCOG data center, tax parcel data  | 2019      | Point    |
| Food Access                  | Food stores (grocery stores, wholesale) locations from NCTCOG data center                                         | 2019      | Point    |
| Trail Access                 | Bike & pedestrian trails from City of Dallas ArcGIS REST Service                                                  | 2019      | Polyline |
| Street Tree Canopy           | Tree coverage from Texas Tree Foundation <sup>5</sup>                                                             | 2018      | Polygon  |
| Sidewalk                     | City of Dallas ArcGIS REST Service - Public Works                                                                 | 2017      | Polyline |

<sup>1</sup> City of Dallas REST Service: <https://gis.dallascityhall.com/wwwgis/rest/services/>.

<sup>2</sup> Federal Communication Commission: <https://www.fcc.gov/reports-research/maps/residential-fixed-internet-access-service-connections-per-1000-households-by-census-tract/>.

<sup>3</sup> North Central Texas Council of Governments (NCTCOG) data center: <https://data-nctcogis.opendata.arcgis.com/>.

<sup>4</sup> Yelp search listing: <https://www.yelp.com>.

<sup>5</sup> Texas Tree Foundation (TTF): <https://www.texastrees.org>

A.2. Table. Descriptive statistics for the substandard percentage ( $\mu$ ) of individual infrastructure type.

|                              | Census Block Groups (n) | Standard Deviation | Min   | 25 %  | 50 %  | 75 %  | Max   |
|------------------------------|-------------------------|--------------------|-------|-------|-------|-------|-------|
| Pavement                     | 790                     | 0.150              | 0.000 | 0.136 | 0.246 | 0.346 | 0.794 |
| Crosswalk                    | 790                     | 0.281              | 0.000 | 0.000 | 0.333 | 0.500 | 1.000 |
| Noise Wall                   | 70                      | 0.446              | 0.000 | 0.013 | 1.000 | 1.000 | 1.000 |
| Internet Service             | 790                     | 0.195              | 0.100 | 0.100 | 0.300 | 0.500 | 0.700 |
| Bank Access                  | 790                     | 0.439              | 0.000 | 0.000 | 0.458 | 1.000 | 1.000 |
| Medical Facility Access      | 790                     | 0.427              | 0.000 | 0.066 | 0.835 | 1.000 | 1.000 |
| Public Transportation Access | 790                     | 0.322              | 0.000 | 0.000 | 0.138 | 0.473 | 1.000 |
| Gathering Place Access       | 790                     | 0.349              | 0.000 | 0.000 | 0.114 | 0.552 | 1.000 |
| Food Access                  | 790                     | 0.416              | 0.000 | 0.029 | 0.529 | 1.000 | 1.000 |
| Trail Access                 | 790                     | 0.361              | 0.000 | 0.341 | 0.765 | 1.000 | 1.000 |
| Street Tree Canopy           | 790                     | 0.131              | 0.286 | 0.763 | 0.856 | 0.943 | 1.000 |
| Sidewalk                     | 786                     | 0.145              | 0.101 | 0.768 | 0.874 | 0.933 | 1.000 |

A.3. Dataset. A Shapefile consisting information of all assessed infrastructure types is included as separated file with submission.

**Filename:** Infrastructure\_assessment\_Dallas.zip

**File Format:** ArcGIS Shapefile (zipped)

**Attributes contains in the shapefile:**

BLOCKGROUP: 12 digits Census block Group ID.

Income3: Categorical income class based on tertiles: Low, Middle, High.

IncomeLog: Log value of annual household median income.

Overall\_IF: Overall infrastructure deficiency - integer.

IF\_5: Categorical overall infrastructure deficiency: Excellent, Good, Moderate, Deficient, Highly Deficient.

PCNG\_PAVE: Percentage of substandard pavement segments.

PCNG\_SDWK: Percentage of residential street segments that has substandard sidewalks.

PCNG\_CRWK: Percentage of intersections that do not have crosswalk present.

PCNG\_MEDL: Percentage of residential households that don't have access\* to medical service facilities.

PCNG\_GATH: Percentage of residential households that don't have access\* to gathering places.

PCNG\_BANK: Percentage of residential households that don't have access\* to local bank branches.

PCNG\_INTT: Percentage of residential households with substandard internet service.

