Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

journal homepage: www.elsevier.com/locate/ceus

Desirable streets: Using deviations in pedestrian trajectories to measure the value of the built environment

Arianna Salazar Miranda[a,*], Zhuangyuan Fan[a], Fabio Duarte[a,b], Carlo Ratti[a]

[a] Senseable City Lab, Department of Urban Studies and Planning, Massachusetts Institute of Technology, United States
[b] Pontifícia Universidade Católica do Paraná, Curitiba, Brazil

ARTICLE INFO

Keywords:
Built environment
Computer vision
Deep learning
Pedestrian choice

ABSTRACT

The experience of walking through a city is influenced by amenities and the visual qualities of its built environment. This paper uses thousands of pedestrian trajectories obtained from GPS signals to construct a desirability index for streets in Boston. We create the index by comparing the actual paths taken by pedestrians with the shortest path between any origin-destination pairs. The index captures pedestrians' willingness to deviate from their shortest path and provides a measure of the scenic and experience value provided by different parts of the city. We then use computer vision techniques combined with georeferenced data to measure the built environment of streets. We show that desirable streets have better access to public amenities such as parks, sidewalks, and urban furniture. They are also sinuous, visually enclosed, have less complex facades, and have more diverse business establishments. These results further our understanding of the value that the built environment brings to pedestrians, enhancing our capacity to design more lively and functional environments.

1. Introduction

The importance of streets as a critical public space for social exchange is widely recognized by urban planners and cities alike. Urban planners have proposed ambitious strategies to promote pedestrian-friendly environments to foster active transportation alternatives (Calthorpe, 1993; Duany, Plater-Zyberk, Krieger, & Lennertz, 1991). Cities are investing vast amounts of resources in improving the quality of their streets, in part motivated by the conviction that well-designed streets will improve their citizens' quality of life. This growing demand for street upgrading and renewal warrants further study on what makes streets desirable to pedestrians. A more thorough understanding of how people use streets will allow urban planners to design guidelines and interventions that lead to streets being more lively and pedestrian-friendly.

An important step towards better street design is to understand *where people walk and whether these choices are influenced by the built environment*. To address these questions, the literature on pedestrian choice has relied on modeling approaches that predict mobility based on the street network (Hillier, Penn, Hanson, Grajewski, & Xu, 1993; Law & Traunmueller, 2017), estimates of individual preferences (Ben-Akiva & Bierlaire, 2006), or has relied on self-reported walking behavior to understand the reasoning and actions of pedestrians (Pikora et al., 2006). In addition, the literature linking pedestrian choices to characteristics of the built environment has relied on trained volunteers and are typically limited to small geographic contexts, such as neighborhood blocks (Ewing, Handy, Brownson, Clemente, & Winston, 2006; Purciel & Marrone, 2006). More recent work has incorporated advanced data collection methods to shed light on the relationship between pedestrian walking and the built environment. Examples include using GPS data to track pedestrian movements and using 2D and 3D GIS and machine learning techniques to construct more comprehensive measures of the built environment (Yin, 2017; Yin & Wang, 2016).

In this paper, we address these questions by adopting a complementary approach that leverages GPS data to study how people move in the city and combine it with computer vision techniques to measure the built environment's features that might motivate these choices. Besides the novel data collection methods, a crucial conceptual difference in our exercise is that we emphasize the fact that not all pedestrians take their shortest route but instead systematically deviate from it. The aggregation of these individual deviations provides information on the built environment's features that pedestrians deem desirable. Moreover, these deviations allow us to construct a measure of street desirability that is independent of the potential number of people (e.g., simulated walk scores) who would use that street segment for reasons unrelated to its location in the city or its position on the street network. That is, our

* Corresponding author.
E-mail address: ariana@mit.edu (A. Salazar Miranda).

https://doi.org/10.1016/j.compenvurbsys.2020.101563
Received 19 March 2020; Received in revised form 20 October 2020; Accepted 22 October 2020

measure accounts for the fact that some street segments are heavily used by pedestrians just because they happen to be centrally located or are very well connected.

The first component of our analysis is to construct a desirability index for every street in the city of Boston that summarizes pedestrian preferences. To measure the desirability of streets, we use a revealed preference approach that considers whether pedestrians are willing to endure longer commutes to enjoy desirable places and avoid undesirable ones. We use the term desirability to describe street segments that pedestrians systematically use when they deviate from the shortest path, regardless of the origin or destination of their journey.[1] This approach is grounded on the common assumption in urban models that the shortest route can predict routing behavior because it maximizes the utility of a trip (Meyer, 2016). Thus a pedestrian's decision to walk longer (or deviate from the shortest route) translates into an additional cost that must be compensated by route-specific factors that make the route attractive and, therefore, worth traveling. Conceptually, the desirability index we propose provides a revealed preference measure of the scenic and experience value provided by different parts of the city. The measure aims to capture those unplanned journeys in which pedestrians "let themselves be drawn by the attractions of the terrain and the encounters they find there" (Debord, 1958), leading them to deviate from their shortest path.[2] Besides capturing a meaningful concept, the desirability index also provides practical value for urban planners. For example, it can be used to identify areas with high potential to attract pedestrians that currently have low desirability and would be good candidates to intervene. Likewise, the desirability index can be used to identify areas with high desirability that can teach practitioners lessons about successful urban design interventions. Finally, desirability maps can also be produced over time to track how changes to the built environment might affect pedestrians' trajectories throughout a city.

The second component of our analysis is to measure a wide range of built environment characteristics that urban planners have long believed to be associated with pedestrians' experience while walking (Jacobs, 1993; Jacobs & Appleyard, 1987; Krier, 1984; Lynch, 1960). To obtain the location of business establishments, cultural and recreation facilities, route-specific services, and parks, we use georeferenced data collected from OpenStreetMap. We also process Google Street View images using computer vision techniques to construct urban design measures, including urban furniture, sidewalks, facade complexity, and visual enclosure. Finally, we use the street network to compute geometric measures, such as the average slope and sinuosity for all street segments.

We present three main results. First, we show that desirable streets are geographically dispersed, and that people systematically deviate from the shortest path to frequent them. This pattern is consistent across different types of trips, including trips taking place at different times of the day (morning, noon, afternoon, and night trips) and trips of varying length (defined by the total length of the shortest path). Second, by comparing street segments with varying levels of desirability and different built environment characteristics, we show that more desirable streets are associated with greater provision of public amenities. In particular, desirable streets tend to have neighboring parks and more urban furniture. In addition, desirable streets tend to be shorter, more sinuous, more visually enclosed, and have more homogeneous facades. Finally, we provide evidence showing that the number of retail establishments and the variety of businesses also contribute to a street's desirability.

