EMPIRICAL MEASURES OF PARK USE IN AMERICAN CITIES, AND THE DEMOGRAPHIC BIASES OF SPATIAL MODELS

JAMES SAXON

Center for Spatial Data Science and Harris School of Public Policy, University of Chicago

Abstract. City planners have a professional and ethical responsibility to provide public goods equitably. Parks improve mental and physical health by nurturing social cohesion and enabling physical activity. So who gets parks? Park access has traditionally been evaluated using constructed variables of *potential* access: distances, buffers, and gravity models. These models have major limitations: they ignore commutes and other more intricate mobility behaviors. To address these issues, I propose a nationally scalable, empirical measure of realized use. Using a dataset of smartphone locations, I identify visits to parks in the twenty largest American cities. I use these data to calibrate existing models, and then contrast the models with realized use. The spatial models are not simply imprecise; they systematically over-estimate realized access by minority populations. In other words, they understate inequity.

1. Introduction

This paper presents scalable measures of park visitation at the neighborhood level, constructed using GPS locations of smartphone users in tandem with park and green space boundaries from OpenStreetMap. The need for and role of parks is fundamentally different in urban, suburban, and rural environments, and this paper focuses on the twenty largest American cities. Park visits are identified as locations recorded within parks. This empirical measure of use automatically accounts for the spatial and economic complexity of urban routines. This approach is less expensive, more replicable and scalable, and spatially more precise than the survey measures used in the past to ascertain actual use. It is distinguished by unprecedented sample size, within and across cities. I share these aggregated visit data.

Parks' value arise from both the intrinsic pleasure of their use as well as their derived benefits. Recent work on park access has had two major focusses: the equity with which it is provisioned, and the impacts on mental health and physical activity. Most studies evaluate access by relying on distances, buffers, and gravity potentials to measure the spatial accessibility of resources (Markevych et al., 2017; Talen, 1998; Wolch et al., 2014). These data are generally derived from GIS data or acquired through expensive surveys. They assess *potential* access: resources that *could* be used. But for practical planning, academic studies, and consideration of equity, "potential access" is often the wrong metric.

Scholars have long defined equity as realized *use* rather than simply expenditures – in the context of parks, visitation rather than facilities. In the epidemiological space, the primary causal mechanism by which greenspaces impact population physiology is through use and direct contact. The importance of measuring actual spatial exposures is widely recognized by geographers (Kwan, 2012). There is good reason to expect potential access and realized use to differ dramatically. Space is just one of many barriers to access: geographers dating to Hägerstrand (1970) have recognized time constraints, and meaningful access also

E-mail address: jsaxon@uchicago.edu.

Date: October 22, 2020.

depends on population social, economic, and cultural factors (Penchansky and Thomas, 1981). The paper thus explores the agreement between use and the spatial access often used to proxy it. It suggests that proxying "access" through proximity results in a measure that is demographically biased rather than simply noisy.

After reviewing the literature (Section 2), I describe the data and its initial processing (Section 3). Supplementary materials extend the validation of their performance in this application. Section 4 describes the methods for calibrating models and contrasting them with data. Sections 5 and 6 present and discuss results: they show that demography predicts empirical use, even after controlling for spatial accessibility. I conclude with a discussion of the role and limitations of this new form of data for future work.

2. Literature Review: Measuring Access to Greenspace

2.1. **The current state of the art.** Two major geographic literatures focus on the spatial accessibility of parks and greenspaces. The first evaluates territorial or spatial equity in the provision of public services. The second measures the impacts of greenspace on health, chiefly though not exclusively through physical activity and stress reduction. Excellent recent reviews include Wolch et al. (2014) on park equity and Markevych et al. (2017) on the epidemiological consequences. Table A.1 of the supplementary materials provides an extensive but certainly incomplete listing of methodologies used over the past two decades. Seminal early work on the impact of various methodologies is by Talen and Anselin (1998).

In both literatures, park access is associated with spatial proximity, usually relying on one of three methods: the distance to the nearest facility, a gravity potential model, or the green area within a buffer of neighborhood. Label parks and greenspaces by p, and call their areas $A(p)$. Denote the network distance from neighborhood ℓ to park p by $\delta_{\ell p}$. The minimum distance to a park can then be written as

$$A_d(\ell) \equiv \min_p \delta_{\ell p}. \tag{1}$$

This minimum may be modified to consider only those facilities whose areas exceed a minimum threshold, $\{p|A(p) > A_{\min}\}$. Using identical notation, the gravity potential is a weighted sum over facilities, with parks weighted up by their area and down by the distance required to reach them, to some power α:

$$A_g(\ell, \alpha) \equiv \sum_p A(p)/\delta_{\ell p}^{\alpha}. \tag{2}$$

The buffered area $A_b(\cdot)$ of location ℓ is defined simply as the area $A(\cdot)$ of the city's greenspace G that intersects a circle of radius R centered at ℓ, $C_\ell(R)$:

$$A_b(\ell, R) \equiv A(G \cap C_\ell(R)). \tag{3}$$

These constructions are illustrated in Figure 1. All three appeal to a primacy of proximity natural to geographers familiar with Tobler's First Law, that "near things are more related than distant things" (Tobler, 1970).

The gravity and buffer models each have an explicit, tunable parameter: the radius R of the buffered area, and the decay rate α of the gravity potential. The minimum area threshold $A_{\min}$ of the minimum-distance metric is less-standard, although regularly recognized (see Logan et al., 2017). Beyond these parameters are many implicit choices regarding the origins and destinations between which to measure distance, and indeed how to calculate these distances. The choices made over both methods and parameters impact substantive conclusions. This indeterminacy in results according to methods was first articulated by Talen and Anselin (1998).

Recent methodological work has advocated refinements or sensitivity analyses for network as opposed to Euclidean distance and the use of park boundaries or formal entrances instead of centroids (Higgs et al., 2012), finer, parcel-level granularity for walking environments (Logan et al., 2017), comparisons of types of land cover (greenspace, trees, or grass) (Gascon et al., 2016; Reid et al., 2017), and the scale of buffer or shape of aggregation regions (Giles-Corti et al., 2005; Reid et al., 2018). On the other hand, this work has demonstrated mainly that results differ according to the choice of method, and they have provided less guidance as to the "right" choice. Lacking clear guidance, methodological consistency across applied work has been poor. Reviews have noted this as a major limitations of the literature (Brownson et al., 2009; Markevych et al., 2017; Wolch et al., 2014).

2.2. **Visitation, not proximity, is the relevant concept.** But how can the "right" choice be evaluated? The answer lies in the notion that *spatial access* is often not the intended concept in the first place. There are two levels of access – potential and realized – which may be thought of as ease of use and use itself (Khan and Bhardwaj, 1994). Theoretically, it is not potential access but *realized use* that is generally appropriate to measurements of the equity or impacts of greenspace. In many cases, proximity is just a simple, limited "proxy" for this more-fundamental quantity of use or exposure (Markevych et al., 2017; Sister et al., 2010).

First consider equity. From the outset, work on (territorial) equity of resource provision has focussed on the relation between need and *use*, rather pro forma availability (Davies, 1968). Indeed, Harvey emphasized in early work that equity can entail uneven distribution: "individuals have rights to equal levels of benefits, which means that there is unequal allocation according to need" (2009, p. 100). What matters is the distribution not of expenditures but of benefits – in the case of parks, not land but visitation (Boyne and Powell, 1991). This view is echoed resoundingly in the foundations and literatures of health access (Aday and Andersen, 1974; Donabedian, 1972; Institute of Medicine, 1993) and planning (American Institute of Certified Planners, 2016, §A1f). To translate the argument to the context of planning: the goal is not simply to spend money but to deliver a service. The park that lies vacant fails to do so.

In the health literature, park proximity is studied as a "treatment" affecting physical activity, and physical and mental health. Though parks also provide sound barriers (Van Renterghem et al., 2015), attenuate urban heat islands (Gronlund et al., 2015), and offer restorative views (Ulmer et al., 2016), the primary implicit causal mechanism of exposure is mediated by *use*. To adopt the language of medicine, the park "dose" delivered is not the amount of spatial access but the amount of use. Focusing on proximity instead can be likened to an "intent to treat" (ITT) rather than a "treatment on the treated" (TOT). The compliance, the relationship between the ITT and TOT, may not be constant across or within cities. This issue parallels the Uncertain Geographic Context Problem (UGCoP) articulated by Kwan (2012): measures of spatial "exposures" should reflect the contexts that individuals actually experience.

A major line of inquiry is the relationship between proximity and use, and activity or other "downstream" effects like general health. The causal chain linking proximity and activity and health frequently elides mediation through use. All three variables are commonly evaluated through survey reports (see supplement for table of methods); proximity and accessibility are also often derived using GIS, and equated. Survey measures, both generally and of activity and proximity in particular, have well-established limitations: high costs, poor instrument design and high recall error (van Poppel et al., 2010), and plunging response rates (Pew Research Center, 2019). Comparisons of objective and perceived proximity to parks have shown poor to moderate correspondence (Jilcott et al., 2007; Lackey and Kaczynski, 2009; Maddison et al., 2010).

In short, for equity and exposures the apposite quantity is more often *use* than *proximity*. Studies evaluating *use* directly need scalable, objective measures. Likewise, those focused on downstream outcomes like physical activity or general health, would benefit from isolating the causal step from park presence to use.

2.3. **Aspatial determinants of access.** This paper measures the use of parks, and evaluates the empirical relationship between use and proximity. Differences between the two are driven by temporal and social constraints on individuals' routines. These constraints have been recognized and studied for decades in literatures on activity spaces, time geography, and service provision. This work shares and advances those literatures' aim to understand individuals' real routines and access to resources.

Half a century ago, Horton and Reynolds (1970, 1971) proposed the study of activity spaces as "the subset of all urban locations with which the individual has direct contact as the result of day-to-day activities" (1971, p. 37). The present paper takes up this mantle, to understand "direct contact" with greenspaces. It also advances this field. Despite Horton and Reynold's intent, activity spaces have traditionally been operationalized using standard-deviational ellipses, minimum convex polygons, and buffers (Patterson and Farber, 2015). These regions can hardly be said to capture the spaces with which individuals come into "direct contact," in their routines. It is only recently that work has begun to itemize the individual locations – the actual contacts and specific locations.

In the same year that Horton and Reynolds introduced activity spaces, Hägerstrand (1970) proposed a parallel project to understand "the fate of the individual human being in an increasingly complicated environment." Hägerstrand weighed two possible strategies: first, to "sample life paths" as biologists band birds and second, to itemize the constraints – especially temporal – on individuals' routines. He chose to pursue constraints and laid the course for space-time geography. Work in the intervening decades has provided theoretical clarity to the temporal constraints on individuals and the "poles" – home and work – in their daily routines (Dijst, 1999a). It has shown how differences in individuals' temporal constraints can create differential accessibility, for instance by gender (Dijst, 1999b; Kwan, 1998). The value of the space-time approach is that it makes its assumptions explicit; the challenge is that analysts must explicitly model each realistic refinement. Often, the refinements lack realism. For example, time space geometries tend to view locations near home, work, or commute as accessible, since they entail brief deviations from fixed routines. In reality, a location halfway between home and work may never be visited, and the next neighborhood over may never be explored. Hägerstrand recognized from the outset that the choice among constraints was methodologically "piecemeal" and theoretically "artistic" (p. 20-21). Further, demanding data and computational requirements, have stymied adoption in applied and professional work (Geurs and van Wee, 2004).

Beyond time, a collection of social and economic barriers constrict access. Penchansky and Thomas (1981) famously outlined the "5As" of access to a resource – availability, accessibility, accommodation, affordability, and acceptability – encompassing its total supply, spatial distribution, opening hours, economic costs, and the cultural match between resource and users. For an example in the context of parks, Byrne and Wolch (2009) documented how design choices have historically been driven by majority preferences, privileging the *acceptability* to those groups. This dynamic continues to play out in American cities (Hui, 2017; Hyra, 2015). Social conditions also affect use. Sociologists have of late renewed their focus on the resources available in urban neighborhoods, and the interplay among forms of disadvantage (Desmond and Western, 2018; Sharkey and Faber, 2014). These factors affect residents' ability to use public resources; they therefore affect equity and dose.

A broad array of temporal, social, and economic factors affect the places that people use. Today, Hägerstrand's "banding" is done: five-sixths of adults in American cities own a smartphone (Pew Research Center, 2018). GPS data are often now available for research quite inexpensively or for free (as in the present case). These data record empirical activity spaces – visit locations – at scale. Geographers now can and should engage the "sampling" strategy that Hägerstrand rejected, to measure the "direct contacts" that Horton and Reynolds sought. This approach sweeps straight from a spatial distribution to the realities of routines. It does not isolate the impact of each constraint, but gathers them into a realistic view of the full effect.

2.4. **Measuring routines with GPS and social data.** It has been apparent for the past decade that smartphone GPS locations and other big data hold transformative potential for measuring and understanding individuals' routines. There are many explicit calls to use GPS data to understand activity spaces in general (Kwan, 2012) and parks in particular (James et al., 2015; Lachowycz and Jones, 2011; Markevych et al., 2017). These calls mirror a broader push for "big data" methods in the social sciences (e.g., Lazer et al., 2009).

The move to these data has been somewhat slow. Initial forays used fairly small, purpose-built datasets (Donaire-Gonzalez et al., 2016; Järv et al., 2014; Palmer et al., 2013; Zenk et al., 2011). In other cases, GPS trajectories were compared with accelerometer records to show that physical activity often takes place within parks, but these studies were limited in geographic scope and sample size, to just a few hundred subjects (Almanza et al., 2012; Rodríguez et al., 2012). More recently, larger scale smartphone and Twitter data have been applied to understand segregation in quotidian routines (Athey et al., 2019; Wang et al., 2018) and develop increasingly accurate and affordable models of aggregate mobility (Jiang et al., 2016; Song et al., 2010a,b). These data have recently entered the mainstream consciousness in the context of the coronavirus pandemic, as a resouce for contact tracing and mobility modelling. The work most relevant to the present application of *parks* leverages geotagged tweets and Flickr images to measure visitation at scale. However, that work has focused primarily on facilities rather than on neighborhood-scale geographies or the accessibility experienced by users (Donahue et al., 2018; Sessions et al., 2016).

To conclude: models of "accessibility" constructed purely from proximity dominate work on the equity and epidemiological consequences of park access, even though it is widely acknowledged that realized use is the more relevant concept. Decades of work on activity spaces, time constraints, and cultural and economic compatibility of resources and users have produced clear theory and evidence to expect an empirical divergence between these concepts. Over the past decade, there have been widespread calls to use GPS data to measure this. That is the project of this paper.

3. Data

This study uses cell phone data to identify locations visited by individual users. Users' home locations are imputed from their modal nighttime location, at the level of Census tracts. Destinations are drawn from OpenStreetMap data.

3.1. Spatial Context.

3.1. Spatial Context. The need for and provision of green and public spaces differs markedly in high- and low-density environments. This paper focuses on major American cities. These are defined as the twenty most-populous Census places; their boundaries and the shapes of tracts within them are both drawn from the Census. These cities are defined jurisdictions with a responsibility to provide equitable public resources.

The boundaries of parks and open spaces are drawn from OpenStreetMap data (OpenStreetMap contributors, 2018). These data include all levels of park systems – local, state, and national. They are available for the entire United States and they are increasingly used for accessibility research (Logan et al., 2017). Parks are defined from the following tags:

- `leisure` any of park, dog park, playground, nature reserve, garden, or golf course; or
- `landuse` is recreation ground, `natural` is beach, or `boundary` is a protected area.

Except for museums, buildings within parks are included. So too are pitches and plazas that may not have any greenspace. Because the focus of this paper is on *realized* access, golf courses are included and I do not attempt to remove other paying facilities. OpenStreetMap ways (roads) are used to calculate network distances from tracts to parks. The locations of major highways (motorways, trunk, primary, and secondary roads) and rail and subway lines are also used to eliminate park "visits" from arterials traversing parks. Like the parameters of traditional models, each of these decisions may be contested.

3.2. Cell Phone Locations. The dataset of behaviors consists of 27 billion GPS locations or "pings" generated on active applications on users' smartphones in May 2017. Each line of data includes identifiers for the device and application, a timestamp, a location, and an estimated precision for that location. Data were supplied by Carto. Figure 2a shows the raw locations. Home locations, constructed according to methods described below, are suppressed from this illustration to preserve anonymity.