PCNG\_TRIL: Percentage of residential households that don't have access\* to bicycle & pedestrian trails.

PCNG\_TRAN: Percentage of residential households that don't have access\* to bus stops nor rail stations.

PCNG\_TREE: Percentage of residential street segments with substandard tree canopy percentage (below 25 %).

PCNG\_NSWL: Percentage of residential households near highways that do not have noise wall present.

Geometry: Geometry of census block group.

\*: Based on corresponding substandard criteria table (see Table 1 in the main manuscript for more details).

A.4. Code. Pseudocode of the method used to compute 12 deficient infrastructure types

Pavement

```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: for each Neighborhood ( $i$ ) do
    Find all pavement segments  $C_i$  within/intersect with the neighborhood boundary.
    Initialize Measurable Components  $M = 0$ ; Substandard Measurable Components
     $M_{std} = 0$ .
    for each segment  $C_{ij}$  do
        Calculate segment length  $L_{ij}$ .
        Count Measurable Components in length  $M = M + L_{ij}$ .
        if segment  $C_{ij}$ 's Pavement Condition Index  $\leq 55$ 
            then  $M_{std} = M_{std} + L_{ij}$ .
        end
    Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $M_{std} / M$ 
    if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end

```

## Sidewalks

```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: for each Neighborhood ( $i$ ) do
    Find residential street segments  $C_i$  within/intersect with the neighborhood
    boundary.
    Initialize Measurable Components  $M = 0$ ; Substandard Measurable Components
     $M_{std} = 0$ .
    for each street segment  $C_{ij}$  do
        Calculate segment length  $L_{ij}$ .
        Count Measurable Components in length  $M = M + L_{ij}$ .
        if segment  $C_{ij}$  has no sidewalks on both sides
            then  $L_{missing} = L_{ij}$ 
        end
        if segment has sidewalk on at least one side
            then calculate the portion length ( $L_{std}$ ) of segment that has been obstructed
            or damaged sidewalks;
             $M_{std} = M_{std} + \max(L_{missing}, L_{std})$ 
        end
    end
    Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $M_{std} / M$ 
end

```

```

if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end

```

## Noise walls



```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: for each Neighborhood ( $i$ ) do
    Find residential households  $C_i$  located within 200 feet (61m) from major highways.
    Initialize Measurable Components  $M = 0$ ; Substandard Measurable Components
     $M_{std} = 0$ .
    for each household  $C_{ij}$  do
        Count Measurable Components  $M = M + 1$ .
        if no noise walls existed within 200 feet (61m) from  $C_{ij}$ 
            then  $M_{std} = M_{std} + 1$ 
        end
    end
    Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $M_{std} / M$ 
    if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end

```

### Crosswalks

```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: for each Neighborhood ( $i$ ) do
    Within Neighborhood boundary, find all crosswalk intersections  $C_i$  intersections
    that are either:
        1) Intersections between residential streets
        2) Intersections between school zones.
    Initialize Measurable Components  $M = 0$ ; Substandard Measurable Components
     $M_{std} = 0$ .
    for each crosswalk intersection  $C_{ij}$  do
        Create a search buffer region (34m radius)  $b_{ij}$  given its coordinates.
        Count Measurable Components  $M = M + 1$ .
        if no crosswalks existed within  $b_{ij}$ 
            then  $M_{std} = M_{std} + 1$ 
        end
    end
    Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $M_{std} / M$ 

    if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end

```

### Street Tree Canopy coverage

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: **for** each *Neighborhood* ( $i$ ) **do**

Final all street segments  $C_i$  within the neighborhood

Create street buffer polygons  $C'_i$  (use city-wide median width: 6.5 feet or 2 meter radius)

Initialize *Measurable Components*  $M = 0$ ; *Substandard Measurable Components*  $M_{std} = 0$ .

**for** each street polygon  $C'_{ij}$  **do**

Count *Measurable Components*  $M = M + 1$

Compute the area of street polygon  $A_{ij}$ .

Compute the area of the tree canopy  $A_{ij}^t$  within  $C'_{ij}$ .

Compute the street tree canopy percentage as  $p_{ij} = A_{ij}^t / A_{ij}$ .