This paper makes four contributions to the literature. The first is methodological in the construction of the desirability index. Our methodology of exploiting deviations from the shortest path in large-scale GPS data, complements existing papers that rely on virtual audits or surveys, which are expanded using machine learning algorithms. Researchers have relied on crowd-sourcing to quantify visual appearance, with the help of trained experts (Kelly, Wilson, Baker, Miller, & Schootman, 2013), surveys (Saiz, Salazar, & Bernard, 2018), or online crowds (Naik, Raskar, & Hidalgo, 2016; Salesses, Schechtner, & Hidalgo, 2013). Quercia, O'Hare, and Cramer (2014), for example, use image processing techniques to identify the visual cues that may be associated with a street in London being perceived as beautiful, quiet, or happy. Their work identifies objective attributes that are systematically associated with different subjective experiences—as measured by an accessory survey. Ordonez and Berg (2014) also show that computer vision techniques can use physical attributes of the built environment to predict perceptual characteristics of places. They use perception surveys from one city to train an algorithm used to predict perceptions and judgments at a larger scale and across cities. Our approach differs from the work mentioned above in that it incorporates the revealed preferences of thousands of pedestrian trips instead of relying on surveys, which are usually constrained to smaller geographical areas and can suffer from response bias. Moreover, our approach circumvents limitations of scale and can be easily implemented in multiple areas in a comparable way, as long as GPS trajectories are available. More broadly, our paper also contributes to the growing literature focused on using automatically-generated data produced by GPS technologies to improve our measurement and understanding of how people use city space (González, Hidalgo, & Barabási, 2008; Jiang, Ferreira, & González, 2012; Ratti, Frenchman, Pulselli, & Williams, 2006; Sevtsuk & Ratti, 2010; Shoval, 2008).

The second contribution of this paper is to shed light on some of the features of the built environment that make street segments more desirable to pedestrians. Most of this literature relies on the socio-ecological theory of human behavior, which suggests that environmental factors can influence the individual behavior of people in cities (Sallis, Floyd, Rodríguez, & Saelens, 2012; Wolff, 1973). Work along this vein has focused on identifying the types of environments, both at the macro and micro scale, that are preferred by pedestrians using travel diary data or simulations. Overall, the documented relationships between the built environment and walking are stronger for macro-level characteristics such as density (Cervero & Radisch, 1996), business clusters (Sevtsuk, 2014), block proportions (Sevtsuk, Kalvo, & Ekmekci, 2016), and the properties of street networks (Hillier et al., 1993; Law & Traunmueller, 2017), including multiple centrality assessments (Porta, Crucitti, & Latora, 2008), and function connectivity (Shen & Karimi, 2016). Complementary to our findings, access to open space (Giles-Corti et al., 2005) and compact development patterns (Giles-Corti & Donovan, 2003; Hess, Moudon, Snyder, & Stanilov, 1999; Talen & Koschinsky, 2013), have been associated with positive walking levels and healthier lifestyles. Evidence of the effect of land uses on walking is mixed. Some studies find that more mixed uses are negatively associated with walking (Ewing & Handy, 2009), while others have found a positive association (Moudon, Hess, Snyder, & Stanilov, 1997). At the micro-scale, studies have focused on understanding which characteristics of the built environment are conducive to more walkable environments. Evidence for the micro-scale design features, such as enclosure and transparency, remains mixed. However, at the micro scale, there seems to be a consensus that the attractiveness of the built environment is associated with walking (Saelens & Handy, 2008a, 2008b). For instance, Vanky, Verma, Courtney, Santi, and Ratti (2017) show that a 10% increase in the number of retail businesses is associated with a 1.2% and 0.85%

[1] We use the term "desirability" as a descriptive label to indicate when a street is visited systematically by many pedestrians who deviate from their shortest path. This term describes streets that pedestrian crowds find "desirable" and is therefore informative of the quality of the built environment offered. Since we use the term as a label, our approach does not compete with other efforts that make normative statements about how cities and streets should be (see, for example, (Jacobs, 1961; Montgomery, 1995))

[2] Our approach is based on the premise that the aggregation of individual choices provides information on the features of the environment that pedestrians deem desirable. While our methodology is designed to identify these commonalities, it abstracts from the specific cognitive and psychological processes determining each individual's decision to detour.

increase in the number of pedestrian trips in Boston and San Francisco, respectively. These studies explain observed pedestrian activity—measured by the total flow of pedestrians— as a function of global and micro characteristics of the built environment. Instead, our methodology zooms in to pedestrian deviations from their shortest path—rather than the total flow. This approach allows us to remove structural and macro factors affecting observed pedestrian activity (like density, street network structure, and others) and isolate the factors that make a given route more attractive than an alternative one. Finally, we enrich the existing measures of the built environment by incorporating deep learning techniques to construct objective urban design metrics that characterize the built environment of streets. This approach distinguishes from previous efforts that have typically relied on subjective perceptions of the built environment collected via surveys and that do not focus on objectively measured built environment characteristics (Saelens & Handy, 2008a, 2008b). Moreover, the scene understanding algorithm we use resembles human eye-level perception, which provides a more direct reflection of pedestrians' visual reality.[3]

Third, our paper complements existing literature on pedestrian behavior. This literature shows that pedestrians deviate from the shortest path for multiple reasons, including perception (Manley, Addison, & Cheng, 2015; Ye, Zeng, Shen, Zhang, & Lu, 2019), visibility (Natapov & Fisher-Gewirtzman, 2016), physical effort (Greenberg, Natapov, & Fisher-Gewirtzman, 2020), reasoning about distance and time (Golledge, 1995; Li, Thrash, Hölscher, & Schinazi, 2019; Manley & Cheng, 2018), reliability and safety (Guo & Loo, 2013; Nasar & Jones, 1997; Olszewski & Wibowo, 2005), among others.[4] A key difference between the work mentioned above and ours is that this literature mostly focuses on explaining specific cognitive and psychological processes that determine pedestrian behavior. Although we acknowledge that pedestrians deviate from the shortest path for multiple reasons, including travelers' perception and preferences, our work is more closely related to empirical studies using large data sets that evaluate the built environment characteristics of routes taken by people (Yin, 2017; Zhu & Levinson, 2015). For example, Zacharias (1997a) assesses the role of visual information acquired in the course of moving through a large outdoor market in Montreal and finds that the layout and visual stimuli of a place contribute to the presence and movement patterns of people. Similarly, Foltête and Piombini (2010) show that variables describing the visual aspect of urban settings may influence the route choices of pedestrians. More recently, Yin and Wang (2016) applied machine learning algorithms on Google Street View images and found that visual variables are significantly associated with observed pedestrian counts. This paper complements this literature by using thousands of pedestrian trips automatically traced using GPS location (instead of relying on self-reported or modeled walking behavior) and taking into account a comprehensive set of built environment characteristics that include the geometric properties of streets (slope and sinuosity), access to business establishments, cultural and recreation facilities, route-specific services, public amenities, such as urban furniture and sidewalks, and urban design features of streets, such as facade complexity and visual enclosure.