Some applications have access only to a user's approximate location. I first suppress these data, as well as points with estimated precision worse than half of a kilometer. I also delete duplicate points, recorded simultaneously by several applications on a single device. I then trim this dataset to 10-kilometer buffers around the twenty largest American cities, using the Census boundaries already mentioned. This ensures a generous buffer around every neighborhood while factorizing the data into separate, computationally manageable jobs for each city.

Using point-in-polygon merges, I record the Census tract that contains each point, and flag points that fall within parks, as previously defined. I suppress visits within 10 meters of major roads and railways. Beautified parkways may make commutes and transit more enjoyable, but they are not "park visits." Spatial processing amounting to around 10,000 CPU hours is performed using the OpenScienceGrid. (Pordes et al., 2007; Sfiligoi et al., 2009) Figure 2b shows the classified points.

For each device, I identify the modal night-time Census tract (midnight to 6 am), and call this tract its "home" if it registers multiple (night-time) locations there. To ensure adequate data to accurately identify the residence, I select a "restricted" sample, including those devices that

(1) are observed in all three "thirds" of the month,
(2) record at least three nights at home, and
(3) are observed at least 100 times.

These requirements do not substantively affect either the baseline visit rates or subsequent findings. Results derived with looser or more stringent requirements are presented in supplementary materials. After these requirements, 455 thousand devices are retained that collectively record over a billion unique locations. The median device records 549 locations. These data record devices' direct contacts in the city: empirical activity spaces and visits to parks.

Park visitation rates, defined as days with a park visit, are averaged over devices in the restricted sample, at the Census tract level. I discard tracts outside the city limits. I also record the specific parks visited; one visit is counted per park, per day. Because the data has finite precision, I avoid spurious park visits by rejecting locations recorded within 100 meters (Euclidean) of a device's home. For this purpose only, the point-location of the home is defined as the centroid of the night-time locations within the home tract.

This study relies on smartphone GPS locations as a record of human mobility. Though this form of data is no longer new, it has not been applied to the measurement of the use of greenspace. It is critical to this work to probe and understand its representativeness. National surveys show that over three quarters of Americans have smartphones; this rises to better than five sixths in urban areas. Smartphone adoption is fairly evenly balanced by race and ethnicity, but the poor, less educated, and elderly (particularly over 65) are under-represented. (Pew Research Center, 2018)

A consistent picture emerges by contrasting the cell phone sample with existing datasets. Each cell phone user is anonymous; all that is known are times and locations, from which homes are imputed. Figure 3 shows normalized histograms of race, ethnicity, poverty, and educational attainment at the Census tract level, for the largest six American cities. These histograms are weighted by official population estimates from the ACS, as well as by the number of devices assigned as "resident" within each tract. These histograms contrast where ACS populations and cell phones devices "live." They probe whether the normalized sample rate is systematically high or low in Black versus non-Black, or high versus low-poverty neighborhoods. The device weights exhibit bias towards whiter, more educated, and less poor neighborhoods, as predicted by the Pew reports, but the histograms are, on the whole, strikingly consistent. Any sample biases *between* tracts presumably also hold *within* them, but this is not measurable.

An appendix describes continued validation of the dataset for the present application. It includes regression-based approaches to the sample composition, and shows that phone use (number of locations recorded) is consistent between white and minority neighborhoods. Within the restricted sample, phone use is weakly correlated with park visits. Controlling for phone usage does not affect subsequent results.

3.3. **Demographic Covariates.** Tract populations, racial and ethnic composition, poverty rates, and educational attainment are drawn from the five-year estimates of the US Census's 2017 American Community Survey (ACS). Race is defined as the fraction of the tract population that is Black and ethnicity is the fraction Hispanic. The poverty rate is the fraction of households below the Federal poverty line. Education attainment is the share of the adult population (over 25 years old) who hold a bachelor's degree.

4. Analysis Methods

This paper proceeds through three steps. The first, already described, is the significant computational task of extracting park visitation rates from location data. This Section describes the analysis of these measurements. I first calibrate spatial access models to data, and then use simple regressions to ask for whom those models over- or under-predict use.

4.1. **Evaluating Models of Spatial Accessibility.** I first aim to evaluate correlations between spatial access and use, and to understand the sensitivity of these correlations on model parameters. To do this, I must first evalute the spatial models.

A number of methods have been used to measure neighborhoods' potential access to parks. I focus on the minimum distance, gravity potential, and buffered area as expressed in Equations 1-3. As already noted, each of these has one tunable parameter: the minimum area of the closest park, the distance decay rate of the gravity model, and the radius of the buffer. The distance buffer is the simple, Euclidean (i.e., non-network) construction of a disk centered at tracts' centroids, in a local coordinate reference system.

For the minimum-distance and gravity models, distances are measured between the tract centroid and the park, along a walking road network that excludes motorways and trunk roads (OpenStreetMap contributors, 2018). To calculate park destinations, a 10 m grid is overlaid on the park boundaries and each intersection between the grid and the boundary is considered, along with the park centroid and a "representative point." The "distance to the park" is the minimum network distance from the tract centroid to any of these park points, evaluated using postgres version 10, postgis 2.5.2, and pgrouting version 2.3. I account for the distances to "snap" the centroids and park edges to the road network. Parks are included that fall within a (Euclidean) distance of up to 10 km of the home. This is the same buffer used to divide the data among cities in the first place. Because the gravity potential diverges when distances approach 0, distances less than 100 m are set to 100 m.

In addition to these "classical" models of access, I fit Poisson models to the usage data, along the lines of Sessions et al. (2016) and Donahue et al. (2018). Tract-park pairs are included that are separated by a Euclidean distance of less than 10 km. I also exclude "severe" network distance outliers, for which the walking distance exceeds 15 km (more than the maximal Manhattan distance on a grid). This requirement is binding mainly in New York City. For example, it is less than 9 km from Staten Island to Battery Park in Manhattan, but it is impossible to reach it with only the walking network (the Staten Island Ferry is not included).

The models are estimated using Generalized Estimating Equations (GEE), clustered at the Census tract level and weighted by Census tract populations. The dependent variable is the visits per device to park or greenspace p from neighborhood location ℓ. Mirroring the traditional models, I include only area and distance as independent variables. These models use a log link, so the lambda parameter of the Poisson (its mean and width) scales exponentially with the regressors ($\lambda \sim e^{\boldsymbol{\beta X}}$). The logarithm of the area is used, since it outperforms the raw value. Three functional forms $f(\cdot)$ are used for the distance: (1) its raw value, (2) its logarithm, and (3) "fixed effect" indicators in 0.1 km steps up to a final "overflow" bin at > 10 km (in which case the function yields a vector). The final method offers a non-parametric approach, and performs on par with the others. All three use network distance to park boundaries, as described for the gravity potential. The model is thus of the form,

$$\text{visits per device}_{\ell,p} \sim \text{Pois}\left(\exp\left(\beta_0 + \beta_a \log(A_p) + \beta_d f(d_{\ell p})\right)\right), \tag{4}$$

where β_i are parameter coefficients and the intercept. Aggregating fitted values over parks available to each tract yields access measures.

I present the correlation to use for each spatial model, and show the correlations' dependence on the parameters of the distance, gravity, and buffer models.

4.2. **Equity in Potential and Realized Access.** One can then ask where the calibrated models over- and under-predict use. To do this, I perform (ecological) regressions of neighborhoods' potential and realized access, as a function of their demographic characteristics. I evaluate three models in each of twenty cities. The models are:

(a) spatial access $\sim$ race + ethnicity
(b) realized access $\sim$ race + ethnicity
(c) realized access $\sim$ race + ethnicity + spatial access

The intent of this sequence is, first, to contrast the relationships with race and ethnicity of potential versus realized access. Second, by adding potential access – the dependent variable from (a) – as an explanatory variable in model (c), I ask whether that control fully explains differences in realized use by minorities.

The dependence of park use on area and proximity is different in different cities, reflecting the scale and spatial structures of inhabitants' routines. Accordingly, potential access is fit to data separately in each city. These calibrated models represent the "best case scenario" of spatial potential access. It is the use that would be derived from a map, if each city's activity space *scale* were known. The baseline model of potential, spatial access – the dependent variable of model (a) and the explanatory spatial access of model (c) – is the GEE network distance model. (Alternatives are included in the supplementary materials.) Race and ethnicity are defined as tracts' fraction Black or Hispanic, as drawn from the ACS.

The comparison is within and not *among* cities, so potential and realized access are normalized in each city to a population-weighted mean access value of 1. The unit of observation is the population-weighted Census tract. Each model is estimated using weighted least squares. Because the aim is to show what is experienced by the population, rather than any causal mechanism, spatial dependencies are not included. Supplementary materials show that point estimates and uncertainties are qualitatively consistent when estimated with a unweighted, heteroskedastic and spatial autocorrelation (HAC) robust method.

5. Results

Figure 5 displays realized park use data for Chicago. Tract-level data for all other cities is available at the journal site.

5.1. **The Performance of Spatial Models of Accessibility.** Figure 5 also displays access modeled for Chicago, using the parameters that maximize the correlations to observed use. The agreement between use and spatial accessibility is far from perfect. Devices from tracts on the wealthy North Side of the city are seen in parks more often than expected, while major parks on the West and South Sides fail to generate the predicted use. The correlations to realized use of the spatial models are plotted in Figure 4 as a function of their parameters, for the six largest American cities. The maximal correlations for all cities are available in supplementary materials.

In all cities, the performance of the traditional proxies is sensitive to their parameters: the minimum scale of the closest park, the decay parameter for the gravity potential, and the radius of the buffers. In short, calibration matters. Further, the maximal correlations are reached at different parameter values in different cities. This reflects different scales of activity spaces in different cities, and suggests that using simple buffers for epidemiological studies will capture different exposures or treatments in different cities. The GEE models are more performant in Chicago, Houston, and Philadelphia, and comparable in New York, Los Angeles, and Phoenix. Among the GEE models, those with distance and distance fixed effects out-perform the one using $\log(\text{distance})$, with the exception of the distance-based model in New York City.

In most cases, the overall correlations between realized use and classical models are moderate, between 0.3 and 0.5. In only Chicago, San Francisco, and Charlotte do the correlations exceed 0.6, for the buffer and gravity models. The Poisson models fare somewhat better, with the distance-based model achieving a 0.75 correlation to use in Chicago, but the average across cities hovers between 0.46 and 0.53. (The correlations are actually slightly higher for models using Euclidean distance; see supplement.)

A hint at the reason for these moderate correlations is shown in Figure 6, with the distances travelled by Chicago residents to Chicago parks. Especially for minority populations, a substantial fraction of park use is far from home. Seen within the framework of activity spaces, this is no surprise. Most adults commute, and may use parks en route to or near their place of work, as well as at home. Taking as examples two important parks in Chicago's Loop (business district), visitors to Grant Park have traveled an average of 7 km, and those to the River Walk have come 5.3 km. That visitors travel so far likely reflects the convenience of use (at work) or appeal of distinctive amenities (like the river) that purely spatial models cannot capture. They assume that public resources are consumed close to home, at levels that fall off uniformly with distance and which are enhanced only by area.

5.2. **Equity in Potential and Realized Access.** Figure 7 presents point estimates and 95% confidence limits for the regression models, across cities. It is a visual depiction of a specification table.

Replacing potential with realized access as the dependent variable (a to b), the parameter estimates for race and ethnicity become more negative (shift left) in every instance except for race in Seattle – although many of these shifts are not significant. In other words, inequality in park access is more severe than would be expected from a purely spatial analysis. Model (c) controls for potential, spatial accessibility using the GEE distance model. As can be seen in Figure 4, the quality of this control – its correlation to use – varies between cities.

The spatial control makes the parameter for fraction Black and Hispanic less negative in a majority of cities. In other words, the spatial accessibility of parks accounts for some of the reduced park use in neighborhoods with high minority fractions. This is expected. Spatial access to parks can be "bought" through housing, and it may be expensive to live by a desirable park. If minority populations are financially worse off, they may be unable to do so. However, these differences are usually small, and the parameters remain negative. In other words, access is worse for minorities than the spatial measure predicts. The supplement includes results with the fixed effects model of potential access, with Euclidean instead of network distance (as well as with modified requirements on the device sample). It is worth emphasizing that while the spatial models are sensitive to their parameters, the findings regarding race, space, and park use are robust to the model chosen.

6. Discussion

6.1. **Methodological Implications.** This paper has had three methodological aims:

(1) to derive and share empirical measures of use of urban resources,
(2) to calibrate existing models of potential access, and
(3) to probe residual discrepancies between measures and models.

The paper shows how modern data sources can be used to characterize activity spaces and "direct contacts" at scale. I have suggested that (2) is possible only to a limited degree because urban scale changes between cities. Planners and others who consider the effects of parks should be aware that (3) the models may be biased and overstate access by minorities.

The intent of this work is not to criticize standard models of potential access derived from spatial proximity. The models are necessary metrics for purely spatial analysis of access. For example, parks departments may have limited influence over neighborhoods' safety or foot traffic, and so the spatial distribution of facilities may be the only salient lever in a limited context. Further work is needed to quantify the impact of each aspatial determinant of use, and to evaluate in particular the extent to which modern, space-time measures capture the discrepancies observed here.

But proximity should never be confused for a delivered service. When spatial measures are used as predictors of use, they exhibit regular biases – both along class lines and because they ignore (measurable) out-of-home activity spaces as fundamental as commutes. While commutes can be implemented through space-time measures, these data also capture the results of social forces, intricate routines, and the distinctive appeal of certain parks. This work reemphasizes the critical impact on access of well-known and theoretically well-developed aspatial dimensions, and provides measurements when direct exposures are the relevant parameter. While these limitations are theoretically well-established, these data facilitate a transition to this intended concept.

For those interested in the impacts of exposures to parks as a "treatment" affecting other downstream conditions, the spatial models exhibit major limitations. I had hoped to measure the "empirical parameters" of the distance, gravity, and buffer models, to provide guidance for other studies where usage data are not available. This is not possible. The correspondence between between modeled and measured (potential and realized) access varies across cities. Different cities have different, endemic spatial structures. Activity spaces change non-trivially from place to place, and a one-off calibration offers no panacea. For example, the correlation of the area buffer to realized use is maximized at radii from 0.7 km in Denver and 0.8 km in New York, to the close to the upper bound limit tested of 10 km, in Austin, Jacksonville, and Seattle. This variable correspondence between spatial proximity and park exposures hampers comparisons between cities of buffer-based findings on health or activity.

It is also worth noting that the buffers that maximize correlations to use are generally, larger than the commonly used buffer of 400 m (0.25 mi). That buffer is sometimes justified from the perceived mobility of young children and the elderly (e.g., Wolch et al., 2005), whose daily trajectories may be less far reaching than working adults. Because these populations have low smartphone ownership (Pew Research Center, 2018; Rideout et al., 2010), they are likely underrepresented in this sample. There are thus valid reasons for retaining traditional buffer radii, in view of the present evidence.

The basic routines of urban residents are well measured, in the sense that the core business districts are known and employment origin-destination matrices are available in products like the Census Bureau's LEHD Origin-Destination Employment Statistics (LODES), and these can be incorporated into space-time measures of accessibility. Other routines and realities – like neighborhood extent or crime – are difficult or expensive to measure. The dataset presented here offers the means of doing this. Measurements of park visits by GPS devices "resident" in each neighborhood, across cities are released as a supplement to this work. These data complement existing methods and data sources – LIDAR, shapefiles, and surveys. GPS data are increasingly available, but have high computational overhead and are usually not "public," though derived datasets have become far more available through the coronavirus pandemic. This paper aims to facilitate and stimulate their use – for parks in large cities – through the release of appropriately anonymized, aggregate data.

6.2. **Substantive Implications for Equity in Urban Environmens.** This paper's findings also contribute to the broader literature on resource equity in urban neighborhoods. In recent decades, sociologists and economists have emphasized the multiple, dimensions of poverty (Sen, 1980, 1993) and the mechanisms through which these influences themselves interact (Desmond and Western, 2018; Sharkey and Faber, 2014). Similarly, Sampson (2011) has advocated ecological measures of neighborhoods that are not mere aggregates over residents' Census characteristics, but which reflect the social, physical, and institutional assets of the neighborhood. Park access is a clear example of such a measure. It is a valuable resource in its own right, and a driver of downstream conditions like health. This study quantifies inequity in provisioned resources, and illustrates the systematic interplay between demographic factors and the consumption of (spatially) available resources.