**if**  $p_{ij} \leq 0.25$  **then**  $M_{std} = M_{std} + 1$  **end**

**end**

Calculate *Substandard Measurable Component Percentage* ( $\mu_i$ ) =  $M_{std} / M$ .

**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**

**end**

#### Pedestrian & bicycle trail access

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: Break the pedestrian & bicycle trails into points using 600-meter intervals.

Use points to create service area  $S$  for pedestrian & bicycle trails (0.8 km travel distance).

Step-3: **for** each *Neighborhood* ( $i$ ) **do**

Find all residential households  $C_i$  within the neighborhood.

Initialize *Measurable Components*  $M = 0$ ; *Substandard Measurable Components*  $M_{std} = 0$ .

**for** each residential household  $C_{ij}$  **do**

Count *Measurable Components*  $M = M + 1$ .

**if**  $C_{ij}$  is spatially outside of  $S$  **then**  $M_{std} = M_{std} + 1$  **end**

Calculate *Substandard Measurable Components Percentage* ( $\mu_i$ ) =  $M_{std} / M$

**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**

**end**

#### Medical facility access

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: Create service area  $S_1$  for major hospitals (2-mile or 3.2 km travel distance).

Create service area  $S_2$  for walk-in clinics and urgent care (1-mile or 1.6 km travel distance).

Step-3: **for** each *Neighborhood* ( $i$ ) **do**

Find all residential households  $c_i$  within the neighborhood.

Initialize *Measurable Components*  $M = 0$ ; *Substandard Measurable Components*  $M_{std} = 0$ .

**for** each residential household  $C_{ij}$  **do**

Count *Measurable Components*  $M = M + 1$ .

**if**  $C_{ij}$  is spatially not in  $S_1$  nor  $S_2$  **then**  $M_{std} = M_{std} + 1$  **end**

Calculate *Substandard Measurable Components Percentage* ( $\mu_i$ ) =  $M_{std} / M$

**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**

**end**

#### Public transportation access

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: Create service area  $S_1$  for rail stations (0.8 km travel distance).  
Create service area  $S_2$  for bus stops (0.4 km travel distance).

Step-3: **for** each *Neighborhood* ( $i$ ) **do**  
Find all residential households  $C_i$  within the neighborhood.  
Initialize *Measurable Components*  $M = 0$ ; *Substandard Measurable Components*  $M_{std} = 0$ .  
**for** each residential household  $C_{ij}$  **do**  
Count *Measurable Components*  $M = M + 1$ .  
**if**  $C_{ij}$  is spatially not in  $S_1$  nor  $S_2$  **then**  $M_{std} = M_{std} + 1$  **end**  
Calculate *Substandard Measurable Components Percentage* ( $\mu_i$ ) =  $M_{std} / M$   
**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**  
**end**

### Food access

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: Create service area  $S$  for fresh food stores (1-mile or 1.6 km travel distance).

Step-3: **for** each *Neighborhood* ( $i$ ) **do**  
Find all residential households  $C_i$  within the neighborhood.  
Initialize *Measurable Components* ( $M$ ) = 0; *Substandard Measurable Components* ( $M_{std}$ ) = 0.  
**for** each residential household  $C_{ij}$  **do**  
Count *Measurable Components*  $M = M + 1$ .  
**if**  $C_{ij}$  is spatially not in  $S$  **then**  $M_{std} = M_{std} + 1$  **end**  
Calculate *Substandard Measurable Components Percentage* ( $\mu_i$ ) =  $M_{std} / M$   
**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**  
**end**

### Bank access

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).

Step-2: Create service area  $S$  for bank branches (1-mile or 1.6 km travel distance).

Step-3: **for** each *Neighborhood* ( $i$ ) **do**  
Find all residential households  $C_i$  within the neighborhood.  
Initialize *Measurable Components*  $M = 0$ ; *Substandard Measurable Components*  $M_{std} = 0$ .  
**for** each residential household  $C_{ij}$  **do**  
Count *Measurable Components*  $M = M + 1$ .  
**if**  $C_{ij}$  is spatially not in  $S$  **then**  $M_{std} = M_{std} + 1$  **end**  
Calculate *Substandard Measurable Components Percentage* ( $\mu_i$ ) =  $M_{std} / M$   
**if**  $\mu_i \geq 0.5$  **then**  $\theta_i = 1$  **else**  $\theta_i = 0$  **end**  
**end**