Finally, our findings also inform the literature on place-making in urban design and planning that seeks to understand how well-designed places can attract people to relax, socialize, and be part of urban life. Recent work in place-making acknowledges that recreational purposes dominate the everyday use of public open spaces (Gehl & Matan, 2009). Thus, one of the key metrics to determine the success of public open space is to measure how often it is being used (Shaftoe, 2012). Practitioners such as Studio Gang and Gehl, for example, use observation to understand the frequency and type of activities of people in public space. Similarly, New Urbanist studies have highlighted the importance of "compact, pedestrian-friendly, and mixed-use design" (Talen, 2013), and argued that urban design plans should include perceptual, visual, temporal, behavioral, and environmental considerations (Carmona, Heath, Tiesdell, & Oc, 2010; Duany et al., 1991; Talen & Ellis, 2002). Empirical evidence on New Urbanism has documented the importance of offering public spaces and amenities to attain walkable streets with a strengthened sense of place and community (Talen, 1999; Trudeau, 2013). For example, Lund (2003) used surveys to study eight neighborhoods in Portland and found that placing amenities such as parks and retail shops within walking distance of homes increases pedestrian travel. Similarly, Beidler (2007) documents the importance of public space to encourage social encounters, which coincides with other studies that survey the landscape and use walk-along questionnaires to document the importance of parks and sidewalks for walkability (Ewing & Handy, 2009; Forsyth & Southworth, 2008). This paper complements the place-making literature by developing a street desirability measure for the whole city that quantifies how attractive every street is for pedestrians. By doing so, we strengthen the case study evidence on the importance of urban design to create places that can attract people.

The remainder of the paper is organized as follows. Section 1 details the data sources used in the empirical analysis. Section 2 describes the construction of the desirability index and illustrates its application to the city of Boston. Section 3 presents the results exploring which characteristics of the built environment determine the desirability of streets. Finally, Section 4 concludes with a discussion of the results and limitations of our approach.

2. Data description

For the empirical analysis, we use data on pedestrian trajectories, the street network, as well as data on the location of public and private amenities and built environment attributes that characterize streets.

2.1. Pedestrian trajectories

To measure the desirability of street segments, we use data on the trajectories of pedestrians, including their origin, destination, and path taken. We obtain these data from a company that aggregates anonymous data from a range of smartphone apps. The data consists of 127,082,227 GPS "pings" from 120,910 trips in Boston collected between May 2014 and May 2015. Since pings are logged whenever an application on a smartphone requests location information, they occur at irregular intervals. For instance, pings occur both while the user is using a particular application and also while the application runs in the background but continues to request information.[5] In our sample, the ping interval ranges from 1 to 10 seconds, which we believe provides a comprehensive picture of when and where individuals spend time.[6] For each ping, we observe a timestamp, a location, and a unique device identifier.

The origin and destination of each trip was inferred using a "stay-

[3] This approach builds on the theory developed by Brossard and Wieber (1984) "visible landscape", which studies the landscape as the sum of the parts that are visible to observers.

[4] See models of route choice such as the Stochastic User Equilibrium proposed by (Daganzo & Sheffi, 1977), or the Random Utility Maximization, focused on multiple plausible preferences (Abdelghany, Abdelghany, Mahmassani, Al-Ahmadi, & Alhalabi, 2010).

[5] The smartphone application was designed for walkers to keep track of their physical activity. This does not preclude app users from collecting data when driving, riding a bus, or cycling. To ensure that the data we used was indeed pedestrian activity, we focused our analysis on trips that were classified by the data provider as 'walking.' This categorization was done using speed-based filters.

[6] GPS data has shown to record data on individual users with superior accuracy and reliability compared to asking the users to recall their past activities (Forrest & Pearson, 2005; Pettersson & Zillinger, 2011)

time" algorithm implemented by the data provider.[7] A random distance between 0 and 100 meters was removed from the origin and destination of each trip to anonymize the users. Due to this censoring processing, very short distance-trips might be misrepresented in the data. Therefore, in our empirical analysis, we focus on pedestrian trajectories longer than 200 meters and shorter than 7 kilometers (from 2.5 minutes to around 1.5 hours of walking). In addition, we treat each recorded trip (even if sequential) as an independent trip.[8] The average trip length is 1,056 meters, and the average trip duration is 19 minutes.

To aggregate the GPS pings to individual trajectories, we implement the following procedure. First, we obtain the street network for Boston's metropolitan area from OpenStreetMap— an open-source mapping platform that collects geographic information. Second, we simplify the OSM street network so that a street is always defined as the street segment between street intersections.[9] Third, we assign each GPS ping to its closest street segment using the Graphhoper matching algorithm that allows us to snap measured GPS points to a street network (Malleson et al., 2018).[10]

We conduct most of our analysis at the street segment level. For each segment, we create two measures of pedestrian activity: a count of the total number of pedestrians passing through every street segment, and a count of the number of pedestrians that should pass through each street segment if they were to follow the shortest path from their origin to their destination. The former is a measure of the total observed flow of people passing by each street, and the latter is a measure of its potential— the number of people that would pass by that segment based on the location of origins, destinations, as well as the structural properties of the street network. To calculate the theoretical shortest path for each trip and count the number of shortest paths passing through each segment, we use Dijkstra's algorithm, which calculates the shortest distance between two vertices on a graph.

Our final sample includes 28,363 segments that cover a significant part of Boston's metropolitan area. For all of these segments, we compute the number of observed and shortest trips that passed through them, and also measure key attributes of their built environment, as we discuss next.

2.2. Measuring attributes of the built environment

For each street segment, we measure the following attributes: number of business establishments, number of cultural and recreation facilities, access to route services, neighboring parks, urban furniture, sidewalks, facade complexity, and visual enclosure. In addition, we measure attributes of the street network geometry such as the average sinuosity and slope. We construct these measures using data obtained from two sources: OpenStreetMap and Google Street View images. To measure the number of business establishments, cultural and recreation facilities, and access to route services in every street segment, we use the point data obtained from OpenStreetMap in 2014. The OpenStreetMap dataset contains the names, classification tag, and geographic coordinates for approximately 42 thousand points of interest in our study area. We first assign each point of interest to a given street segment by calculating its distance to the closest street segment and assigning it if it lies within a range of 100 meters or closer (the average length of an urban block). We then classify all the points of interest into one of the three categories using the OSM classification tags. For example, the business establishment measure includes OSM classification tags related to shops, bars, convenience stores, among others. The category of cultural and recreation facilities includes OSM classification tags related to art centers, libraries, museums, and sports facilities. The measure of access to route services includes OSM classification tags related to water fountains and toilets. The measure is a dummy variable that takes a value of one when there are route services available in a given street and a value of zero otherwise. Table A1 in the Appendix shows a comprehensive list of all the categories used in our classification.