6.3. **The Limitations and Potential of the Data.** It must by this point be clear that the new data also carry significant challenges. They are a convenience sample, biased towards wealthier and whiter populations; that bias is not corrected for here. National surveys and comparisons to Census data both suggest that disadvantaged populations are also underrepresented. If this is true within tracts, then the worst off are the most likely missing within each tract – and they are more likely to be missing in tracts with high levels of disadvantage. Race and class are tightly intertwined in the United States. Since the data suggest that disadvantage limits park consumption, if the sampled populations are better off than the tract as a whole, the demographic biases between potential and realized park access described in this work would tend to be understated.

Moreover, the scale of biases in the sample rate across tracts is moderate. An appendix shows that levels of app use are not predictive of neighborhood demographics, and are weakly correlated park visits. Still, it is not possible with these data to measure the representativeness of device owners' behaviors. Of particular concern for parks, children and the elderly have lower smartphone ownership so, as already mentioned, these more-sensitive populations are likely under-represented.

Although baseline estimates are not substantially affected by changes in the sample requirements, the choices that plague models are altered and not avoided. Defining use still requires a sequence of decisions: what a park is, how long one must remain there for a visit to "count," how to define the home, and so forth. These decisions echo choices in past work over the appropriate buffer (or park definition). Further, the visits measured in this paper are not the only possible "exposure" to a park. A park can be seen and enjoyed from an apartment tower, and woods along a highway mitigate noise in adjacent communities. This study has understood "presence" in a binary sense – a location is either inside a park polygon or it is not, and vetoed only within a fixed 100 meter buffer of the home. Potential refinements including treating locations probabilistically as a function of their estimated precision, or accounting for the workplace or other frequent non-home locations. However, these strategies are likely to be overtaken by the inexorable march of technology: GPS chips achieve higher precision today than they did in 2017. In densely populated regions, location measurements also leverage WiFi-based services.

This analysis has aggregated measures of visitation at the Census tract level, allowing the data to be matched to existing Census measures. Studying individual access – as advocated by geographers since Hägerstrand – is tempting, but raises substantial privacy concerns for suppliers, and individuals cannot generally be matched to detailed demographic data. The ability to assemble individual level covariates remains an advantage of surveys. Still, the data have immense promise. They capture empirical, objective human activity spaces, continuously and at scale. They permit comparison of populations and cities with unprecedented sample size. And they are increasingly obtainable.

7. Conclusions

The beneficial effects of parks and open spaces have rigorously established impact on health and well-being; the benefit of this public good is not equally shared. This paper has described a new approach to measuring realized exposures to parks. This analysis is at once at meter-level granularity and national scope; it showcases the potential of new and increasingly available data sources for research in geography, epidemiology, and planning. Responsible use of these data will require continued care for representativeness and user privacy. It is my hope that appropriate, aggregate measures will stimulate attention to realized use in general, and to the factors that generate differences with respect to measures of potential, spatial access to parks in particular.

Contrasting the data with models shows that spatial proxies for park access have different correlations with realized use in different geographic contexts. These models also exhibit consistent biases in regards race and ethnicity. As compared with actual use of parks, purely spatial models of access tend to understate inequity.

References

Aday, L. A. and Andersen, R., "A Framework for the Study of Access to Medical Care.," *Health Services Research,* **9**:3, 208-220 (1974).

Almanza, E., Jerrett, M., Dunton, G., Seto, E., and Pentz, M. A., "A study of community design, greenness, and physical activity in children using satellite, GPS and accelerometer data," *Health & Place,* **18**:1, 46-54 (2012). DOI: 10.1016/j.healthplace.2011.09.003.

American Institute of Certified Planners, *AICP Code of Ethics and Professional Conduct.* (April 2016). American Planning Association, Washington DC.

Athey, S., Ferguson, B., Gentzkow, M., and Schmidt, T., *Experienced Segregation.* (February 2019). Stanford University Working Paper. pp. 1-72

Boyne, G. and Powell, M., "Territorial justice: A review of theory and evidence," *Political Geography Quarterly,* **10**:3, 263-281 (1991). DOI: 10.1016/0260-9827(91)90038-V.

Brownson, R. C., Hoehner, C. M., Day, K., Forsyth, A., and Sallis, J. F., "Measuring the Built Environment for Physical Activity: State of the Science," *American Journal of Preventive Medicine,* **36**:4, Supplement, S99-S123.e12 (2009). DOI: 10.1016/j.amepre.2009.01.005.

Byrne, J. and Wolch, J., "Nature, race, and parks: past research and future directions for geographic research," *Progress in Human Geography,* **33**:6, 743-765 (2009). DOI: 10.1177/0309132509103156.

Davies, B., *Social Needs and Resources in Local Services: A Study of Variations in Provision of Social Services between Local Authority Areas*, (1968). Joseph Rowntree, London UK.

Desmond, M. and Western, B., "Poverty in America: New Directions and Debates," *Annual Review of Sociology,* **44**:1, 305-318 (2018). DOI: 10.1146/annurev-soc-060116-053411.

Dijst, M., "Action space as planning concept in spatial planning," *Netherlands journal of housing and the built environment,* **14**:2, 163-182 (June 1, 1999a). DOI: 10.1007/BF02496820.

Dijst, M., "Two-earner families and their action spaces: A case study of two dutch communities," *GeoJournal,* **48**:3, 195-206 (1999b). DOI: 10.2307/41147371.

Donabedian, A., "Models for Organizing the Delivery of Personal Health Services and Criteria for Evaluating Them," *The Milbank Memorial Fund Quarterly,* **50**:4, 103-154 (1972). DOI: 10.2307/3349436.

Donahue, M. L., Keeler, B. L., Wood, S. A., Fisher, D. M., Hamstead, Z. A., and McPhearson, T., "Using social media to understand drivers of urban park visitation in the Twin Cities, MN," *Landscape and Urban Planning,* **175**, 1-10 (2018). DOI: 10.1016/j.landurbplan.2018.02.006.

Donaire-Gonzalez, D., Valentín, A., de Nazelle, A., Ambros, A., Carrasco-Turigas, G., Seto, E., *et al.*, "Benefits of Mobile Phone Technology for Personal Environmental Monitoring," *JMIR Mhealth Uhealth,* **4**:4, 126 (Nov 10, 2016). DOI: 10.2196/mhealth.5771.

Gascon, M., Cirach, M., Martínez, D., Dadvand, P., Valentn, A., Plasncia, A., *et al.*, "Normalized difference vegetation index (NDVI) as a marker of surrounding greenness in epidemiological studies: The case of Barcelona city," *Urban Forestry & Urban Greening,* **19**, 88-94 (2016). DOI: 10.1016/j.ufug.2016.07.001.

Geurs, K. T. and van Wee, B., "Accessibility evaluation of land-use and transport strategies: review and research directions," *Journal of Transport Geography,* **12**:2, 127-140 (2004). DOI: 10.1016/j.jtrangeo.2003.10.005.

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., *et al.*, "Increasing walking: How important is distance to, attractiveness, and size of public open space?" *American Journal of Preventive Medicine,* **28**:2, Supplement 2, 169-176 (2005). DOI: 10.1016/j.amepre.2004.10.018.

Gronlund, C. J., Berrocal, V. J., White-Newsome, J. L., Conlon, K. C., and O'Neill, M. S., "Vulnerability to extreme heat by socio-demographic characteristics and area green space among the elderly in Michigan, 19902007," *Environmental Research,* **136**, 449-461 (2015). DOI: 10.1016/j.envres.2014.08.042.

Hägerstrand, T., "What about people in regional science?" *Papers in Regional Science,* **24**:1, 7-24 (1970). DOI: 10.1111/j.1435-5597.1970.tb01464.x.

Harvey, D., *Social Justice and the City*, Revised edition, (2009). University of Georgia Press, Athens, Georgia.

Higgs, G., Fry, R., and Langford, M., "Investigating the Implications of Using Alternative GIS-Based Techniques to Measure Accessibility to Green Space," *Environment and Planning B: Planning and Design,* **39**:2, 326-343 (2012). DOI: 10.1068/b37130.

Horton, F. E. and Reynolds, D. R., "Action Space Formation: A Behavioral Approach to Predicting Urban Travel Behavior," *Highway Research Record,* **322**, 136-148 (1970).

Horton, F. E. and Reynolds, D. R., "Effects of Urban Spatial Structure on Individual Behavior," *Economic Geography,* **47**:1, 36-48 (1971). DOI: 10.2307/143224.

Hui, M., "Pups and public space: What D.C.s dog parks say about the changing city," *The Washington Post,* **Local**, (August 27, 2017).

Hyra, D., "The back-to-the-city movement: Neighbourhood redevelopment and processes of political and cultural displacement," *Urban Studies,* **52**:10, 1753-1773 (2015). DOI: 10.1177/0042098014539403.

Institute of Medicine, *Access to Health Care in America*; ed. Millman, M., (1993). National Academies Press, Washington DC. DOI: 10.17226/2009.

James, P., Banay, R. F., Hart, J. E., and Laden, F., "A Review of the Health Benefits of Greenness," *Current Epidemiological Reports,* **2**, 131142 (2015). DOI: 10.1007/s40471-015-0043-7.

Järv, O., Ahas, R., and Witlox, F., "Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records," *Transportation Research Part C: Emerging Technologies,* **38**, 122-135 (2014). DOI: 10.1016/j.trc.2013.11.003.

Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., and González, M. C., "The TimeGeo modeling framework for urban mobility without travel surveys," *Proceedings of the National Academy of Sciences,* **113**:37, E5370-E5378 (2016). DOI: 10.1073/pnas.1524261113.

Jilcott, S. B., Evenson, K. R., Laraia, B. A., and Ammerman, A. S., "Association between physical activity and proximity to physical activity resources among low-income, midlife women.," *Preventing chronic disease,* **4**:1, A04 (January, 2007). PMC: PMC1832127

Khan, A. A. and Bhardwaj, S. M., "Access to Health Care: A Conceptual Framework and its Relevance to Health Care Planning," *Evaluation & the Health Professions,* **17**:1, 60-76 (1994). DOI: 10.1177/016327879401700104.

Kwan, M.-P., "Space-Time and Integral Measures of Individual Accessibility: A Comparative Analysis Using a Point-based Framework," *Geographical Analysis,* **30**:3, 191-216 (1998). DOI: 10.1111/j.1538-4632.1998.tb00396.x.

Kwan, M.-P., "The Uncertain Geographic Context Problem," *Annals of the Association of American Geographers,* **102**:5, 958-968 (2012). DOI: 10.1080/00045608.2012.687349.

Lachowycz, K. and Jones, A. P., "Greenspace and obesity: a systematic review of the evidence," *Obesity Reviews,* **12**:5, e183-e189 (2011). DOI: 10.1111/j.1467-789X.2010.00827.x.

Lackey, K. J. and Kaczynski, A. T., "Correspondence of perceived vs. objective proximity to parks and their relationship to park-based physical activity," *International Journal of Behavioral Nutrition and Physical Activity,* **6**:1, 53 (Aug 11, 2009). DOI: 10.1186/1479-5868-6-53.

Lazer, D., Pentland, A., Adamic, L., Aral, S., Barabási, A.-L., Brewer, D., *et al.*, "Computational Social Science," *Science,* **323**:5915, 721-723 (2009). DOI: 10.1126/science.1167742.

Logan, T., Williams, T., Nisbet, A., Liberman, K., Zuo, C., and Guikema, S., "Evaluating urban accessibility:

leveraging open-source data and analytics to overcome existing limitations," *Environment and Planning B: Urban Analytics and City Science,* , 2399808317736528 (2017). DOI: 10.1177/2399808317736528.

Maddison, R., Jiang, Y., Hoorn, S. V., Mhurchu, C. N., Exeter, D., and Utter, J., "Perceived Versus Actual Distance to Local Physical-Activity Facilities: Does It Really Matter?" *Journal of Physical Activity and Health,* **7**:3, 323-332 (2010).

Markevych, I., Schoierer, J., Hartig, T., Chudnovsky, A., Hystad, P., Dzhambov, A. M., *et al.*, "Exploring pathways linking greenspace to health: Theoretical and methodological guidance," *Environmental Research,* **158**, 301-317 (2017). DOI: 10.1016/j.envres.2017.06.028.

OpenStreetMap contributors, "Planet data retrieved from http://overpass-api.de." March, 2018. Accessed March 7, 2018. https://www.openstreetmap.org

Palmer, J. R. B., Espenshade, T. J., Bartumeus, F., Chung, C. Y., Ozgencil, N. E., and Li, K., "New Approaches to Human Mobility: Using Mobile Phones for Demographic Research," *Demography,* **50**:3, 1105-1128 (June 1, 2013). DOI: 10.1007/s13524-012-0175-z.

Patterson, Z. and Farber, S., "Potential Path Areas and Activity Spaces in Application: A Review," *Transport Reviews,* **35**:6, 679-700 (2015). DOI: 10.1080/01441647.2015.1042944.

Penchansky, R. and Thomas, J. W., "The Concept of Access: Definition and Relationship to Consumer Satisfaction," *Medical Care,* **19**:2, 127-140 (1981). DOI: 10.2307/3764310.

Pew Research Center, "Mobile Fact Sheet." February, 2018. Accessed October 15, 2018.

Pew Research Center, *Growing and Improving Pew Research Center's American Trends Panel.* (February 2019). Pew Research Center, Washington DC.

van Poppel, M. N M., Chinapaw, M. J M., Mokkink, L. B., van Mechelen, W., and Terwee, C. B., "Physical Activity Questionnaires for Adults: A Systematic Review of Measurement Properties," *Sports Medicine,* **40**:7, 565-600 (2010). DOI: 10.2165/11531930-000000000-00000.

Pordes, R., Petravick, D., Kramer, B., Olson, D., Livny, M., Roy, A., *et al.*, "The open science grid," *Journal of Physics: Conference Series,* **78**, 012057 (July, 2007). DOI: 10.1088/1742-6596/78/1/012057.

Reid, C. E., Clougherty, J. E., Shmool, J. L., and Kubzansky, L. D., "Is All Urban Green Space the Same? A Comparison of the Health Benefits of Trees and Grass in New York City," *International Journal of Environmental Research and Public Health,* **14**:11, (2017). DOI: 10.3390/ijerph14111411.

Reid, C. E., Kubzansky, L. D., Li, J., Shmool, J. L., and Clougherty, J. E., "It's not easy assessing greenness: A comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City," *Health & Place,* **54**, 92-101 (2018). DOI: 10.1016/j.healthplace.2018.09.005.

Rideout, V. J., Foehr, U. G., and Roberts, D. F., *Generation M^2: Media in the Lives of 8- to 18-Year-Olds.* (January 2010). Kaiser Family Foundation, Menlo Park, CA. pp. 18

Rodríguez, D. A., Cho, G.-H., Evenson, K. R., Conway, T. L., Cohen, D., Ghosh-Dastidar, B., *et al.*, "Out and about: Association of the built environment with physical activity behaviors of adolescent females," *Health & Place,* **18**:1, 55-62 (2012). DOI: 10.1016/j.healthplace.2011.08.020. Active Living Research

Sampson, R. J., "Neighborhood effects, causal mechanisms and the social structure of the city." In *Analytical Sociology and Social Mechanisms.* Demeulenaere, P. ed. (Cambridge University Press, 2011), pp. 227–249.

Sen, A., "Equality of What?" In *The Tanner Lectures on Human Values.* McMurrin, S. M. ed. (Cambridge University Press, 1980), pp. 196–220.

Sen, A., "Capability and Well-Being." In *The Quality of Life*, Ch. 3. Nussbaum, M. and Sen, A. eds. (Oxford University Press, 1993), pp. 33–48. DOI: 10.1093/0198287976.003.0003.

Sessions, C., Wood, S. A., Rabotyagov, S., and Fisher, D. M., "Measuring recreational visitation at U.S. National Parks with crowd-sourced photographs," *Journal of Environmental Management,* **183**, 703-711 (2016). DOI: 10.1016/j.jenvman.2016.09.018.