### Gathering place access

```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: Create service area  $S_1$  for parks (1-mile or 1.6 km travel distance).
        Create service area  $S_2$  for libraries (1-mile or 1.6 km travel distance).
        Create service area  $S_3$  for community centers (1-mile or 1.6 km travel distance).
        Create service area  $S_4$  for farmers' markets (1-mile or 1.6 km travel distance).
Step-3: for each Neighborhood ( $i$ ) do
        Find all residential households  $C_i$  within the neighborhood.
        Initialize Measurable Components  $M = 0$ ; Substandard Measurable Components
 $M_{std} = 0$ .
        for each residential household  $C_{ij}$  do
                Count Measurable Components  $M = M + 1$ .
        if  $C_{ij}$  is spatially outside of  $S_1$  and  $S_2$  and  $S_3$  and  $S_4$  then  $M_{std} = M_{std} + 1$ 
end
        Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $M_{std} / M$ 
        if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end

```

## Internet Service

```

Step-1: Initiate  $\theta$  as an empty array with size  $N \times I$  ( $N = \text{the number of neighborhoods}$ ).
Step-2: Uniformly disaggregate data from census tract level into neighborhood level (Note:
internet data is only available at Census tract level)
Step-3: for each Neighborhood ( $i$ ) do
        Find attribute (pcat_all)  $S_i$  that represents the households with over 200 kbps in at
least one direction
        Convert  $S_i$  into percentage measure  $s_i$  by taking the average of the range
        Calculate Substandard Measurable Components Percentage ( $\mu_i$ ) =  $1 - s_i$ 
        if  $\mu_i \geq 0.5$  then  $\theta_i = 1$  else  $\theta_i = 0$  end
end
Note: pcat_all represents the number residential households with fixed high-speed connections
over 200 kbps in at least one direction per 1000 households

```

## References

- Abhyanker, R. (2014). *Nextdoor neighborhood social network method, apparatus, and system*. United States Patent.
- Agresti, A. (2003). *Categorical data analysis*. John Wiley & Sons.
- Aldrich, D. P., & Meyer, M. A. (2015). Social capital and community resilience. *American Behavioral Scientist*, 59(2), 254–269.
- Aman, J. J. C., & Smith-Colin, J. (2020). Transit deserts: Equity analysis of public transit accessibility. *Journal of Transport Geography*, 89, Article 102869. <https://doi.org/10.1016/j.jtrangeo.2020.102869>
- Aniello, C., Morgan, K., Busbey, A., & Newland, L. (1995). Mapping micro-urban heat island using LANDSAT TM and a GIS. *Computers & Geosciences*, 21(8), 965–969.
- Ashwood, J. S., Reid, R. O., Setodji, C. M., Weber, E., Gaynor, M., & Mehrotra, A. (2011). Trends in retail clinic use among the commercially insured. *The American Journal of Managed Care*, 17(11), Article e443.
- Bereitschaft, B. (2017). Equity in microscale urban design and walkability: A photographic survey of six Pittsburgh streetscapes. *Sustainability (Basel, Switzerland)*, 9(7), 1233. <https://doi.org/10.3390/su9071233>
- Bersamin, M., Todd, M., & Remer, L. (2011). Does distance matter? Access to family planning clinics and adolescent sexual behaviors. *Maternal and Child Health Journal*, 15(5), 652–659.
- Booth, K. M., Pinkston, M. M., & Poston, W. S. C. (2005). Obesity and the built environment. *Journal of the American Dietetic Association*, 105(5), 110–117. <https://doi.org/10.1016/j.jada.2005.02.045>
- Borowski, E., Ermagun, A., & Levinson, D. (2018). *Disparity of access: Variations in transit service by race, ethnicity, income, and auto availability*.
- Bureau, U. S. C. (2020). *American community survey; Table B19013; Table B19113; Table 03002* (Vol. 2020). <https://data.census.gov/cedsci/>.
- Chau, C. N., Zoellner, J. M., & Hill, J. L. (2013). Availability of healthy food: Does block group race and income matter? *Journal of Hunger & Environmental Nutrition*, 8(1), 22–38. <https://doi.org/10.1080/19320248.2012.758063>
- City of Dallas. (2017). The BIG picture: 2017 capital bond program. <https://www.dallascitynews.net/picturedallas>.
- Commission, F. C. (2019). *Internet access services: Status as of December 31, 2016*.
- Cummins, S., & Macintyre, S. (2002). "Food deserts"—Evidence and assumption in health policy making. *BMJ*, 325(7361), 436–438.
- Cutts, B. B., Darby, K. J., Boone, C. G., & Brewis, A. (2009). City structure, obesity, and environmental justice: An integrated analysis of physical and social barriers to walkable streets and park access. *Social Science & Medicine*, 69(9), 1314–1322. <https://doi.org/10.1016/j.socscimed.2009.08.020>
- Dallas City Hall. (2016). Sidewalk replacement program Dallas Texas (Vol. 2021, Issue May 17.). <https://dallascityhall.com:443/departments/public-works/Pages/SidewalkReplacementProgram.aspx>.
- Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive Justice? *Journal of Social Issues*, 31(3), 137–149. <https://doi.org/10.1111/j.1540-4560.1975.tb01000.x>
- El-Geneidy, A. M., & Levinson, D. M. (2006). *Access to destinations: Development of accessibility measures*.
- El-Geneidy, A. M., Tetreault, P., & Surprenant-Legault, J. (2010). Pedestrian access to transit: Identifying redundancies and gaps using a variable service area analysis. *In Transportation research board 89th annual meeting*.
- Emily Cochrane. (2021, August 10). *Senate passes \$1 trillion infrastructure*. The New York Times. <https://www.nytimes.com/2021/08/10/us/politics/infrastructure-bill-passes.html>.
- Engineering-USACE, U. A. C.o. (2014). *Micro Paver 6.5.7 user manual*.