For the empirical analysis, we also construct measures for the business makeup of the streets. In particular, we use the OSM classification tags to further group the business establishment category into four subcategories: food and drinks, retail, services, and hotels. In addition, we also construct a measure of business establishment variety, which captures the mix of uses in a given street. To construct this measure, we sum the number of times each unique OSM classification tag appears in each street segment.

To construct a measure for neighboring parks, we use the polygon data obtained from OpenStreetMap in 2014. We compute a 20-meter buffer around all the street segments in our sample and consider a street adjacent to a park if the park geometry overlaps the buffer. The neighboring park measure is a dummy variable that takes a value of one when parks are neighboring a given street segment and zero otherwise. This measure captures the value of a park for leisure, as well as the additional scenic value that parks provide to streets via abundant trees and vegetation.

We also construct measures of the sinuosity and slope for all street segments. To calculate the average slope of each street segment, we use the raster layer from the digital elevation surface model (DEM) at a resolution of 5 meters by 5 meters.[11] We first sample points along each street segment at 50-meter intervals. We then overlay the sampled points and the DEM model, assigning each point to a slope value and then taking the average for each street segment. To calculate the sinuosity of each street segment—a measure of how much a street segment deviates from a straight line—we use the street vertices obtained from the OSM street network. First, we calculate the distance between each pair of street vertices to get the great-circle distance of each street segment, and then we divide it by each street segment's length to derive its sinuosity. Sinuosity is computed as 1-distance/length. This measure is between 0 and 1, with higher values indicating more sinuosity.

The remaining measures are constructed by extracting visual attributes from images collected from Google Street View images via their Application Protocol Interface (API)[12]. In particular, we construct measures of urban furniture, sidewalks, facade complexity, and visual enclosure—a set of objective urban design attributes that have shown to be closely related to the pedestrian experience; but have proven difficult to measure at scale (Ewing and Handy, 2009). To collect the images, we use the 50-meter sample points along the OSM street network to query the Google API, which returns four images of size 400 × 400 pixels corresponding to images collected in 2014 or before.[13] Each four-image

[7] The details of the "stay-time" algorithm are proprietary information of the company and were not shared with researchers.

[8] In principle, each trip could be a multi-purpose trip with more than one stop. In this exercise, each stop marks the beginning or end of a trip. Conceptually, our definition of trip provides the relevant information because an origin and a destination already define trips, and hence, allow us to understand how people deviate when moving from a given origin to a given destination. Instead, if we treated a multi-purpose trip as the unit of analysis, we wouldn't be able to tease apart deviations driven by the need to pass through a given intermediate stop, from deviations driven by the built environment.

[9] To ensure that each street intersection is only represented once in the street graph, we use a hierarchical clustering algorithm to group intersections within 30 m or less.

[10] The positional accuracy of the GPS coordinates in our data is within 10 meters.

[11] The raster was obtained from MassGIS. https://docs.digital.mass.gov/dataset/massgis-data-digital-elevation-model-15000

[12] This approach has been shown to represent characteristics of the built environment accurately. See, for example, Naik et al., 2016; Salesses et al., 2013; Zhang et al., 2018.

[13] The Google API returns images taken between 2011 and 2019. Since our analysis period is 2014, we exclude from our analysis images taken after 2014 (10.5% of total images).