Sfiligoi, I., Bradley, D. C., Holzman, B., Mhashilkar, P., Padhi, S., and Wurthwein, F.. "The Pilot Way to Grid Resources Using glideinWMS," at *WRI World Congress on Computer Science and Information Engineering,* (March 2009). 428–432 DOI: 10.1109/CSIE.2009.950.

Sharkey, P. and Faber, J. W., "Where, When, Why, and For Whom Do Residential Contexts Matter? Moving Away from the Dichotomous Understanding of Neighborhood Effects," *Annual Review of Sociology,* **40**:1, 559-579 (2014). DOI: 10.1146/annurev-soc-071913-043350.

Sister, C., Wolch, J., and Wilson, J., "Got green? addressing environmental justice in park provision.," *GeoJournal,* **75**, 229242 (2010). DOI: 10.1007/s10708-009-9303-8.

Song, C., Qu, Z., Blumm, N., and Barabási, A.-L., "Limits of Predictability in Human Mobility," *Science,* **327**:5968, 1018-1021 (2010a). DOI: 10.1126/science.1177170.

Song, C., Koren, T., Wang, P., and Barabási, A.-L., "Modelling the scaling properties of human mobility," *Nature Physics,* **6**, 818-821 (September, 2010b). DOI: 10.1038/nphys1760.

Talen, E. and Anselin, L., "Assessing Spatial Equity: An Evaluation of Measures of Accessibility to Public Playgrounds," *Environment and Planning A: Economy and Space,* **30**:4, 595-613 (1998). DOI: 10.1068/a300595.

Talen, E., "Visualizing Fairness: Equity Maps for Planners," *Journal of the American Planning Association,* **64**:1, 22-38 (1998). DOI: 10.1080/01944369808975954.

Tobler, W., "A Computer Movie Simulating Urban Growth in the Detroit Region," *Economic Geography,* **46**, 234-240 (1970). DOI: 10.2307/143141.

Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J., ONeil-Dunne, J. P., *et al.*, "Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription," *Health & Place,* **42**, 54-62 (2016). DOI: 10.1016/j.healthplace.2016.08.011.

Van Renterghem, T., Forssén, J., Attenborough, K., Jean, P., Defrance, J., Hornikx, M., *et al.*, "Using natural means to reduce surface transport noise during propagation outdoors," *Applied Acoustics,* **92**, 86-101 (2015). DOI: 10.1016/j.apacoust.2015.01.004.

Wang, Q., Phillips, N. E., Small, M. L., and Sampson, R. J., "Urban mobility and neighborhood isolation in America's 50 largest cities," *Proceedings of the National Academy of Sciences,* **115**:30, 7735-7740 (2018). DOI: 10.1073/pnas.1802537115.

Wolch, J., Wilson, J. P., and Fehrenbach, J., "Parks and Park Funding in Los Angeles: An Equity-Mapping Analysis," *Urban Geography,* **26**:1, 4-35 (2005). DOI: 10.2747/0272-3638.26.1.4.

Wolch, J. R., Byrne, J., and Newell, J. P., "Urban green space, public health, and environmental justice: The challenge of making cities just green enough," *Landscape and Urban Planning,* **125**, 234-244 (2014). DOI: 10.1016/j.landurbplan.2014.01.017.

Zenk, S. N., Schulz, A. J., Matthews, S. A., Odoms-Young, A., Wilbur, J., Wegrzyn, L., *et al.*, "Activity space environment and dietary and physical activity behaviors: A pilot study," *Health & Place,* **17**:5, 1150-1161 (2011). DOI: 10.1016/j.healthplace.2011.05.001.

Figure Captions

(1) Walking paths to parks and buffered areas ($R = 2$ km) are shown for a tract centroid in Washington DC.
(2) Raw and classified data in Philadelphia. Locations in home Census tracts are suppressed.
(3) Race, ethnicity, and socioeconomic status by Census tract, weighted by population estimates from the ACS or the number of devices "resident" in the tract. This shows the consistency of the sampled and Census distributions.
(4) Correlations between access models and realized use, for the six largest American cities, weighted by Census tract populations. For the traditional models, correlations exceed 0.6 only in Chicago, San Francisco, and Charlotte (not shown).
(5) Use and modeled accessibility of parks in Chicago, by decile. The minimum distance model requires a park area of 0.3 square kilometers, the gravity potential uses a decay parameter of 0.6, and the buffer radius is set to 2.4 km. Airports and offshore tracts are removed.
(6) Kernel density estimate and cumulative distribution for park visits in Chicago, according to the fraction of the neighborhood population that is Black. More than a sixth of park visits observed are to parks more than 10 km from the home. This fraction is even larger for neighborhoods whose residents are overwhelmingly Black. For scale reference, it is about 25 km from the Loop (CBD) to the southern edge of the city.
(7) Point estimates (circles) and 95% confidence limits (bars) are presented for three models of access across twenty cities. All models are at the Census tract level, and weighted by official population estimates. Each color represents a separate model of accessibility to parks in each city. The dependent variables and spatial accessibility (third column) are all normalized per city. In model (a), potential access is regressed on race and ethnicity. Potential access is assessed with a GEE model using log area and distance. In the remaining models realized access is the dependent variable. Model (b) parallels the first, with simply race and ethnicity. Model (c) adds the GEE accessibility from the first model as an explanatory variable for realized use.

Figure 1. Walking paths to parks and buffered areas (R = 2 km) are shown for a tract centroid in Washington DC.

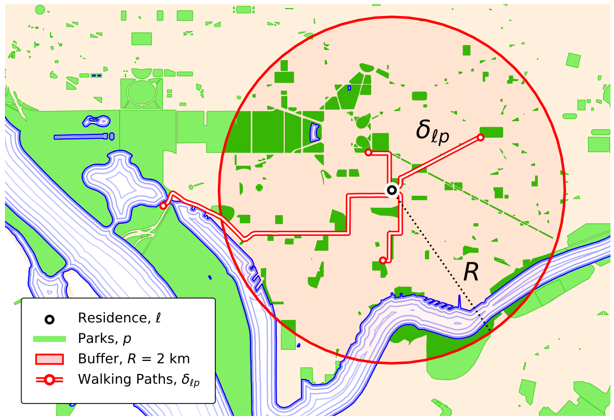


Figure 2. Raw and classified data in Philadelphia. Locations in home Census tracts are suppressed.

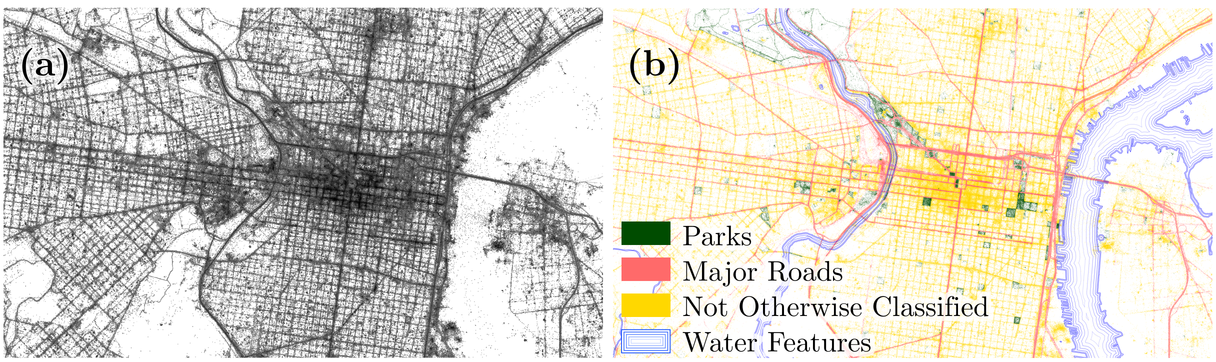


Figure 3. Race, ethnicity, and socioeconomic status by Census tract, weighted by population estimates from the ACS or the number of devices "resident" in the tract. This shows the consistency of the sampled and Census distributions.

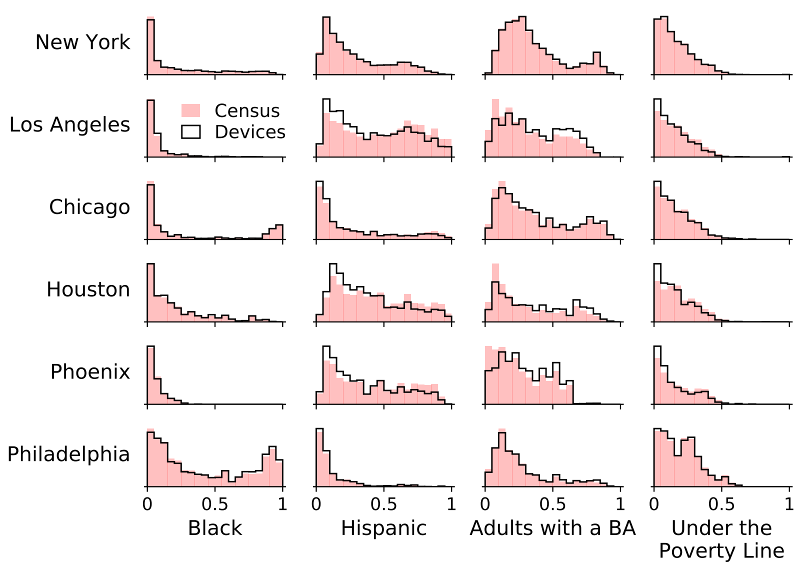


FIGURE 4. Correlations between access models and realized use, for the six largest American cities, weighted by Census tract populations. For the traditional models, correlations exceed 0.6 only in Chicago, San Francisco, and Charlotte (not shown).

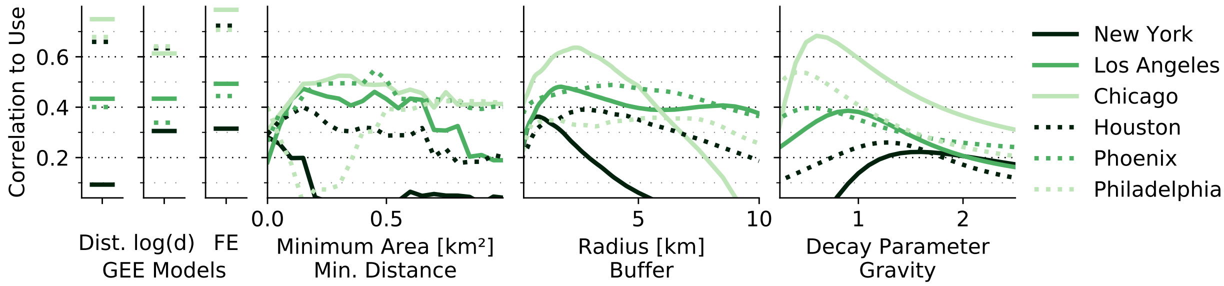


FIGURE 5. Use and modeled accessibility of parks in Chicago, by decile. The minimum distance model requires a park area of 0.3 square kilometers, the gravity potential uses a decay parameter of 0.6, and the buffer radius is set to 2.4 km. Airports and offshore tracts are removed.

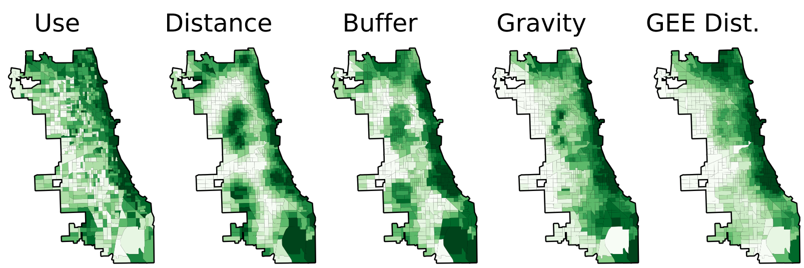


FIGURE 6. Kernel density estimate and cumulative distribution for park visits in Chicago, according to the fraction of the neighborhood population that is Black. More than a sixth of park visits observed are to parks more than 10 km from the home. This fraction is even larger for neighborhoods whose residents are overwhelmingly Black. For scale reference, it is about 25 km from the Loop (CBD) to the southern edge of the city.

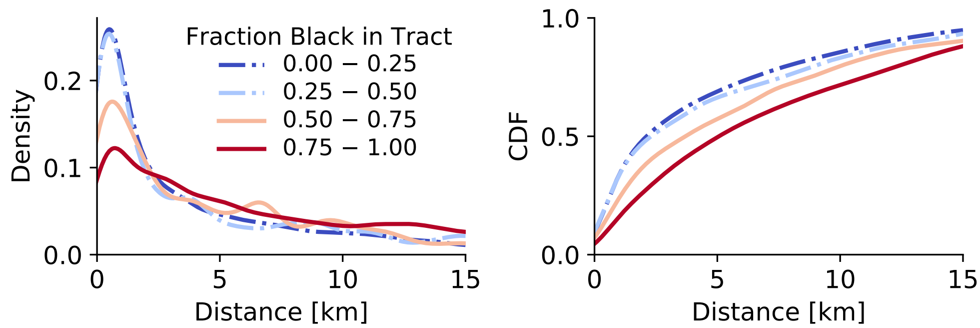
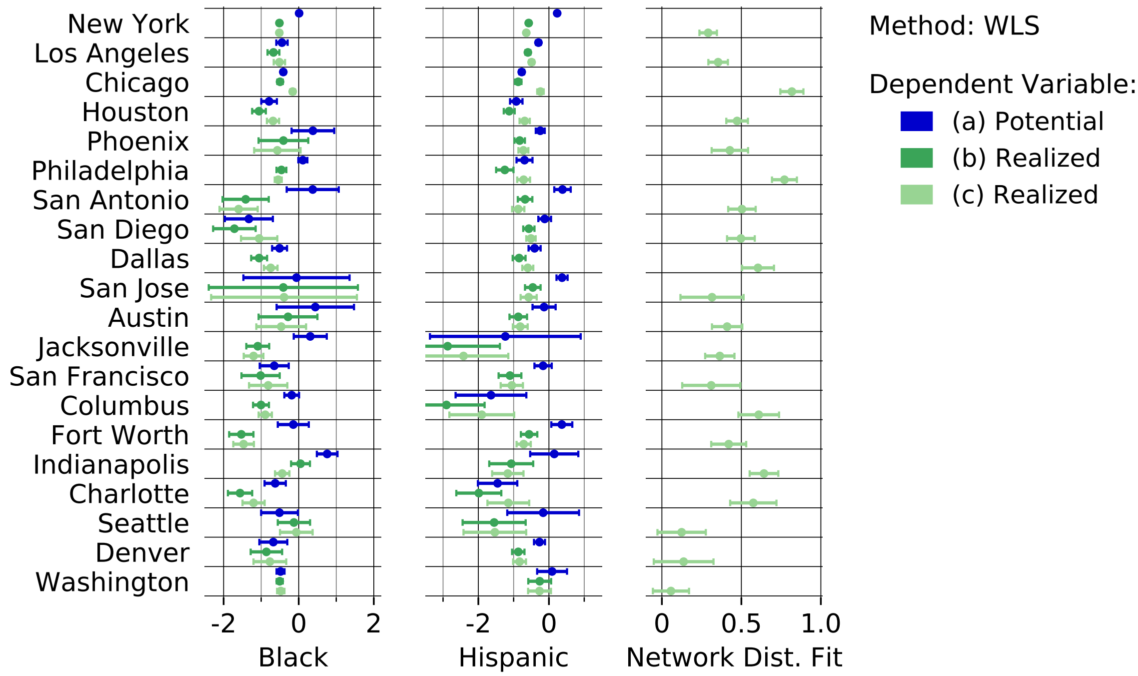


Figure 7. Point estimates (circles) and 95% confidence limits (bars) are presented for three models of access across twenty cities. All models are at the Census tract level, and weighted by official population estimates. Each color represents a separate model of accessibility to parks in each city. The dependent variables and spatial accessibility (third column) are all normalized per city. In model (a), potential access is regressed on race and ethnicity. Potential access is assessed with a GEE model using log area and distance. In the remaining models realized access is the dependent variable. Model (b) parallels the first, with simply race and ethnicity. Model (c) adds the GEE accessibility from the first model as an explanatory variable for realized use.