- Ermini, L., & Hendry, D. F. (2008). Log income vs. linear income: An application of the encompassing principle\*. *Oxford Bulletin of Economics and Statistics*, 70(s1), 807–827. <https://doi.org/10.1111/j.1468-0084.2008.00531.x>
- Ewing, R., Handy, S., Brownson, R. C., Clemente, O., & Winston, E. (2006). Identifying and measuring urban design qualities related to walkability. *Human Kinetics Journals*, 3(s1), S223–S240.
- Farrington, D. P., & Welsh, B. C. (2002). Improved street lighting and crime prevention. *Justice Quarterly*, 19(2), 313–342. <https://doi.org/10.1080/07418820200095261>
- Florida, R. (2017). *The new urban crisis: How our cities are increasing inequality, deepening segregation, and failing the middle class-and what we can do about it*. Basic Books.
- Flowerdew, R., Manley, D. J., & Sabel, C. E. (2008). Neighbourhood effects on health: Does it matter where you draw the boundaries? *Social Science & Medicine* (1982), 66(6), 1241–1255. <https://doi.org/10.1016/j.socscimed.2007.11.042>
- Forkenbrock, D. J., Sheeley, J., & Program, N. C. H. R. (2004). *Effective methods for environmental justice assessment*. Transportation Research Board.
- Frackelton, A., Grossman, A., Palinginis, E., Castrillon, F., Elango, V., & Guensler, R. (2013). Measuring walkability: Development of an automated sidewalk quality assessment tool. *Suburban Sustainability*, 1(1). <https://doi.org/10.5038/2164-0866.1.1.4>
- Franco, M., Roux, A. V. D., Glass, T. A., Caballero, B., & Brancati, F. L. (2008). Neighborhood characteristics and availability of healthy foods in Baltimore. *American Journal of Preventive Medicine*, 35(6), 561–567. <https://doi.org/10.1016/j.amepre.2008.07.003>
- Gordon, C., Purciel-Hill, M., Ghai, N. R., Kaufman, L., Graham, R., & Wye, G. V. (2011). Measuring food deserts in New York City's low-income neighborhoods. *Health & Place*, 17(2), 696–700.
- Gordon-Larsen, P. (2006). Inequality in the built environment underlies key health disparities in physical activity and obesity. *Pediatrics (Evanston)*, 117(2), 417–424. <https://doi.org/10.1542/peds.2005-0058>
- Gunn, L. D., Mavoa, S., Boulang, C., Hooper, P., Kavanagh, A., & Giles-Corti, B. (2017). Designing healthy communities: Creating evidence on metrics for built environment features associated with walkable neighbourhood activity centres. *The International Journal of Behavioral Nutrition and Physical Activity*, 14(1). <https://doi.org/10.1186/s12966-017-0621-9>
- Habeb, D., Vargo, J., & Stone, B. (2015). Rising heat wave trends in large US cities. *Natural Hazards*, 76(3), 1651–1665.
- Handy, S. L., Boarnet, M. G., Ewing, R., & Killingsworth, R. E. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventive Medicine*, 23(2, Supplement 1), 64–73. [https://doi.org/10.1016/S0749-3797\(02\)00475-0](https://doi.org/10.1016/S0749-3797(02)00475-0)
- Hilmers, A., Hilmers, D. C., & Dave, J. (2012). Neighborhood disparities in access to healthy foods and their effects on environmental justice. *American Journal of Public Health*, 102(9), 1644–1654.
- Hirsch, J. A., Green, G. F., Peterson, M., Rodriguez, D. A., & Gordon-Larsen, P. (2017). Neighborhood sociodemographics and change in built infrastructure. *Journal of Urbanism*, 10(2), 181–197. <https://doi.org/10.1080/17549175.2016.1212914>
- Hong, J., & Chen, C. (2014). The role of the built environment on perceived safety from crime and walking: Examining direct and indirect impacts. *Transportation (Dordrecht)*, 41(6), 1171–1185. <https://doi.org/10.1007/s11116-014-9535-4>
- Hwang, J., Joh, K., & Woo, A. (2017). Social inequalities in child pedestrian traffic injuries: Differences in neighborhood built environments near schools in Austin, TX, USA. *Journal of Transport & Health*, 6, 40–49. <https://doi.org/10.1016/j.jth.2017.05.003>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.
- Jiao, J. (2017). Identifying transit deserts in major Texas cities where the supplies missed the demands. *Journal of Transport and Land Use*, 10(1), 529–540.