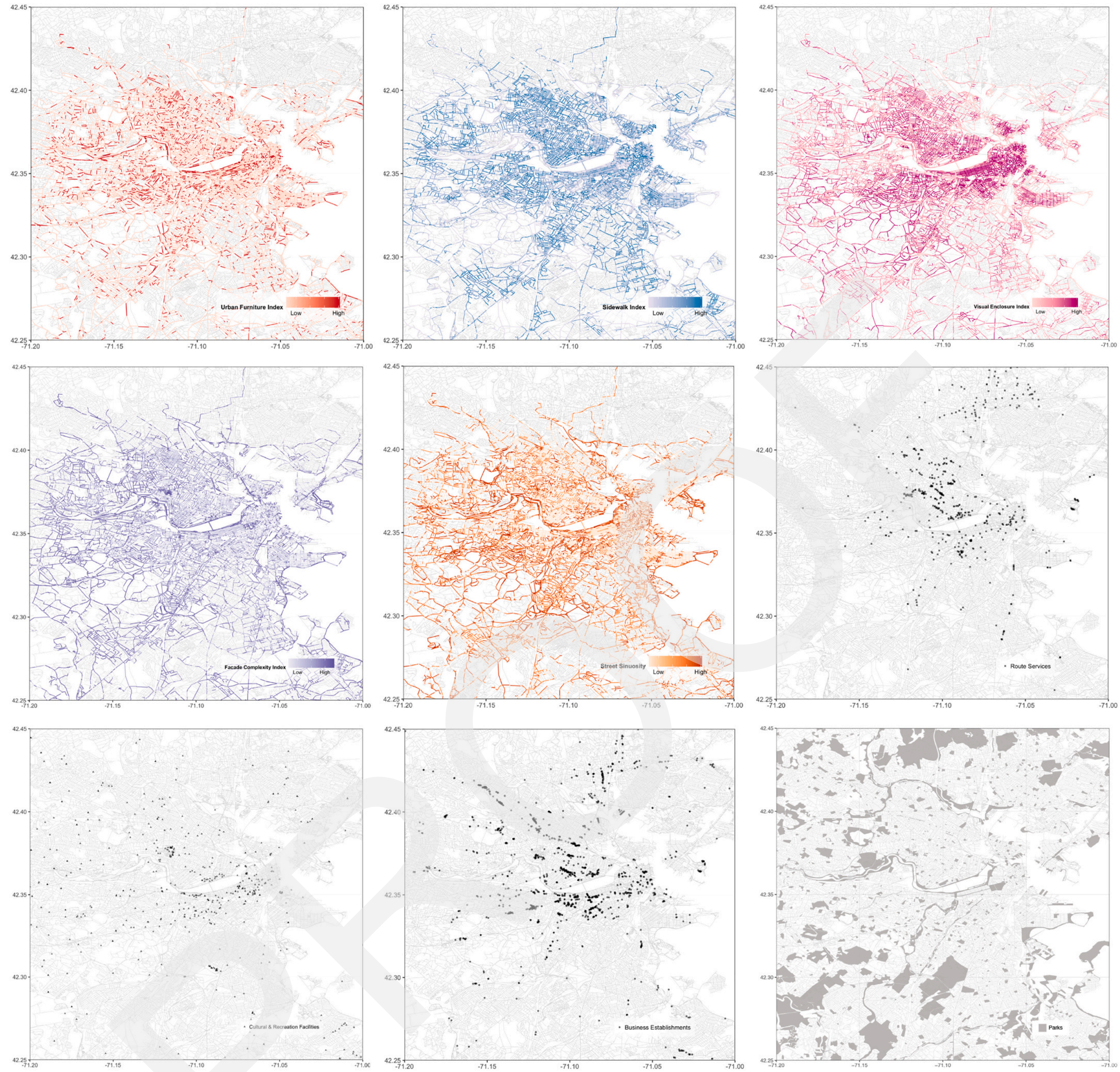

Fig. 1. Maps showing the indices constructed using images obtained from Google Street View (urban furniture, sidewalks, visual enclosure, and facade complexity) and the measures constructed using OSM data (street sinuosity, route services, cultural and recreational facilities, business establishments, and parks.)

set constitutes a street view panorama captured at a single point in time. Altogether, we obtained 1,520,054 images for the entire study area.[14]

To extract the urban design attributes from each image, we use computer vision techniques. In particular, we use an image semantic segmentation algorithm that assigns each pixel in an image to a given category (Zhu et al., 2020). This model is trained using the ADE20K dataset, which has a total of 150 categories available.[15] Out of the 150 categories, we focus on 15 categories that capture relevant features of the built environment and combine them to create four metrics: urban furniture, sidewalks, facade complexity, and visual enclosure[16]. These metrics measure the share of pixels of a given category relative to the total pixels in each image. The urban furniture metric measures the share of street furniture available in a given street segment and is

[14] Since Google Street View images can include both interior and exterior photos, we use the sky label to exclude interior images from our analysis. This reduced the sample size from 1,831,209 to 1,520,054. For street segments longer than 50 meters with more than one sampled point, we compute each metric's mean value.

[15] According to the ResNet18 prediction model specification, our overall pixel-wise accuracy was approximately 78.64%

[16] The 15 categories are benches, chairs, trash bins, sidewalks, sky, window, wall, building, tower, door, column, house, fence, skyscraper, handrails.

constructed by including pixels in an image that are classified as benches, chairs, or trash bins. The sidewalk metric measures the share of pixels covered by sidewalks in each street segment. We focus on these two measures motivated by the observation that streets with more furnishings and pedestrian infrastructure can help create safe environments for pedestrians to walk and are also critical motivating factors for pedestrian exploration (Zacharias, 1997b). The facade complexity metric measures the entropy of different materials in building facades. To compute this metric, we calculate the proportion of all facade materials relative to the total number of pixels for each image and sum these proportions multiplied by their natural logarithm. Facade complexity is defined as the negative of this sum so that higher values indicate more uneven materials. Conceptually, this measure captures the visual richness of building facades, which has been highlighted as a critical attribute of engaging and attractive environments (Nasar, 1994). Finally, the visual enclosure metric measures how well streets are visually defined by buildings, walls, trees, and other vertical elements. We compute this metric as the share of pixels that are not sky in an image. We focus on the visual enclosure of streets because it has implications for the quality of urban environments. In particular, less visually enclosed urban environments are thought to be less inviting for pedestrians (Southworth & Owens, 1993). Fig. A1 in the appendix provides an example of an image retrieved from Google Street View and the resulting image after the labeling procedure.

To ease the interpretation of these measures, we standardize them and work with indices defined in terms of standard deviations from the mean. These indices are available for a subset of 24,182 street segments (out of 28,363).[17]

To obtain controls for the primary use of street segments, we use the OSM categories to group street segments into three categories: heavy streets, moderate streets, and light streets. Heavy streets include street segments with OSM categories related to heavy car usages, such as trucks and highways. Light streets include street segments with OSM categories related to slow activity, such as walking. Finally, moderate streets include street segments with OSM categories that lie in-between the light and the heavy, such as secondary roads and cycleways. Fig. 1 maps the geographic distribution of all the variables used in our empirical analysis.

3. Constructing the desirability index

We compute a street segment's desirability by measuring whether pedestrians systematically deviate from their shortest path to frequent it. Our approach to using the shortest path as a natural baseline to compare observed trajectories is grounded on the common assumption that the shortest path can predict routing behavior because it maximizes the utility of a trip (Meyer, 2016). More broadly, this choice also aligns with human behavior models grounded in the random utility theory that emphasizes the importance of the shortest path in anchoring trajectories (Ben-Akiva, Lerman, & Lerman, 1985), making deviations from the shortest path meaningful to study.[18]

To estimate the desirability of the street segments, we first regress the total flow of pedestrians observed in each street segment on the number of pedestrians that should pass through each street segment if they were to follow the shortest path to their destination, as shown in equation (1).

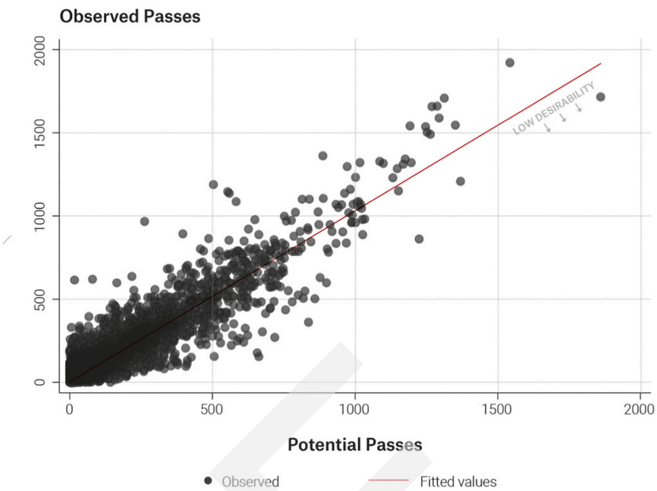

Fig. 2. Scatterplot of the potential and observed passes for each street segment.

$$\text{Log Observed Passes}_i = \gamma_0 + \gamma_1 \times \text{Log Potential Passes}_i + \upsilon_i. \tag{1}$$

Here, Log Observed Passes$_i$ is the natural logarithm of the total flow of people in street segment i, and Log Potential Passes$_i$ is the natural logarithm of the number of passes of people who would have passed if they would have taken the shortest route, and υ_i is the error term of the regression model.