Empirical Measures of Park Use in American Cities, and the Demographic Biases of Spatial Models: Supplemental Materials

James Saxon

Appendix A. Methodologies of the Recent Literature on Park Accessibility

Table A.1. Methodologies used in studies of parks, playgrounds, greenspace, and environmental accessibility. The third through fifth columns refer to the minimum-distance to a facility, buffered areas (or sometimes counts), and gravity potential models. The minimum-distance and buffer models can be dichotomized to be functionally identical: there is non-zero park area within a buffer of radius R, and the nearest park is less than R away.

Authors	Year	Dist.	Buff.	Grav.	Survey	Params / Other Methods	Subject	Outcomes and Notes
Talen	1998	✓	✓	✓		Container, average travel cost.	Equity	
Talen and Anselin	1998	✓		✓		Average travel cost	Equity	
Lindsey et al.	2001		✓			Buffer: 0.5 mi. Dichotomized.	Equity	
Troped et al.	2001	✓			✓		Health	Activity
Nicholls	2001	✓	✓				Equity	
Takano et al.	2002				✓		Health	Mortality
Giles-Corti and Donovan	2003			✓	✓	Surveyer-recorded built env.	Health	Individual, social, and environmental walking.
Hartig et al.	2003					Experimental: view and walk	Psych.	Blood pressure & mental test
Huston et al.	2003				✓		Health	Physical activity
Leyden	2003				✓		Health	Walkability → social cohesion
Witten et al.	2003		✓			Buffer: 0.75 & 2 km. Count	Planning	
Powell et al.	2003				✓		Health	Walkability → activity
Humpel et al.	2004				✓		Health	Perceived env. → Walking
Smoyer-Tomic et al.	2004	✓	✓			Buffer: 0.8 km	Equity	Playgrounds
Bedimo-Rung et al.	2005	✓				Other: total acreage, equitability.		Conceptual study emphasizes use.
Hoehner et al.	2005		✓		✓	Buffer: 0.4 km	Health	Perceived and objective env. → activity.
Wolch et al.	2005		✓			Buffer: 0.25 mi. Area & Dichotomized.	Equity	
Cohen et al.	2006	✓	✓		✓	Buffer: 1 mi (count). Grav. decay from fit.	Health	Accelerometer-based activity.
Gordon-Larsen et al.	2006		✓			Buffer: 5 mi	Health	Obesity and activity
Hillsdon et al.	2006			✓		Decay param. fit from data.	Health	Activity
Maas et al.	2006		✓			Buffer: 1 km and 3 km.	Health	Perceived general health.
Barbosa et al.	2007	✓	✓			Buffer: 0.5 km	Planning	Greenspace chronically underprovided.
Cohen et al.	2007	✓			✓	Self-reported proximity + user counts.	Health	Activity
Hume et al.	2007				✓	Perceived walkability	Health	Accelerometer-derived activity
Jilcott et al.	2007	✓			✓		Health	Activity. Poor agreement perceived v. objective.
Nielsen and Hansen	2007	✓			✓	Reported distance	Health	Stress and Obesity
Berman et al.	2008					Experimental: view and walk	Psych.	Directed attention mental tasks
Boone et al.	2009		✓			Buffer: 400 m. Dichotomized.	Equity	
Lackey and Kaczynski	2009		✓		✓	Dichotomized buffer: 0.75 km	Health	Poor agreement reported v. objective access.
Maas et al.	2009		✓			Buffer: 1 km and 3 km.	Health	Social cohesion as a mediator of health.
Sister et al.	2010					Thiessen polygons!	Equity	
Maddison et al.	2010	✓			✓		Health	Poor agreement reported v. objective distances.
Singh et al.	2010				✓		Health	Obesity
Dai	2011					Two-stage floating catchment	Equity	
Higgs et al.	2012	✓					GIS	Distance calculation methods are consequential.
Thompson et al.	2012					Percent greenspace in UK Census ward.	Planning	Stress: self-report & cortisol.
de Vries et al.	2013		✓			Buffer: 0.5 km, area per person	Health	Stress and social cohesion.
Wen et al.	2013			✓		Grav. $\beta = 2$, restricted to 7 nearest.	Equity	
Vaughan et al.	2013					Number and area of facilities in tract.	Equity	
Rigolon and Flohr	2014	✓				Modified to account for walkability	Equity	
Astell-Burt et al.	2014		✓			Buffer: 1 km	Equity	
Reklaitiene et al.	2014	✓				Three cateogories, cutoffs at 0.3 and 1 km	Health	General health and depression
Reyes et al.	2014		✓			Buffer derived from trip data	Equity	
Hughey et al.	2016					Area intersecting Census block group.	Equity	
Dadvand et al.	2016		✓		✓	Buffer: 0.3 km, dichotomized. Survey: 10 min.	Health	Subjective general health
James et al.	2016		✓			Buffer: 0.25 & 1.25 km, mean NDVI	Health	Mortality
Ulmer et al.	2016		✓			Buffer: parks in 0.5 km, trees in 0.25 km	Health	Overall health
Logan et al.	2017	✓				Dist. refinments: parcel level network dist.	GIS	
Tan and Samsudin	2017		✓			Buffer: 0.4 km. Other: population ratios.	Equity	
Reid et al.	2018		✓			Buffer: 5 radii, 0.1 to 2 km, & subjective.	Health	Self-rated health. Larger buffers ⇒ bigger effects.
Engemann et al.	2019		✓			Buffer: 210×210 m square.	Health	Mental health

Appendix B. Sample Validation

The main text presents histograms of Census characteristics weighted by Census populations and device counts. That comparison suggests that though device populations are shifted towards whiter and wealthier tracts, this bias is moderate. This appendix presents continued validation of the sample, with particular attention to sample composition and the variables used in the present analysis.

I begin with an alternative, regression-based approach to the sampling rate. Call the number of devices per tract d_t and that for the city $d_c = \sum_t d_t$. Likewise, call the ACS populations p_t and p_c. The tract sampling rate is $s_t = d_t/p_t$, and the city-wide rate is $s_c = d_c/p_c$. These city-wide rates are shown in the left panel of Table B.1. I define the rate normalized with respect to the city-wide rate, $\overline{s}_t = s_t/s_c$.

Note that the city-wide rate can be thought of as the population-weighted average of the tract rates:

$$s_c = \frac{\sum_t p_t s_t}{\sum_t p_t} = \frac{\sum_t p_t d_t/p_t}{\sum_t p_t} = \frac{\sum_t d_t}{\sum_t p_t} = \frac{d_c}{p_c}.$$

A tract's normalized rate can also be thought of as its share of the citywide devices as compared with its share of the total population:

$$\overline{s}_t = \frac{s_t}{s_c} = \frac{d_t/p_t}{d_c/p_c} = \frac{d_t/d_c}{p_t/p_c}.$$

To evaluate the dependence of the city-normalized tract sampling rates $\overline{s}_t$ on tract-level population characteristics X, I evaluate single-parameter, robust regressions using iteratively reweighted least squares (IRLS). The population characteristics X are, at turns, race, ethnicity, bachelor's attainment of the adult population, and household poverty rate. Each is expressed as a fraction ranging from 0 to 1. Call α the intercept and β the coefficient of X to estimate. The regression takes the form

$$\overline{s}_t \sim \alpha + \beta X.$$

Tracts like airports and parks may have zero formal residents and therefore extremely high sampling rates. I exclude these and other extreme outliers from the regression by dropping the tracts with the highest 5% of sampling rates. The best-fit parameters $\hat{\beta}$ are reported in the right panel of Table B.1. The intercepts are not reported.[1]

Parameters are significant and as high as 1. This means that a 1% change in a tract characteristic from its citywide mean value corresponds to a 1% change in the sampling rate. In most cities, however, the dependence is much smaller, and the domain of values observed for each characteristic should also be considered (see Figure 3). More typical is a fractional difference in the sampling rate of 0.5, between neighborhoods with no or all households in poverty, or all non-Black or Black, etc. It is notable that the differences in sampling rates are far less severe than those observed with other digital records like Twitter or Foursquare. (Anselin and Williams, 2016) Smartphone usage is more widespread than location-tagged

[1]The population-weighted average value of $\overline{s}_t$ is 1 in each city, but the value of X for which $\alpha + \beta X = 1$ depends on the population composition of the city along X, and the intercept α does as well. For example, consider where X represents tracts' fraction Hispanic, varying between 0 and 1. Then contrast two cities: one with almost all tracts completely White (non-Hispanic, $X = 0$) and the other with most tracts mainly Hispanic ($X = 1$). In the city with very few Hispanics, $1 \approx \alpha + \beta \times 0 = \alpha$, so α will be close to 1 (its value for the non-Hispanic tracts). In the Hispanic city, the average value of X approaches 1, so $1 = \alpha + \beta \times 1$ and $\alpha = 1 - \beta$.

Table B.1. Sampling rates by city for the baseline sample, and its dependence on race, ethnicity and socioeconomic status. The left-hand side presents regression coefficients for single-parameter, robust regressions of unweighted Census tracts, after dropping the five percent of tracts with the highest sampling rate – usually airports, uninhabited locations, etc.

	Pop. [k]	Dev. [k]	Sample Rate	Linear Regression Parameters Black	Hispanic	Educ.	Poverty
New York	8560	191.4	0.022	0.14*	0.11*	0.08	0.15
Los Angeles	3945	133.9	0.034	0.10	−0.44*	0.61*	−0.43*
Chicago	2722	95.1	0.035	0.14*	−0.39*	0.34*	−0.18
Houston	2231	90.6	0.041	−0.20	−0.63*	0.90*	−1.24*
Phoenix	1567	51.2	0.033	−0.18	−0.59*	0.74*	−0.76*
Philadelphia	1570	42.1	0.027	0.19*	−0.29*	0.16	0.14
San Antonio	1433	72.6	0.051	0.97*	−0.82*	0.91*	−1.19*
San Diego	1370	53.5	0.039	−0.96*	−0.63*	0.76*	−1.15*
Dallas	1340	61.6	0.046	−0.07	−0.71*	0.74*	−1.25*
San Jose	991	29.5	0.030	2.35	−0.46*	0.62*	−1.05*
Austin	900	44.1	0.049	−0.64	−0.48*	0.40*	−0.66*
Jacksonville	873	54.1	0.062	−0.24*	0.66	0.32	−0.76*
San Francisco	864	51.1	0.059	−0.14	−0.62	0.75*	−0.54
Columbus	776	50.1	0.065	−0.40*	−0.39	0.33*	−0.50*
Fort Worth	820	41.9	0.051	−0.10	−0.53*	0.56*	−0.96*
Indianapolis	856	55.1	0.064	−0.36*	−0.62	0.39*	−0.75*
Charlotte	802	39.5	0.049	0.10	−0.59*	0.22	−0.40
Seattle	688	34.9	0.051	−0.78	−0.83	0.56*	−0.68
Denver	678	29.2	0.043	0.10	−0.48*	0.48*	0.07
Washington	672	56.8	0.084	0.13	0.08	−0.19	0.55*

* $p < 0.001$

tweeting or Foursquare check-ins. The present comparisons are also derived from homes instead of destinations, and are normalized accordingly.

Table B.2 shows the same information, but as the slope to intercept ratios β/α. This represents the fractional change from a hypothetical tract with $X = 0$ to 1, changing for example from a tract with no Blacks to a completely Black one. Tables B.3 and B.4 show sampling rates and dependence on neighborhood characteristics, for the restricted sample.

Figure B.1 shows the physical locations of over- and under-represented neighborhoods, according to the unrestricted sample. Over-represented regions are of three stripes: city-centers (perhaps due to tourists or hotels), places with very low (or even 0) official population and locations with 24-hour operations (hospitals, airports, transportation hubs, and major factories).

It is worth noting the large differences in the overall sampling rate between cities, ranging from 0.022 devices per person in New York City to 0.084 in Washington DC. Note that these values should *not* be conceptualized as the fraction of residents in the sample. It is the number of devices per person. Residents may have multiple devices, and visitors to the city may also be included. The data supplier provides limited information on the sampling structure, and does not explain this variation. One natural explanation is that different cities

TABLE B.2. This table displays the slope over intercept (β/α) for the regressions of Table B.1. This presentation affords the simpler interpretation of a percentage change from a constant baseline.

	Linear Regressions: Slope / Intercept			
	Black	Hispanic	Educ.	Poverty
New York	0.15*	0.12*	0.08	0.16
Los Angeles	0.10	−0.36*	0.76*	−0.41*
Chicago	0.14*	−0.35*	0.38*	−0.17
Houston	−0.19	−0.50*	1.26*	−1.02*
Phoenix	−0.18	−0.48*	0.94*	−0.68*
Philadelphia	0.20*	−0.28*	0.17	0.14
San Antonio	1.07*	−0.54*	1.21*	−1.03*
San Diego	−0.96*	−0.56*	1.22*	−1.10*
Dallas	−0.07	−0.56*	0.98*	−1.03*
San Jose	2.60	−0.42*	0.85*	−1.01*
Austin	−0.62	−0.43*	0.52*	−0.63*
Jacksonville	−0.23*	0.74	0.37	−0.72*
San Francisco	−0.15	−0.59	1.40*	−0.54
Columbus	−0.38*	−0.41	0.40*	−0.48*
Fort Worth	−0.10	−0.47*	0.72*	−0.88*
Indianapolis	−0.35*	−0.61	0.47*	−0.70*
Charlotte	0.10	−0.57*	0.25	−0.39
Seattle	−0.80	−0.85	0.98*	−0.70
Denver	0.11	−0.44*	0.65*	0.07
Washington	0.15	0.09	−0.19	0.63*

* $p < 0.001$

have different amounts of business and tourist travel and shipping hubs within city limits. These visitors would affect the observed population of "residents." This explanation does not appear to be correct, however. For one thing, New York City has one of the lowest rates despite high levels of visitors. For another, the over-sampling in Washington DC does not seem to stem from an excess of visitors.

The paper adopts a "restricted sample" for the nominal analysis, from devices less likely to be travelers. These devices must

(a) show up at their imputed "home" on at least three distinct nights,
(b) be present within the city in the first, middle and final third of the month, and
(c) record more than 100 unique locations recorded across the month.

These requirements reduce the sample size by nearly two thirds, but do not affect the high rate in Washington or the low rate in New York. After these requirements the highest sample rates are in Columbus and Indianapolis, which are not known as hubs of tourism.

It is worth emphasizing that differences in the sampling rate are analytically distinct from biases in the averaged rates in park use. It matters less if thirty or three hundred devices are observed in tract, than if the device owners' behaviors are representative of those of their neighbors. This is not assessed in the preceding exercises. It is of course impossible to compare those who are observed with those who are *not* observed. What *can* be said is

Table B.3. Sampling rates by city for the restricted sample, and its dependence on race, ethnicity and socioeconomic status. The left-hand side presents regression coefficients for simple, robust regressions of unweighted Census tracts, after dropping the five percent of tracts with the highest sampling rate – usually airports, uninhabited locations, etc.

	Pop. [k]	Dev. [k]	Sample Rate	Linear Regression Parameters Black	Hispanic	Educ.	Poverty
New York	8560	64.1	0.007	0.22*	0.04	−0.12*	0.01
Los Angeles	3945	49.0	0.012	−0.06	−0.77*	1.05*	−1.37*
Chicago	2722	35.6	0.013	0.18*	−0.37*	0.11	−0.14
Houston	2231	31.3	0.014	−0.09	−0.75*	0.96*	−1.42*
Phoenix	1567	19.5	0.012	−0.60	−0.82*	1.04*	−1.27*
Philadelphia	1570	15.3	0.010	0.26*	−0.49*	0.08	−0.09
San Antonio	1433	24.6	0.017	0.48	−1.04*	1.18*	−1.75*
San Diego	1370	19.6	0.014	−0.85	−0.67*	0.76*	−1.25*
Dallas	1340	22.4	0.017	−0.16	−0.91*	0.98*	−1.65*
San Jose	991	11.7	0.012	2.53	−0.62*	0.79*	−1.53*
Austin	900	15.3	0.017	−1.34*	−0.79*	0.71*	−1.38*
Jacksonville	873	19.0	0.022	−0.39*	0.69	0.66*	−1.14*
San Francisco	864	15.4	0.018	−0.04	−0.48	0.65*	−0.64
Columbus	776	21.5	0.028	−0.40*	−0.36	0.31*	−0.71*
Fort Worth	820	17.2	0.021	−0.24	−0.81*	1.09*	−1.48*
Indianapolis	856	23.5	0.027	−0.35*	−0.87*	0.52*	−1.05*
Charlotte	802	15.6	0.019	−0.01	−0.72*	0.31	−0.68*
Seattle	688	12.7	0.018	−0.61	−0.97	0.54*	−0.43
Denver	678	8.1	0.012	0.09	−0.57*	0.55*	−0.32
Washington	672	14.2	0.021	0.38*	−0.24	−0.47*	0.93*

* $p < 0.001$

that, conditional on being in the sample at all, the rate of app use is independent of the demographic composition of the device's home location, and that this rate is only weakly correlated with park visitation.