- Jiao, J., & Dillivan, M. (2013). Transit deserts: The gap between demand and supply. *Journal of Public Transportation*, 16(3). <https://doi.org/10.5038/2375-0901.16.3.2>
- Jiao, J., Moudon, A. V., Ulmer, J., Hurvitz, P. M., & Drewnowski, A. (2012). How to identify food deserts: Measuring physical and economic access to supermarkets in King County, Washington. *American Journal of Public Health* (1971), 102(10). <https://doi.org/10.2105/ajph.2012.300675>. e32–e39.
- Kaczynski, A., Potwarka, L., Smale, B., & Havitz, M. (2009). Association of Parkland proximity with neighborhood and park-based physical activity: Variations by gender and age. *Leisure Sciences*, 31, 174–191. <https://doi.org/10.1080/01490400802686045>
- LaPlante, J., & McCann, B. (2008). Complete streets: We can get there from here. *ITE Journal*, 78(5), 24.
- Leventhal, T., & Brooks-Gunn, J. (2000). The neighborhoods they live in: The effects of neighborhood residence on child and adolescent outcomes. *Psychological Bulletin*, 126(2), 309–337. <https://doi.org/10.1037/0033-2909.126.2.309>
- Lonescu, D. (2022, January 3). *How to prioritize equity in the infrastructure bill*. The Brookings Institution. <https://www.planetizen.com/news/2022/01/115696-how-prioritize-equity-infrastructure-bill>.
- Lucy, W. (1981). Equity and planning for local services. *Journal of the American Planning Association*, 47(4), 447–457. <https://doi.org/10.1080/0194368108976526>
- Luo, W., & Wang, F. (2003). Measures of spatial accessibility to health care in a GIS environment: Synthesis and a case study in the Chicago region. *Environment and Planning B: Planning and Design*, 30(6), 865–884.
- MacDonnell, S., Embuldeniya, D., Anderson, J., Roberts, P., Ratanshi, F., & Rexe, K. (2004). *Poverty by postal code: The geography of neighbourhood poverty, 1981–2001*. <https://doi.org/10.13140/rg.2.1.3201.3521>
- Maco, S. E., & McPherson, E. G. (2002). Assessing canopy cover over streets and sidewalks in street tree populations. *Journal of Arboriculture*, 28(6), 270–276.
- Martin, C., Perry, A. M., & Barr, A. (2021, December 17). *How equity isn't built into the infrastructure bill—And ways to fix it*. The Brookings Institution. <https://www.brookings.edu/blog/the-avenue/2021/12/17/how-equity-isnt-built-into-the-infrastructure-bill-and-ways-to-fix-it/>.
- McCormack, G. R., & Shiell, A. (2011). In search of causality: A systematic review of the relationship between the built environment and physical activity among adults. *The International Journal of Behavioral Nutrition and Physical Activity*, 8(1), 125. <https://doi.org/10.1186/1479-5868-8-125>
- Naylor, B., & Walsh, D. (2021, November 15). Biden signs the \$1 trillion bipartisan infrastructure bill into law. NPR. <https://www.npr.org/2021/11/15/1055841358/biden-signs-1t-bipartisan-infrastructure-bill-into-law>.
- Neutens, T. (2015). Accessibility, equity and health care: Review and research directions for transport geographers. *Journal of Transport Geography*, 43, 14–27.
- Nicholls, S. (2001). Measuring the accessibility and equity of public parks: A case study using GIS. *Managing Leisure*, 6. <https://doi.org/10.1080/13606710110084651>
- O'Sullivan, S., & Morrall, J. (1996). Walking distances to and from light-rail transit stations. *Transportation Research Record*, 1538(1), 19–26.
- Parks, J. R., & Schofer, J. L. (2006). Characterizing neighborhood pedestrian environments with secondary data. *Transportation Research. Part D, Transport and Environment*, 11(4), 250–263. <https://doi.org/10.1016/j.trd.2006.04.003>
- Perzynski, A. T., Roach, M. J., Shick, S., Callahan, B., Gunzler, D., Cebul, R., Kaelber, D. C., Huml, A., Thornton, J. D., & Einstadter, D. (2017). Patient portals and broadband internet inequality. *Journal of the American Medical Informatics Association*, 24(5), 927–932.
- Peterson, B., & Harrell, F. E. (1988). Partial proportional odds models and the LOGIST procedure. In *Vol. 11. SAS Users Group International, thirteenth annual conference, Orlando FL*.