Then, we calculate the desirability of each street segment as the residual of the following model:

$$\text{Desirability Index}_i = \text{Log Observed Passes}_i - \widehat{\gamma}_0 - \widehat{\gamma}_1 \times \text{Log Potential Passes}_i \tag{2}$$

The desirability index is high for streets that are not on the shortest path, yet they are the route of choice for many pedestrians. Conversely, the desirability index is low for streets *on* the shortest path, yet they are *not* the route of choice for many pedestrians. Importantly, by constructing the desirability index as a residual, we remove mechanical effects that prompt pedestrians to pass through some segments more frequently, just because they are centrally positioned in the street network or happen to be located on several shortest paths. This approach allows us to isolate the factors that make street segments along a given route more attractive relative to others on an alternative route.

Fig. 2 plots the observed passes and the potential passes of each street. The figure shows that the potential passes of a given street explain approximately 80% of the variation in total trips, leaving a 20% difference unexplained. This coincides with existing literature documenting that pedestrians frequently take the shortest route (Hill, 1982; Verlander & Heydecker, 1997), but that also a relatively high number of them deviate from it (Malleson et al., 2018). Motivated by the observation that 20% of the Boston trips do not take the shortest path, we explore the relationship between the observed passes and the constructed desirability index in the next part of the analysis.

Fig. 3 maps the observed passes (left) and the desirability index (right) for the streets in our sample. Values shown in red depict streets with high desirability (or a high number of observed passes in the left panel), and streets shown in blue depict those with low desirability (or a low number of observed passes in the left panel). The figure shows that observed passes are clustered towards the city center, where density is higher. This is to be expected as observed passes depend on global characteristics of the city, such as density, location, proximity to transit stops, and other attributes of the street network. On the other hand, areas of high and low desirability are widely dispersed throughout the study area. Desirable streets are found frequently towards the outskirts of the city, along

[17] We could not compute the image-based measures for 4,184 street segments, either because they only have images collected by Google after 2014 (our baseline year) or because Google has not collected an image for that street.

[18] This approach assumes that pedestrians know all their potential itineraries, which in turn allows them to anticipate the cost of these alternatives and to compare their utility before choosing the "optimal" route. Despite its wide use, this assumption has been criticized due to evidence suggesting that individuals have neither perfect knowledge of the urban environment nor a complete understanding of the possible alternatives offered by the street network (Golledge, 1999).

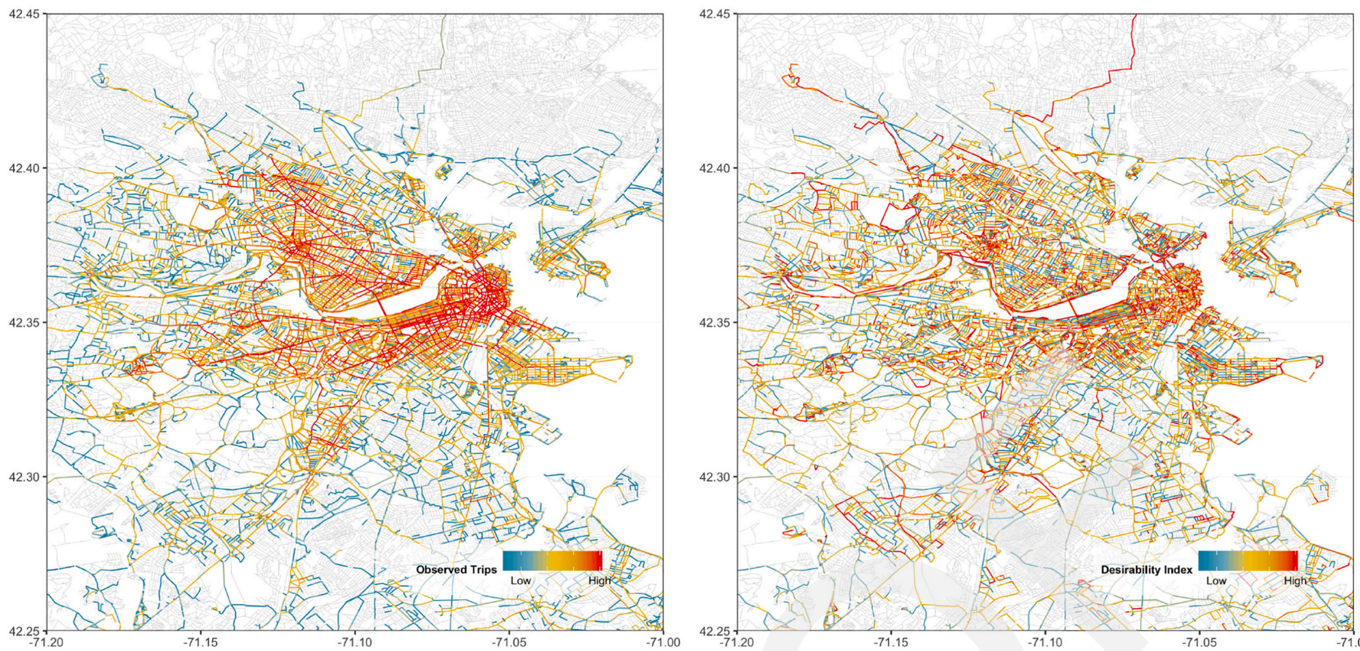

Fig. 3. Maps showing the observed trips (left) and average desirability (right).

Fig. 4. Maps showing the observed paths (in grey) and the top decile of street desirability (in red) in two selected areas. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

corridors connecting nearby towns with the city center, or proximate to parks. Fig. 4 zooms in on two selected neighborhoods and illustrates the observed passes (in grey) and the streets with the highest desirability (in red). The figure shows that the desirability index varies at a very local level. In particular, it captures meaningful variations across streets within neighborhoods, which speaks to the observation that the desirability index exploits very local differences in the environment instead of only capturing differences across large geographic areas.

3.1. *Desirability by trip type*

In the next section, we focus on how desirability changes for different types of trips. Since the nature of trips will likely differ across time of day, we group them into four categories: morning trips, noon trips, afternoon trips, and night trips. The morning and noon trips are defined as trips that happen anytime between 5:00 am and 9:00 am and between 11:00 am and 2:00 pm, respectively. The afternoon and night trips are defined as trips that happen anytime between 4:00 pm and 7:00 pm and between 7:00 pm and 11:00 pm, respectively. Our goal is to distinguish trips that are mostly driven by quickness and speed, usually taken in the morning, from less prescriptive after-work trips were individuals are more likely to visit restaurants, shops, and open spaces (Seneviratne & Morrall, 1985; Willis, Gjersoe, Havard, Kerridge, & Kukla, 2004).