Figure B.2 shows box plots of app use, represented as the number of "pings" (registered locations), as a function of the fraction of the device's home tract that is Black or Hispanic. The inner quartiles and whiskers (10th and 90th percentiles) are both consistent, from bin to bin. To show dependence to extreme outliers that might bias the distribution, Figure B.3 shows a lowess curve for the two box plots. The curves are, on the whole quite flat, except in cities with very small minority populations – Blacks in Denver, for instance. Focusing on tracts instead of users, Figure B.4 shows tracts' average park use against their median pings per device. I use median pings instead of mean pings to avoid the impact of a few very severe outlier devices in ping rate, that otherwise entirely wipe out any correlation. Using the median ping rate, most correlations remain weak: they exceed 0.3 only in Los Angeles (0.37), San Antonio (0.33), Jacksonville (0.43), and Fort Worth (0.37). Figure D.7 shows that the regression results are robust to a control for the ping rate.

TABLE B.4. This table displays the slope over intercept (β/α) for the regressions of Table ??. This presentation affords the simpler interpretation of a percentage change from a constant baseline.

	Linear Regressions: Slope / Intercept			
	Black	Hispanic	Educ.	Poverty
New York	0.23*	0.04	−0.11*	0.01
Los Angeles	−0.06	−0.56*	1.55*	−1.12*
Chicago	0.20*	−0.33*	0.11	−0.13
Houston	−0.09	−0.57*	1.36*	−1.13*
Phoenix	−0.57	−0.61*	1.43*	−1.03*
Philadelphia	0.29*	−0.45*	0.08	−0.09
San Antonio	0.51	−0.62*	1.68*	−1.39*
San Diego	−0.82	−0.56*	1.14*	−1.12*
Dallas	−0.16	−0.66*	1.40*	−1.26*
San Jose	2.74	−0.52*	1.16*	−1.39*
Austin	−1.22*	−0.63*	1.10*	−1.20*
Jacksonville	−0.35*	0.76	0.83*	−1.01*
San Francisco	−0.04	−0.46	1.04*	−0.62
Columbus	−0.36*	−0.36	0.35*	−0.63*
Fort Worth	−0.24	−0.65*	1.61*	−1.24*
Indianapolis	−0.33*	−0.84*	0.64*	−0.92*
Charlotte	−0.01	−0.66*	0.35	−0.64*
Seattle	−0.61	−0.94	0.85*	−0.43
Denver	0.09	−0.49*	0.73*	−0.31
Washington	0.47*	−0.23	−0.38*	1.05*

* $p < 0.001$

FIGURE B.1. Normalized device to population ratio for studied cities. Core business districts, airports and other transportation hubs, and 24-hour facilities tend to be over-represented, presumably due to travel and the use of nocturnal locations for residence assignment.

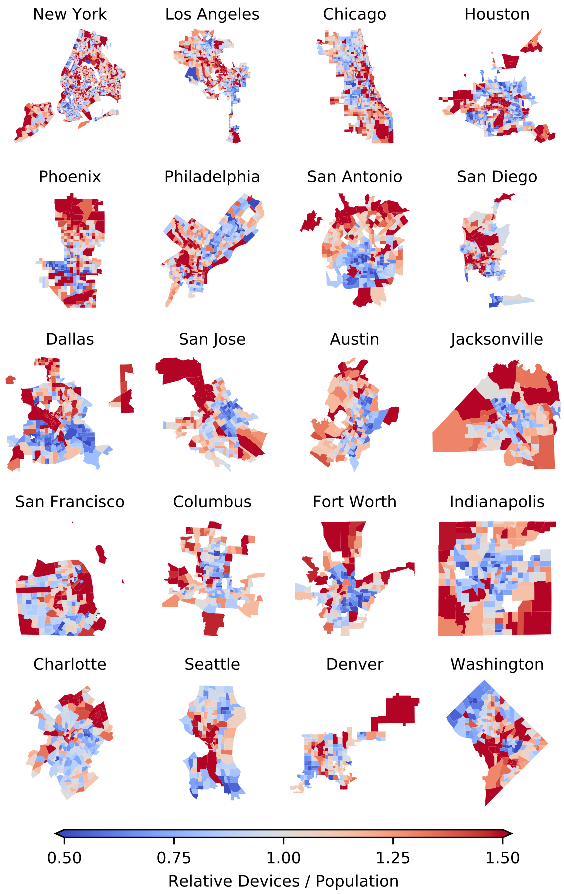
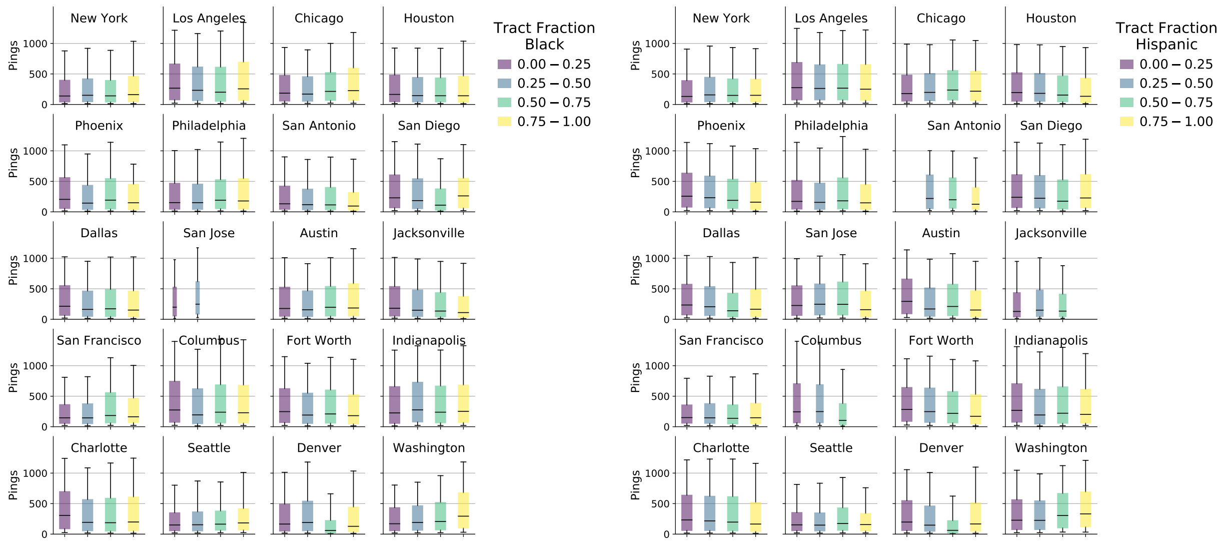


FIGURE B.2. Box plot of pings per person as a function of tract demographic composition. The boxes present the median and inner quartiles, and the whiskers show the 10th and 90th percentiles.



(A) Observation Rate by Fraction Black (B) Observation Rate by Fraction Hispanic

FIGURE B.3. Lowess curves for pings per person, as a function of tract ethnic and racial composition. This presentation is similar to Figure B.2 but is more sensitive to the tails of the distribution.

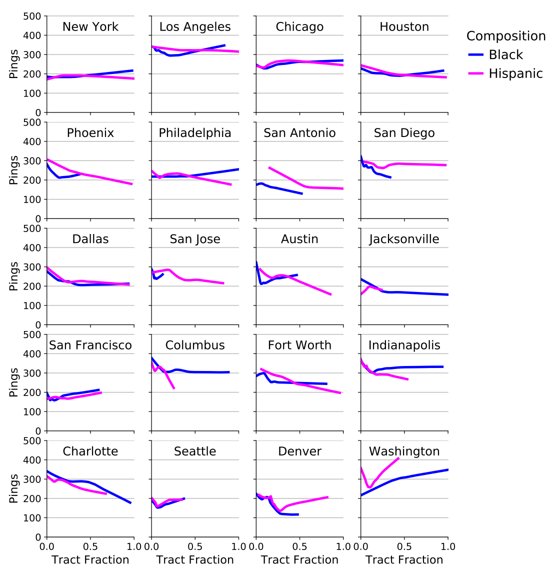
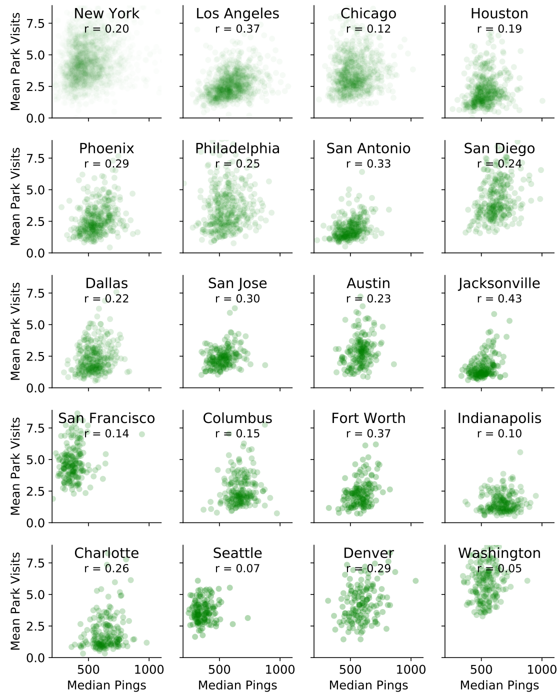


FIGURE B.4. Monthly park visits per person by tract, as a function of tracts' average pings (observations) per person, for the restricted sample.



Appendix C. Correlations Between Constructed Measures and Models

Table C.1 investigates the robustness of park usage variables on the sample restrictions. The nominal, restricted sample required devices to be observed in all three thirds of the month, three nights at home, and at least 100 times. Dropping this requirements makes for a sample three times larger (contrast Tables B.1 and B.3). Using these samples, the lowest correlation in park visitation by tract is in Washington DC, where it remains 0.84. The majority of correlations exceed 0.9. I also define two more-stringent samples: one tight, with 8 observed nights and 150 locations, and one ultra-tight, with 10 nights and 200 locations observed. In all save three cases, the correlations in park use to the nominal sample exceed 0.8. The three outliers are the ultra-tight samples for New York City (0.77), San Jose (0.78) and Washington DC (0.73). In short, the results are quite robust to significant changes in the sample composition.

Figure 4 of the main text presents the correlations between the baseline measurement of park use per deivce, and six spatial models, in the six largest cities in the United States. Table C.1 also shows those population-weighted correlations for the twenty largest cities. The table shows the maximal correlations achieved over a scan of the model parameter (gravity decay, buffer radius, and nearest park size).

Table C.1. Correlations at the Census tract level between measured use based on the nominal, restricted sample and alternative measures and models. For the "traditional" spatial models (distance, gravity, and buffer), the quoted value is the maximum across the scan of parameters as shown in Figure 4. To orient study, cells are shaded in a gradient from dark purple for correlations near zero, to light yellow for those approaching one.

	Measured Use Sample Selection			Traditional Spatial Maximized Correlation			GEE Poisson Fits Euclidean Distance			GEE Poisson Fits Network Distance		
	Loose	Tight	Ultra	Min. Dist.	Buffer	Gravity	Dist.	Log	FE	Dist.	Log	FE
New York	0.88	0.84	0.77	0.28	0.36	0.22	0.17	0.33	0.33	0.09	0.31	0.31
Los Angeles	0.90	0.88	0.81	0.47	0.48	0.39	0.57	0.54	0.59	0.43	0.43	0.49
Chicago	0.90	0.92	0.88	0.53	0.64	0.68	0.74	0.70	0.77	0.75	0.61	0.79
Houston	0.95	0.93	0.88	0.40	0.39	0.26	0.69	0.63	0.73	0.66	0.63	0.72
Phoenix	0.92	0.92	0.88	0.55	0.49	0.40	0.49	0.36	0.49	0.40	0.34	0.44
Philadelphia	0.91	0.92	0.87	0.42	0.36	0.54	0.72	0.67	0.73	0.68	0.64	0.71
San Antonio	0.93	0.88	0.83	0.33	0.44	0.39	0.48	0.48	0.54	0.45	0.47	0.50
San Diego	0.89	0.93	0.89	0.47	0.44	0.36	0.66	0.62	0.68	0.59	0.56	0.60
Dallas	0.94	0.94	0.90	0.29	0.54	0.51	0.60	0.58	0.64	0.63	0.60	0.68
San Jose	0.87	0.85	0.78	0.23	0.35	0.43	0.19	0.27	0.40	0.11	0.19	0.33
Austin	0.91	0.94	0.88	0.36	0.41	0.41	0.52	0.48	0.58	0.48	0.50	0.56
Jacksonville	0.97	0.96	0.94	0.28	0.52	0.46	0.51	0.42	0.55	0.41	0.41	0.47
San Francisco	0.93	0.90	0.86	0.54	0.69	0.62	0.34	0.55	0.55	0.31	0.44	0.51
Columbus	0.90	0.94	0.90	0.36	0.33	0.31	0.60	0.56	0.63	0.57	0.57	0.62
Fort Worth	0.94	0.97	0.94	0.27	0.34	0.20	0.41	0.42	0.50	0.40	0.44	0.48
Indianapolis	0.93	0.94	0.90	0.49	0.49	0.44	0.65	0.64	0.71	0.62	0.62	0.68
Charlotte	0.97	0.97	0.94	0.50	0.60	0.60	0.71	0.64	0.79	0.64	0.64	0.69
Seattle	0.87	0.87	0.81	0.35	0.38	0.37	0.24	0.28	0.33	0.16	0.27	0.27
Denver	0.89	0.90	0.84	0.26	0.18	0.02	0.33	0.37	0.34	0.32	0.36	0.36
Washington	0.84	0.84	0.73	0.44	0.44	0.33	0.44	0.48	0.48	0.45	0.48	0.48

Appendix D. Contrasting Potential and Realized Access with Alternate Spatial Models.

Table D presents the form of regressions used for the generalized estimating equations (GEE) spatial models. Tract-park pairs are considered out to a Euclidean distance of 10 km or a network distance of 15 km. The network requirement "applies" in particular, in New York City, whose boroughs are separated by water, and whose bridges and tunnels are often not accessible for pedestrians. (I exclude motorways and trunk roads from the "walking" network.) For this tract-level analysis, the difference between Euclidean and network distances has limited impact on the pseudo-R^2. Chicago is in fact an unusual case; in most cities the Euclidean distances are slightly more performant (see Table C.1).

Figure D.1 repeats the form of Figure 7, using Euclidean instead of network distances. Figure D.2 does the same, using a heteroskedastic and autocorrelation (HAC) robust OLS method, as implemented in pySAL. (Anselin and Rey, 2014) This analysis uses an adaptive triangular kernel with a number of neighbors taken as the cube root of the number of tracts. Figure D.3 uses distance fixed effects in the GEE approach to spatial access. Figure D.4 uses the loose, unrestricted sample of users: without the requirement of at least 100 locations over the month, three nights, and all three thirds of the month. Figures D.5 and D.6 go in the opposite direction – they are "tight" and "ultra-tight," requiring users to be observed, respectively, 8 or 10 nights in the month and 150 or 200 times. Results are largely unaffected by these changes among spatial methods or regression methods. However, with the unrestricted sample, the distance fit is less effective as a control (its parameter estimates are smaller), and the dependence of use on race and ethnicity is larger. This trend does not continue, however, to the tight and ultra tight samples.

Finally, Figure D.7 again repeats the baseline model, but with a control for tracts' median ping rate. The aim of this paper is to evaluate park use, not phone use. Results should not be sensitive to phone use, and empirically they are not.