- Ploeg, M. V., Breneman, V., Farrigan, T., Hamrick, K., Hopkins, D., Kaufman, P., ... Tuckermant, E. (2009). *Access to affordable and nutritious food: Measuring and understanding food deserts and their consequences: Report to congress*. <https://doi.org/10.22004/ag.econ.292130>. doi.
- Quadros, S. G. R., & Nassi, C. D. (2015). An evaluation on the criteria to prioritize transportation infrastructure investments in Brazil. *Transport Policy*, 40, 8–16.
- Redmon, J., & Farhadi, A. (2018). YoloV3: An incremental improvement. ArXiv Preprint ArXiv:1804.02767.
- Reduction, G. F. D. & Recovery(GFDRR). (2017). *Post-disaster needs assessments guidelines volume B community infrastructure*.
- Retting, R. (2017). *Pedestrian traffic fatalities by state*. Washington, DC, USA: Governors Highway Safety Association.
- Ripley, B., Venables, B., Bates, D. M., Hornik, K., Gebhardt, A., Firth, D., & Ripley, M. B. (2013). In Package 'mass'. *Cran R* (p. 538).
- Roberts, J., Koeser, A., Abd-Elrahman, A., Wilkinson, B., Hansen, G., Landry, S., & Perez, A. (2019). Mobile terrestrial photogrammetry for street tree mapping and measurements. *Forests*, 10(8), 701. <https://doi.org/10.3390/f10080701>
- Rochat, J. L., & Reiter, D. (2016). Highway traffic noise. *Acoustics Today*, 12(4). <https://acousticstoday.org/wp-content/uploads/2018/08/Highway-Traffic-Noise-Judith-L-Rochat.pdf>.
- Rogers, S., Gardner, K., & Carlson, C. (2013). Social capital and walkability as social aspects of sustainability. *Sustainability (Basel, Switzerland)*, 5(8), 3473–3483. <https://doi.org/10.3390/su5083473>
- Rollings, K. A., Wells, N. M., & Evans, G. W. (2015). Measuring physical neighborhood quality related to health. *Behavioral Sciences*, 5(2), 190–202. <https://doi.org/10.3390/bs5020190>
- Ross, C. E. (2000). Neighborhood disadvantage and adult depression. *Journal of Health and Social Behavior*, 41(2), 177–187. <https://doi.org/10.2307/2676304>
- Roux, A. V. D., & Mair, C. (2010). Neighborhoods and health. *Annals of the New York Academy of Sciences*, 1186(1), 125–145. <https://doi.org/10.1111/j.1749-6632.2009.05333.x>
- Shah, Y. U., Jain, S. S., Tiwari, D., & Jain, M. K. (2013). Development of overall pavement condition index for urban road network. *Procedia, Social and Behavioral Sciences*, 104, 332–341. <https://doi.org/10.1016/j.sbspro.2013.11.126>
- Srinivasan, S., O'Fallon, L. R., & Dearth, A. (2003). Creating healthy communities, healthy homes, healthy people: Initiating a research agenda on the built environment and public health. *American Journal of Public Health*, 93(9), 1446–1450. <https://doi.org/10.2105/AJPH.93.9.1446>
- Su, S., Zhou, H., Xu, M., Ru, H., Wang, W., & Weng, M. (2019). Auditing street walkability and associated social inequalities for planning implications. *Journal of Transport Geography*, 74, 62–76. <https://doi.org/10.1016/j.jtrangeo.2018.11.003>
- Talen, E., & Anselin, L. (1998). Assessing spatial equity: An evaluation of measures of accessibility to public playgrounds. *Environment and Planning A*, 30(4), 595–613.
- Tan, Z., Lau, K. K.-L., & Ng, E. (2016). Urban tree design approaches for mitigating daytime urban heat island effects in a high-density urban environment. *Energy and Buildings*, 114, 265–274.
- Taylor, P., & Fry, R. (2012). *The rise of residential segregation by income*. Washington, DC: Pew Research Center.
- Texas Tree Foundation. (2017). *Urban Heat Island Management Study Dallas 2017*. <http://www.texastrees.org/wp-content/uploads/2019/06/Urban-Heat-Island-Study-August-2017.pdf>.
- Texas Trees Foundation. (2019). Dallas urban forest master plan. <https://www.texastrees.org/projects/dallas-urban-forest-master-plan/>.
- Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J., O'Neil-Dunne, J. P., & Frank, L. D. (2016). Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription. *Health & Place*, 42, 54–62. <https://doi.org/10.1016/j.healthplace.2016.08.011>