To formally study the willingness to deviate from the shortest route for different types of trips, we re-estimate equation 1 for each type, which gives us a desirability index for each group. Across these different trips, the share of the variation in observed passes explained by the potential passes ranges from 0.48 (morning trips) to 0.69 (night trips). A low correlation between observed and potential paths indicates that individuals are systematically deviating from the shortest path. This

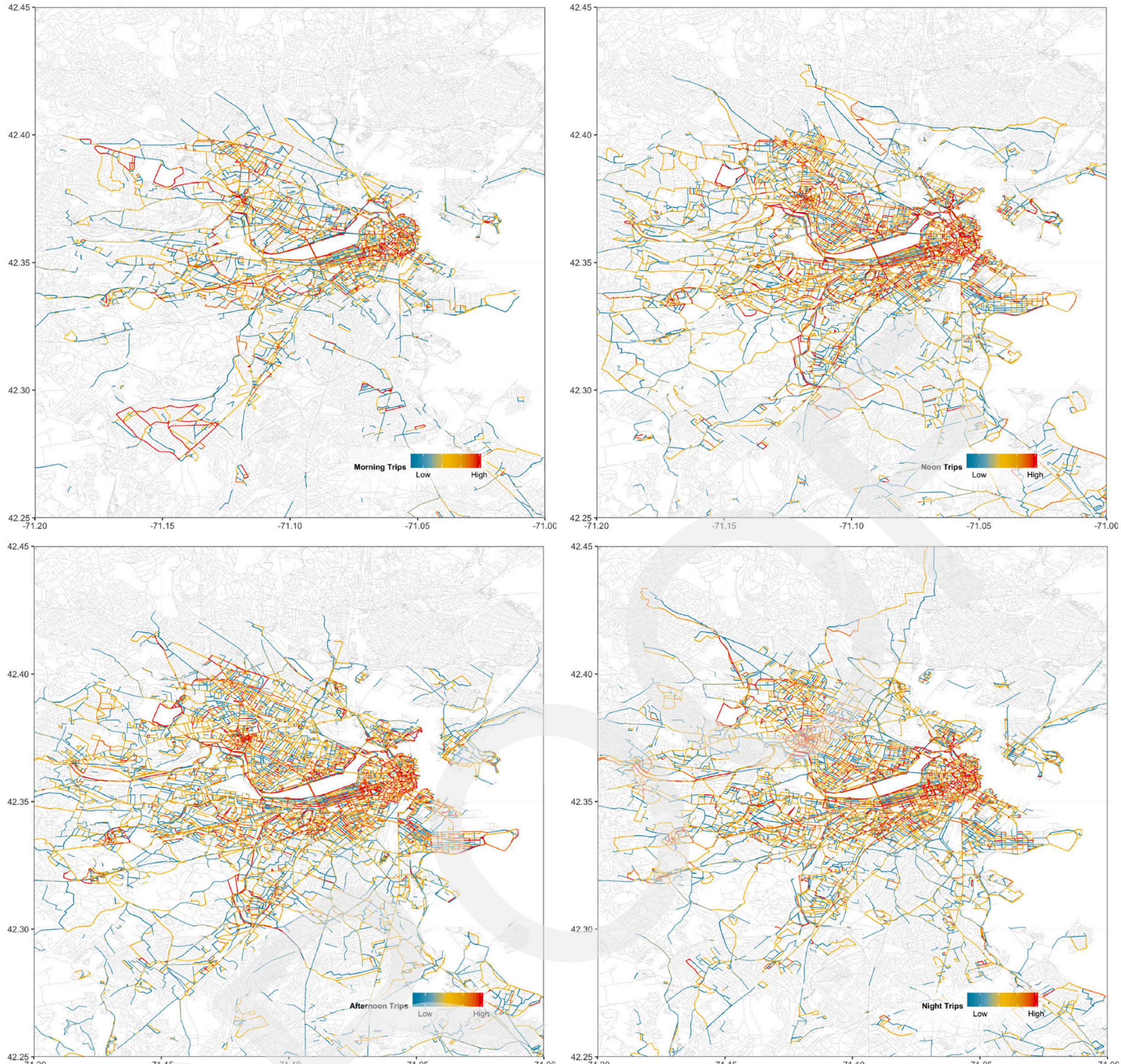

Fig. 5. Maps showing the average desirability index for different trip types.

finding suggests that pedestrians deviate from their shortest path at a similar frequency across trips of different types, with morning trips exhibiting the highest deviations.

Fig. 5 maps the desirability index for each of the four types of trips. Overall, desirable streets for morning trips are more clustered towards the central areas of Boston and Cambridge. In contrast, desirable streets for the rest of the trips (noon, afternoon, and night) show a less dense, more fragmented pattern that extends towards the outskirts of the city center. The desirability of street segments for morning and afternoon trips exhibits the lowest correlation (0.47), while the correlation for afternoon and night trips is the highest (0.69). Despite some geographic variation in the desirability computed for each type of trip, they are highly correlated overall.

4. Empirical patterns

In the next section of the paper, we explore the role of built environment characteristics in determining street desirability. We use all of the attributes introduced in the data section (street sinousity and slope, neighboring parks, urban furniture, sidewalks, facade complexity, visual enclosure, number of business establishments, and number of cultural and recreation facilities) and correlate them with the desirability index. Table 1 summarizes these attributes by splitting segments into high desirability (above 75th percentile) and low desirability (below 75th percentile). Most of the built environment characteristics of interest are more likely to be found in the highly desirable street segments. For example, as we move from streets with high desirability to low desirability, the share of parks drops from 0.35 to 0.30, and the sidewalk index drops from 0.047 to −0.057. The two exceptions are light streets and facade complexity, which are lower in streets with high desirability.

In Fig. 6, we present the built environment characteristics separately by type of trip. For each type of trip, the figure plots the difference in the mean value of each built environment characteristic between streets with high and low desirability. For comparison, the last bar in the figure represents the average difference among all trips, which coincides with the data reported in Table 1. A positive difference indicates that pedestrians systematically choose routes with more of these characteristics relative to the shortest route, contributing to the streets' desirability. On the contrary, a negative difference indicates that a given characteristic of the built environment lowers its desirability. The figure shows that, overall, the built environment characteristics associated with a segment's higher desirability are common across trips taken at different

Table 1
Summary statistics of built environment characteristics.