Table D. Park use in Chicago modeled at the level of Census tract-park pairs, using generalized estimating equations, grouped by Census tract and using and exchangeable covariance structure. Observations are weighted by Census tract population. All parameters shown are significant at the 0.001 level.

	Dep. Variable: Tract to Park Visit Rate					
Log Area	0.55	0.62	0.60	0.58	0.57	0.37
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
Log Eucl. Distance	-0.88					
	(0.03)					
Euclidean Distance		-0.86				
		(0.02)				
Network Distance			-0.75			
			(0.02)			
Distance FE				Eucl.	Net.	Net.
Clusters	Tracts	Tracts	Tracts	Tracts	Tracts	Parks
N Clusters	793	793	793	793	793	1201
Pseudo-R^2	0.53	0.65	0.66	0.63	0.65	0.49

Figure D.1. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that it uses Euclidean instead of network distances to parks, for the spatial models.

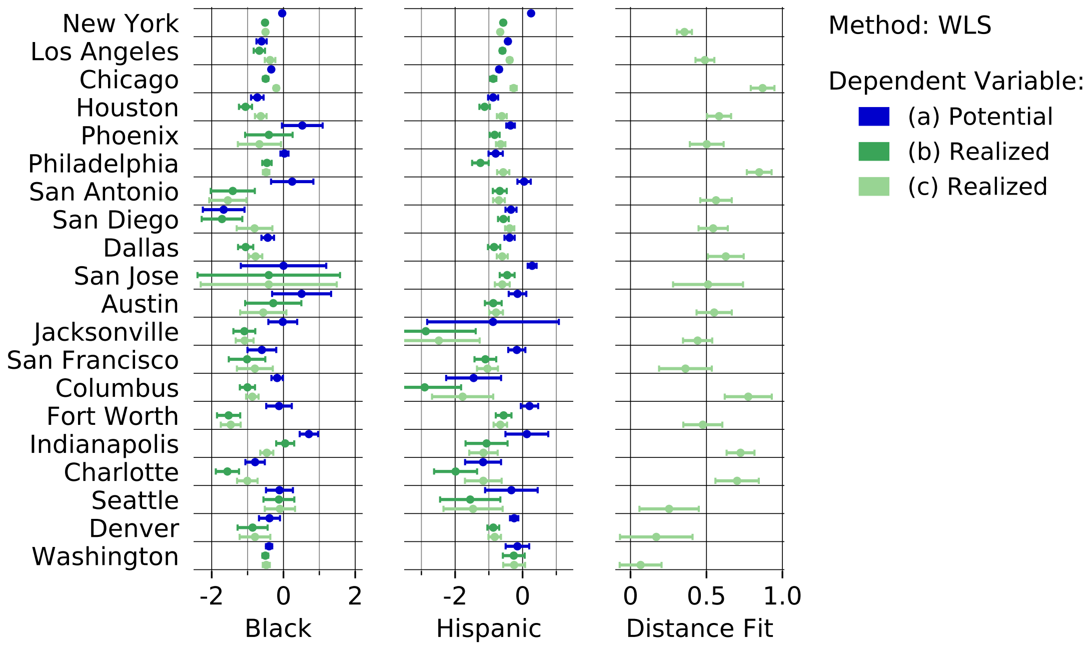


Figure D.2. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that it uses unweighted OLS with Heteroskedastic and Autocorrelation robust standard errors, implemented using pySAL. (Anselin and Rey, 2014)

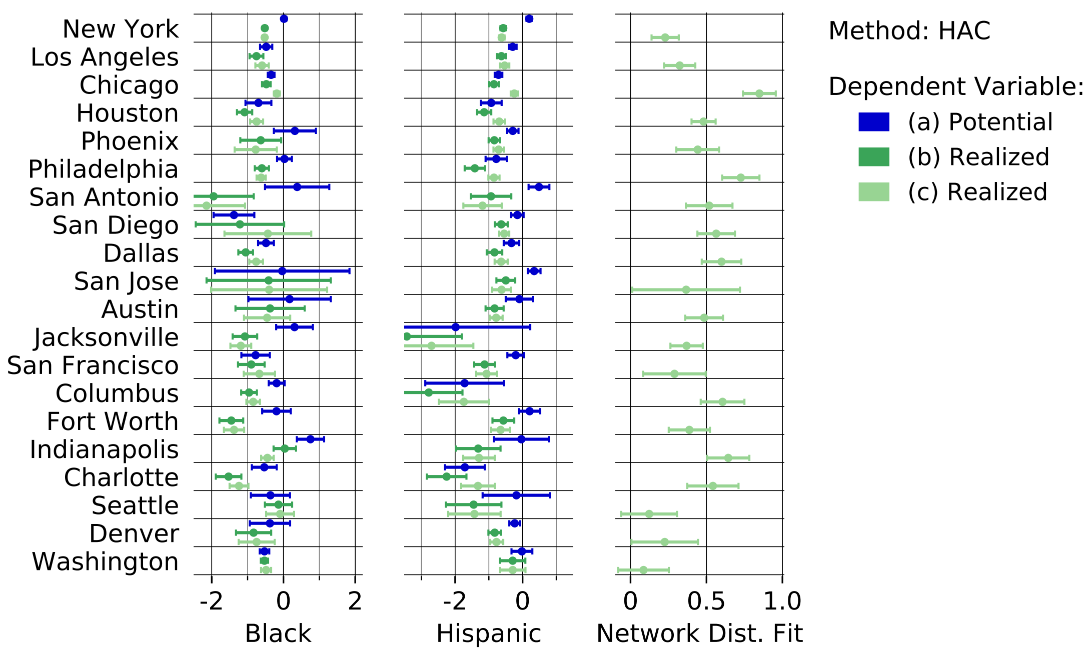


Figure D.3. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that network distance fixed effects are used instead of network distance, in the GEE model.

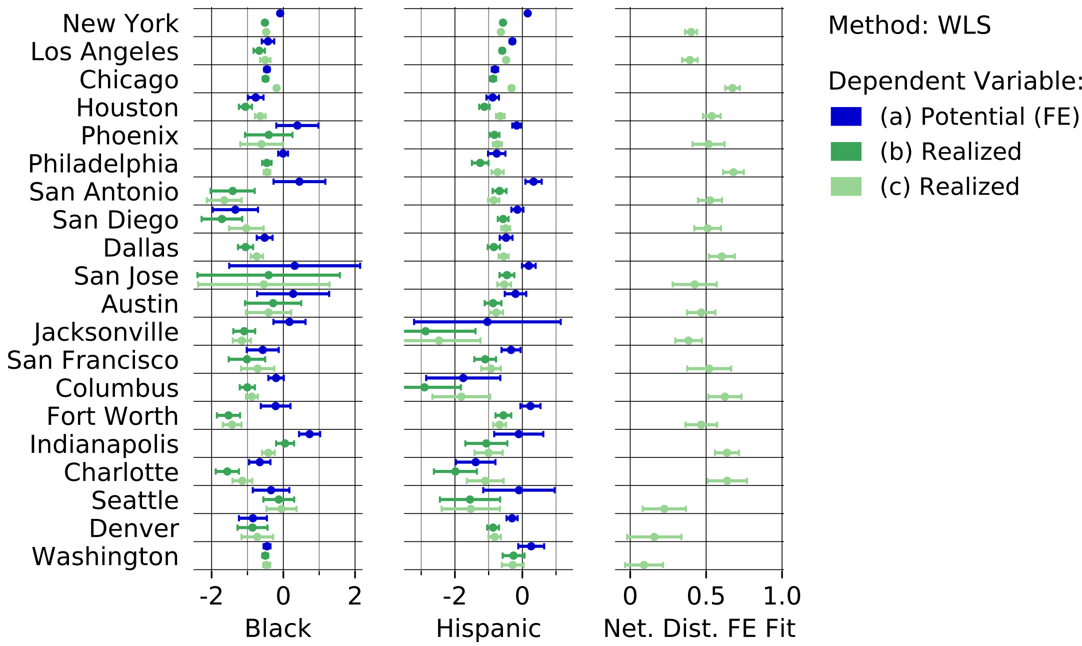


Figure D.4. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that park use is evaluated using the unrestricted sample: without the requirement of at least 100 locations over the month, three nights, and all three thirds of the month.

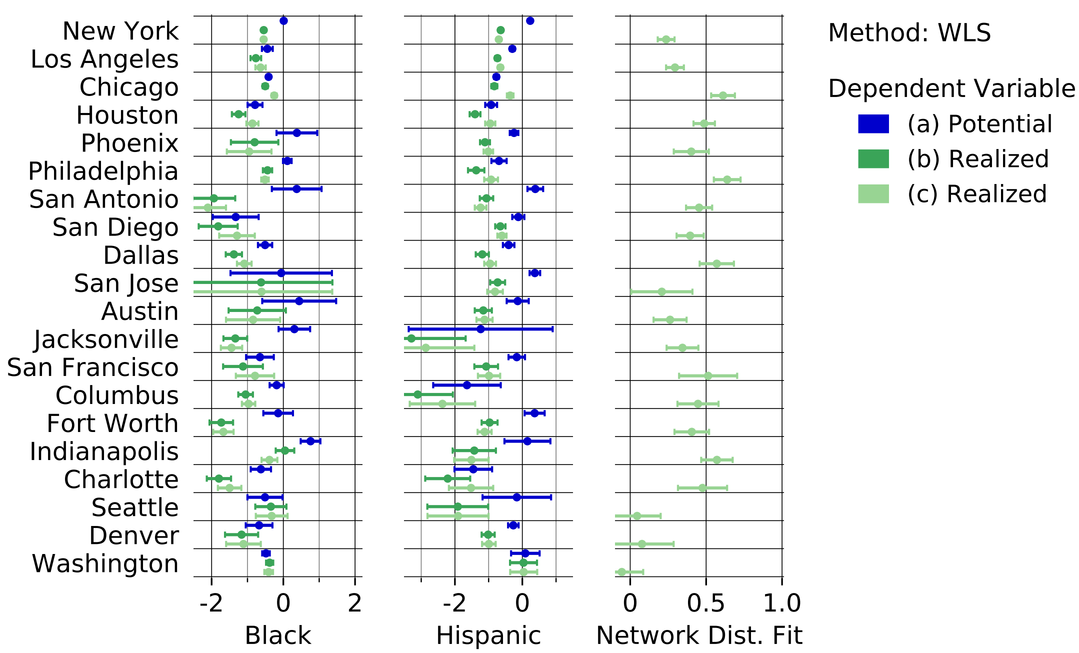


FIGURE D.5. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that park use is evaluated using with a tight sample: requiring at least 8 nights at home and 150 locations over the month (as well as in all three thirds of the month).

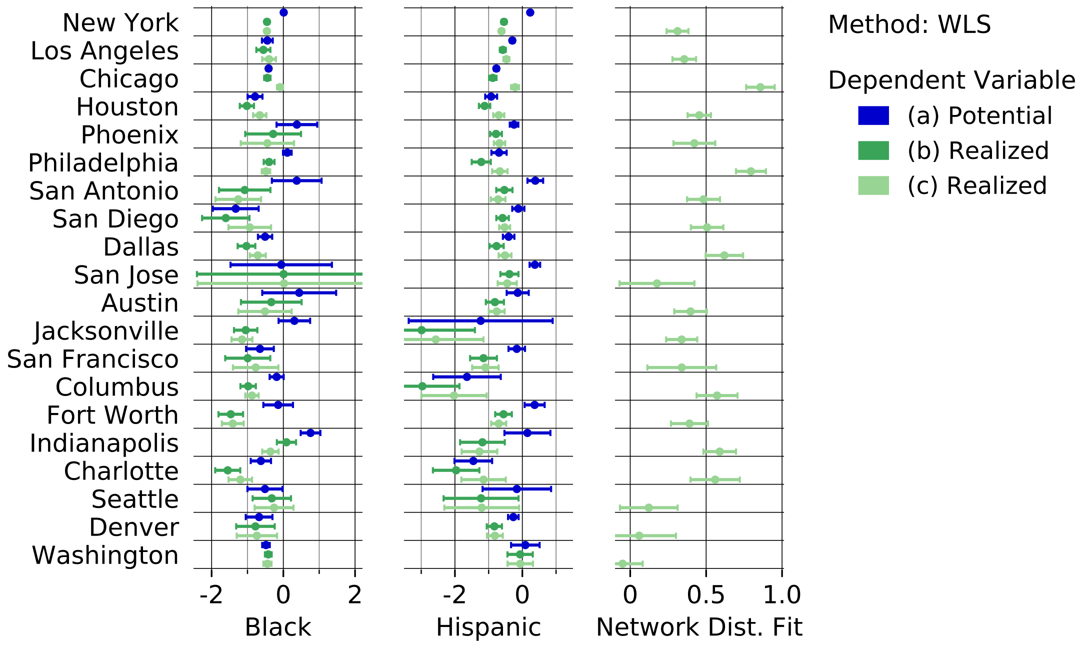


FIGURE D.6. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that park use is evaluated using with a "ultra" tight sample: requiring at least 10 nights at home and 200 locations over the month (as well as in all three thirds of the month).

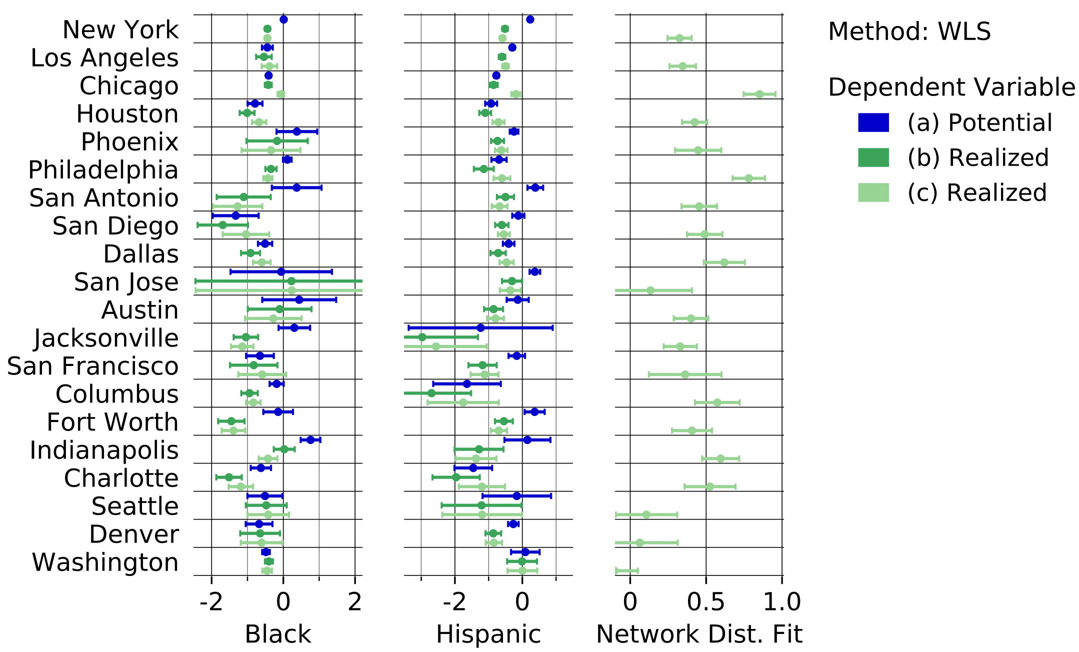
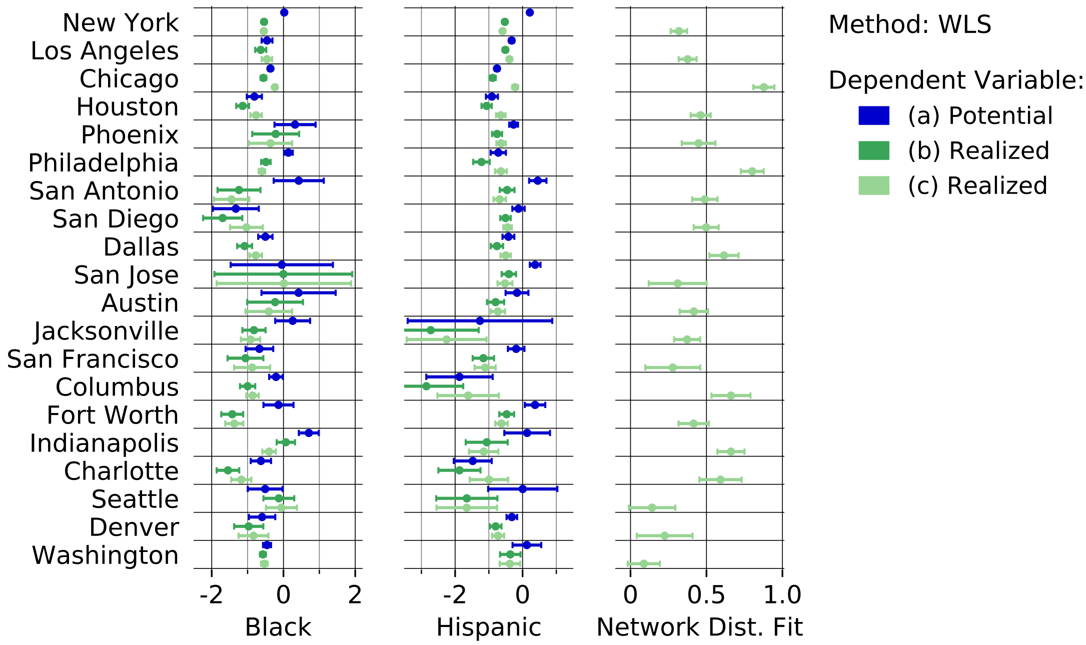


FIGURE D.7. Point estimates (circles) and 95% confidence limits are presented for three models across twenty cities. The Figure is the same as Figure 7, except that it also controls for the median pings per device, in each tract. This has a minimal impact on parameter estimates.