- Walker, S. H., & Duncan, D. B. (1967). Estimation of the probability of an event as a function of several independent variables. *Biometrika*, 54(1/2), 167–179. <https://doi.org/10.2307/2333860>
- Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: Towards an integrated approach to defining health professional shortage areas. *Health & Place*, 11(2), 131–146. <https://doi.org/10.1016/j.healthplace.2004.02.003>
- Weibull, J. W. (1976). An axiomatic approach to the measurement of accessibility. *Regional Science and Urban Economics*, 6(4), 357–379.
- Wells, N. M., & Rollings, K. A. (2012). *The natural environment in residential settings: Influences on human health and function*.
- Wolch, J., Wilson, J. P., & Fehrenbach, J. (2005). Parks and park funding in Los Angeles: An equity-mapping analysis. *Urban Geography*, 26(1), 4–35. <https://doi.org/10.2747/0272-3638.26.1.4>
- Wrigley, N. (2002). “Food deserts” in british cities: Policy context and research priorities. *Urban Studies*, 39(11), 2029–2040.
- Xiao, Y., Wang, Z., Li, Z., & Tang, Z. (2017). An assessment of urban park access in Shanghai – implications for the social equity in urban China. *Landscape and Urban Planning*, 157, 383–393. <https://doi.org/10.1016/j.landurbplan.2016.08.007>
- Zegeer, C. V., Stewart, J. R., Huang, H. H., Lagerwey, P. A., Feaganes, J. R., & Campbell, B. J. (2005). *Safety effects of marked versus unmarked crosswalks at uncontrolled locations final report and recommended guidelines*. United States: Federal Highway Administration.