References for Supplementary Materials

Anselin, L. and Rey, S. J., *Modern Spatial Econometrics in Practice: A Guide to GeoDa, GeoDaSpace and PySAL*, (2014). GeoDa Press LLC, Chicago.

Anselin, L. and Williams, S., "Digital neighborhoods," *Journal of Urbanism,* **9**:4, 305-328 (2016). DOI: 10.1080/17549175.2015.1080752.

Astell-Burt, T., Feng, X., Mavoa, S., Badland, H. M., and Giles-Corti, B., "Do low-income neighbourhoods have the least green space? A cross-sectional study of Australias most populous cities.," *BMC Public Health,* **14**:292, 1-12 (2014). DOI: 10.1186/1471-2458-14-292.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P., *et al.*, "Who benefits from access to green space? A case study from Sheffield, UK," *Landscape and Urban Planning,* **83**:2, 187-195 (2007). DOI: 10.1016/j.landurbplan.2007.04.004.

Bedimo-Rung, A. L., Mowen, A. J., and Cohen, D. A., "The significance of parks to physical activity and public health: A conceptual model," *American Journal of Preventive Medicine,* **28**:28, (2005). DOI: 10.1016/j.amepre.2004.10.024.

Berman, M. G., Jonides, J., and Kaplan, S., "The Cognitive Benefits of Interacting With Nature," *Psychological Science,* **19**:12, 1207-1212 (2008). DOI: 10.1111/j.1467-9280.2008.02225.x.

Boone, C. G., Buckley, G. L., Grove, J. M., and Sister, C., "Parks and People: An Environmental Justice Inquiry in Baltimore, Maryland," *Annals of the Association of American Geographers,* **99**:4, 767-787 (2009). DOI: 10.1080/00045600903102949.

Cohen, D. A., Ashwood, J. S., Scott, M. M., Overton, A., Evenson, K. R., Staten, L. K., *et al.*, "Public Parks and Physical Activity Among Adolescent Girls," *Pediatrics,* **118**:5, 1381-1389 (2006). DOI: 10.1542/peds.2006-1226.

Cohen, D. A., McKenzie, T. L., Sehgal, A., Williamson, S., Golinelli, D., and Lurie, N., "Contribution of Public Parks to Physical Activity," *American Journal of Public Health,* **97**:3, 509-514 (2007). DOI: 10.2105/AJPH.2005.072447.

Dadvand, P., Bartoll, X., Basagaña, X., Dalmau-Bueno, A., Martinez, D., Ambros, A., *et al.*, "Green spaces and General Health: Roles of mental health status, social support, and physical activity," *Environment International,* **91**, 161-167 (2016). DOI: 10.1016/j.envint.2016.02.029.

Dai, D., "Racial/ethnic and socioeconomic disparities in urban green space accessibility: Where to intervene?" *Landscape and Urban Planning,* **102**:4, 234-244 (2011). DOI: 10.1016/j.landurbplan.2011.05.002.

Engemann, K., Pedersen, C. B., Arge, L., Tsirogiannis, C., Mortensen, P. B., and Svenning, J.-C., "Residential green space in childhood is associated with lower risk of psychiatric disorders from adolescence into adulthood," *Proceedings of the National Academy of Sciences,* **116**:11, 5188-5193 (2019). DOI: 10.1073/pnas.1807504116.

Giles-Corti, B. and Donovan, R. J., "Relative Influences of Individual, Social Environmental, and Physical Environmental Correlates of Walking," *American Journal of Public Health,* **93**:9, 1583-1589 (2003). DOI: 10.2105/AJPH.93.9.1583.

Gordon-Larsen, P., Nelson, M. C., Page, P., and Popkin, B. M., "Inequality in the Built Environment Underlies Key Health Disparities in Physical Activity and Obesity," *Pediatrics,* **117**:2, 417-424 (2006). DOI: 10.1542/peds.2005-0058.

Hartig, T., Evans, G. W., Jamner, L. D., Davis, D. S., and Gärling, T., "Tracking restoration in natural and urban field settings," *Journal of Environmental Psychology,* **23**:2, 109-123 (2003). DOI: 10.1016/S0272-4944(02)00109-3.

Higgs, G., Fry, R., and Langford, M., "Investigating the Implications of Using Alternative GIS-Based Techniques to Measure Accessibility to Green Space," *Environment and Planning B: Planning and Design,* **39**:2, 326-343 (2012). DOI: 10.1068/b37130.

Hillsdon, M., Panter, J., Foster, C., and Jones, A., "The relationship between access and quality of urban green space with population physical activity," *Public Health,* **120**:12, 1127-1132 (2006). DOI: 10.1016/j.puhe.2006.10.007.

Hoehner, C. M., Ramirez, L. K. B., Elliott, M. B., Handy, S. L., and Brownson, R. C., "Perceived and objective environmental measures and physical activity among urban adults," *American Journal of Preventive Medicine,* **28**:2, Supplement 2, 105-116 (2005). DOI: 10.1016/j.amepre.2004.10.023.

Hughey, S. M., Walsemann, K. M., Child, S., Powers, A., Reed, J. A., and Kaczynski, A. T., "Using an environmental justice approach to examine the relationships between park availability and quality indicators, neighborhood disadvantage, and racial/ethnic composition," *Landscape and Urban Planning,* **148**, 159-169 (2016). DOI: https://doi.org/10.1016/j.landurbplan.2015.12.016.

Hume, C., Salmon, J., and Ball, K., "Associations of Children's Perceived Neighborhood Environments with Walking and Physical Activity," *American Journal of Health Promotion,* **21**:3, 201-207 (2007). DOI: 10.4278/0890-1171-21.3.201.

Humpel, N., Owen, N., Iverson, D., Leslie, E., and Bauman, A., "Perceived environment attributes, residential location, and walking for particular purposes," *American Journal of Preventive Medicine,* **26**:2, (February, 2004). DOI: 10.1016/j.amepre.2003.10.005.

Huston, S. L., Evenson, K. R., Bors, P., and Gizlice, Z., "Neighborhood Environment, Access to Places for Activity, and Leisure-Time Physical Activity in a Diverse North Carolina Population," *American Journal of Health*

Promotion, **18**:1, 58-69 (2003). DOI: 10.4278/0890-1171-18.1.58.

James, P., Hart, J. E., Banay, R. F., and Laden, F., "Exposure to Greenness and Mortality in a Nationwide Prospective Cohort Study of Women," *Environmental Health Perspectives,* **124**:9, 1344-1352 (2016). DOI: 10.1289/ehp.1510363.

Jilcott, S. B., Evenson, K. R., Laraia, B. A., and Ammerman, A. S., "Association between physical activity and proximity to physical activity resources among low-income, midlife women.," *Preventing chronic disease,* **4**:1, A04 (January, 2007). PMC: PMC1832127

Lackey, K. J. and Kaczynski, A. T., "Correspondence of perceived vs. objective proximity to parks and their relationship to park-based physical activity," *International Journal of Behavioral Nutrition and Physical Activity,* **6**:1, 53 (Aug 11, 2009). DOI: 10.1186/1479-5868-6-53.

Leyden, K. M., "Social Capital and the Built Environment: The Importance of Walkable Neighborhoods," *American Journal of Public Health,* **93**:9, 1546-1551 (2003). DOI: 10.2105/AJPH.93.9.1546.

Lindsey, G., Maraj, M., and Kuan, S., "Access, Equity, and Urban Greenways: An Exploratory Investigation," *The Professional Geographer,* **53**:3, 332-346 (2001). DOI: 10.1111/0033-0124.00288.

Logan, T., Williams, T., Nisbet, A., Liberman, K., Zuo, C., and Guikema, S., "Evaluating urban accessibility: leveraging open-source data and analytics to overcome existing limitations," *Environment and Planning B: Urban Analytics and City Science,* , 2399808317736528 (2017). DOI: 10.1177/2399808317736528.

Maas, J., Verheij, R. A., Groenewegen, P. P., de Vries, S., and Spreeuwenberg, P., "Green space, urbanity, and health: how strong is the relation?" *Journal of Epidemiology & Community Health,* **60**:7, 587-592 (2006). DOI: 10.1136/jech.2005.043125.

Maas, J., van Dillen, S. M. E., Verheij, R. A., and Groenewegen, P. P., "Social contacts as a possible mechanism behind the relation between green space and health," *Health & Place,* **15**:2, 586-595 (2009). DOI: 10.1016/j.healthplace.2008.09.006.

Maddison, R., Jiang, Y., Hoorn, S. V., Mhurchu, C. N., Exeter, D., and Utter, J., "Perceived Versus Actual Distance to Local Physical-Activity Facilities: Does It Really Matter?" *Journal of Physical Activity and Health,* **7**:3, 323-332 (2010).

Nicholls, S., "Measuring the accessibility and equity of public parks: a case study using GIS," *Managing Leisure,* **6**:4, 201-219 (2001). DOI: 10.1080/13606710110084651.

Nielsen, T. S. and Hansen, K. B., "Do green areas affect health? Results from a Danish survey on the use of green areas and health indicators," *Health & Place,* **13**:4, 839-850 (2007). DOI: 10.1016/j.healthplace.2007.02.001.

Powell, K. E., Martin, L. M., and Chowdhury, P. P., "Places to Walk: Convenience and Regular Physical Activity," *American Journal of Public Health,* **93**:9, 1519-1521 (2003). DOI: 10.2105/AJPH.93.9.1519.

Reid, C. E., Kubzansky, L. D., Li, J., Shmool, J. L., and Clougherty, J. E., "It's not easy assessing greenness: A comparison of NDVI datasets and neighborhood types and their associations with self-rated health in New York City," *Health & Place,* **54**, 92-101 (2018). DOI: 10.1016/j.healthplace.2018.09.005.

Reklaitiene, R., Grazuleviciene, R., Dedele, A., Virviciute, D., Vensloviene, J., Tamosiunas, A., *et al.*, "The relationship of green space, depressive symptoms and perceived general health in urban population," *Scandinavian Journal of Public Health,* **42**:7, 669-676 (2014). DOI: 10.1177/1403494814544494.

Reyes, M., Pez, A., and Morency, C., "Walking accessibility to urban parks by children: A case study of Montreal," *Landscape and Urban Planning,* **125**, 38-47 (2014). DOI: 10.1016/j.landurbplan.2014.02.002.

Rigolon, A. and Flohr, T. L., "Access to Parks for Youth as an Environmental Justice Issue: Access Inequalities and Possible Solutions.," *Buildings,* **4**:2, 69-94 (2014). DOI: 10.3390/buildings4020069.

Singh, G. K., Siahpush, M., and Kogan, M. D., "Neighborhood Socioeconomic Conditions, Built Environments, And Childhood Obesity," *Health Affairs,* **29**:3, 503-512 (2010). DOI: 10.1377/hlthaff.2009.0730.

Sister, C., Wolch, J., and Wilson, J., "Got green? addressing environmental justice in park provision.," *GeoJournal,* **75**, 229242 (2010). DOI: 10.1007/s10708-009-9303-8.

Smoyer-Tomic, K. E., Hewko, J. N., and Hodgson, M. J., "Spatial accessibility and equity of playgrounds in Edmonton, Canada," *The Canadian Geographer / Le Gographe canadien,* **48**:3, 287-302 (2004). DOI: 10.1111/j.0008-3658.2004.00061.x.

Takano, T., Nakamura, K., and Watanabe, M., "Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces," *Journal of Epidemiology & Community Health,* **56**:12, 913-918 (2002). DOI: 10.1136/jech.56.12.913.

Talen, E. and Anselin, L., "Assessing Spatial Equity: An Evaluation of Measures of Accessibility to Public Playgrounds," *Environment and Planning A: Economy and Space,* **30**:4, 595-613 (1998). DOI: 10.1068/a300595.

Talen, E., "Visualizing Fairness: Equity Maps for Planners," *Journal of the American Planning Association,* **64**:1, 22-38 (1998). DOI: 10.1080/01944369808975954.

Tan, P. Y. and Samsudin, R., "Effects of spatial scale on assessment of spatial equity of urban park provision," *Landscape and Urban Planning,* **158**, 139-154 (2017). DOI: 10.1016/j.landurbplan.2016.11.001.

Thompson, C. W., Roe, J., Aspinall, P., Mitchell, R., Clow, A., and Miller, D., "More green space is linked to less stress in deprived communities: Evidence from salivary cortisol patterns," *Landscape and Urban Planning,* **105**:3, 221-229 (2012). DOI: 10.1016/j.landurbplan.2011.12.015.

Troped, P. J., Saunders, R. P., Pate, R. R., Reininger, B., Ureda, J. R., and Thompson, S. J., "Associations between Self-Reported and Objective Physical Environmental Factors and Use of a Community Rail-Trail," *Preventive Medicine,* **32**:2, 191-200 (2001). DOI: 10.1006/pmed.2000.0788.

Ulmer, J. M., Wolf, K. L., Backman, D. R., Tretheway, R. L., Blain, C. J., ONeil-Dunne, J. P., *et al.*, "Multiple health benefits of urban tree canopy: The mounting evidence for a green prescription," *Health & Place,* **42**, 54-62 (2016). DOI: 10.1016/j.healthplace.2016.08.011.

VAUGHAN, K. B., KACZYNSKI, A. T., WILHELM STANIS, S. A., BESENYI, G. M., BERGSTROM, R., AND HEINRICH, K. M., "Exploring the Distribution of Park Availability, Features, and Quality Across Kansas City, Missouri by Income and Race/Ethnicity: an Environmental Justice Investigation," *Annals of Behavioral Medicine,* **45**:S1, S28-S38 (01, 2013). DOI: 10.1007/s12160-012-9425-y.

DE VRIES, S., VAN DILLEN, S. M. E., GROENEWEGEN, P. P., AND SPREEUWENBERG, P., "Streetscape greenery and health: Stress, social cohesion and physical activity as mediators," *Social Science & Medicine,* **94**, 26-33 (2013). DOI: 10.1016/j.socscimed.2013.06.030.

WEN, M., ZHANG, X., HARRIS, C. D., HOLT, J. B., AND CROFT, J. B., "Spatial Disparities in the Distribution of Parks and Green Spaces in the USA," *Annals of Behavioral Medicine,* **45**:S1, S18-S27 (01, 2013). DOI: 10.1007/s12160-012-9426-x.

WITTEN, K., EXETER, D., AND FIELD, A., "The Quality of Urban Environments: Mapping Variation in Access to Community Resources," *Urban Studies,* **40**:1, 161-177 (2003). DOI: 10.1080/00420980220080221.

WOLCH, J., WILSON, J. P., AND FEHRENBACH, J., "Parks and Park Funding in Los Angeles: An Equity-Mapping Analysis," *Urban Geography,* **26**:1, 4-35 (2005). DOI: 10.2747/0272-3638.26.1.4.