	Classification of desirability	
	High Desirability	Low Desirability
	($N = 7187$)	($N = 21{,}176$)
Heavy Street	0.146	0.138
	[0.353]	[0.345]
Moderate Street	0.511	0.508
	[0.500]	[0.500]
Light Street	0.343	0.354
	[0.475]	[0.478]
Street Sinuosity	0.043	−0.047
	[1.099]	[0.878]
Street Slope	2.168	2.022
	[3.624]	[3.232]
Neighboring Park	0.348	0.296
	[0.476]	[0.456]
Urban Furniture Index[a]	0.032	−0.036
	[1.201]	[0.717]
Sidewalk Index[a]	0.047	−0.057
	[1.097]	[0.874]
Facade Complexity Index[a]	−0.071	0.146
	[0.925]	[1.080]
Visual Enclosure Index[a]	0.091	−0.105
	[0.972]	[1.020]
Number of Business Establishments	0.166	0.095
	[0.679]	[0.432]
Number of Cultural & Recreation Facilities	0.035	0.022
	[0.207]	[0.167]

Notes: The Table reports the sample means and standard deviation (in brackets) of the built environment characteristics separately for high and low desirability streets. High desirability streets are those above the 75th percentile of the desirability distribution, and low desirability streets are those below the 75th percentile. The sample includes a total of 28,363 street segments for which the desirability measure was computed. The mean and standard deviation of the computer vision variables reported with the superscript [a] are computed for a sample of 24,182 segments. See Data Description Section in the main text for a comprehensive description of each variable.

hours of the day.

The previous exercises focused on trips classified by the time of the day. A complimentary exercise is to define trips using their total length (defined as the length of the shortest path), which accounts for the possibility that, compared to those who travel short distances, people who travel long distances might pay more attention to the shortest distance effect by weighing the built environment's features differently. To address this possibility, we constructed desirability indices separating trips by their total length into three categories: short trips with a total length of less than 1 km, medium trips with a total length between 1 and 3 km, and long trips of 3 or more kilometers. Fig. 7 shows that the factors associated with a segment's higher desirability are common for short and long trips. Thus, in the context of this study, people seem to value the same characteristics of the built environment independently of whether they are traveling a short or long distance. One exception are sidewalks and visual enclosure, which attract pedestrians making short and medium trips but do not seem to be an essential determinant of deviations for long trips (3 km or more).

Besides augmenting the analysis of trips by separating them by their length, we also tested the idea that people are more likely to follow the shortest path when traveling longer distances. We find that the shortest path has greater explanatory power over the observed path in short trips (correlation of 0.80 for trips shorter than 1 km) than in the longer ones (correlation of 0.35 for trips longer than 3 km). This finding suggests that, in our context, pedestrians deviate from the shortest path the most on longer trips. This observation contrasts previous findings suggesting that the proportion of deviations tend to be larger for short walks relative to longer ones (Sevtsuk, 2020).

Our analysis suggests that the characteristics of the built environment that determine a street's desirability are stable and shared across trips of different lengths and taken at different times of the day. In particular, public amenities such as neighboring parks, urban furniture, sidewalks, access to business and cultural facilities, as well as urban design features such as visual enclosure and more homogeneous facades contribute to the desirability of street segments for most types of trips. Motivated by these

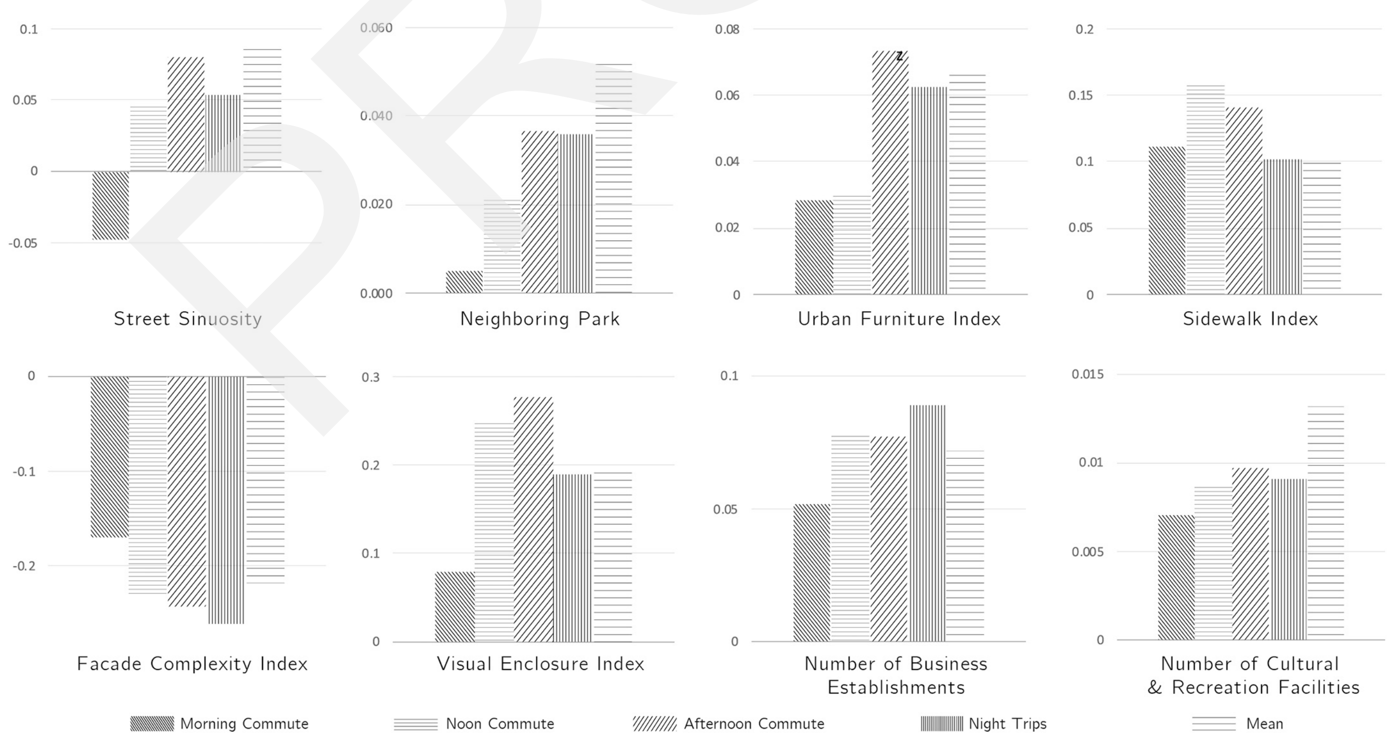

Fig. 6. Bar plots showing the difference between the frequency of each built environment characteristic in high and low desirability segments. Each panel reports the differences for trips taken at different times of the day.

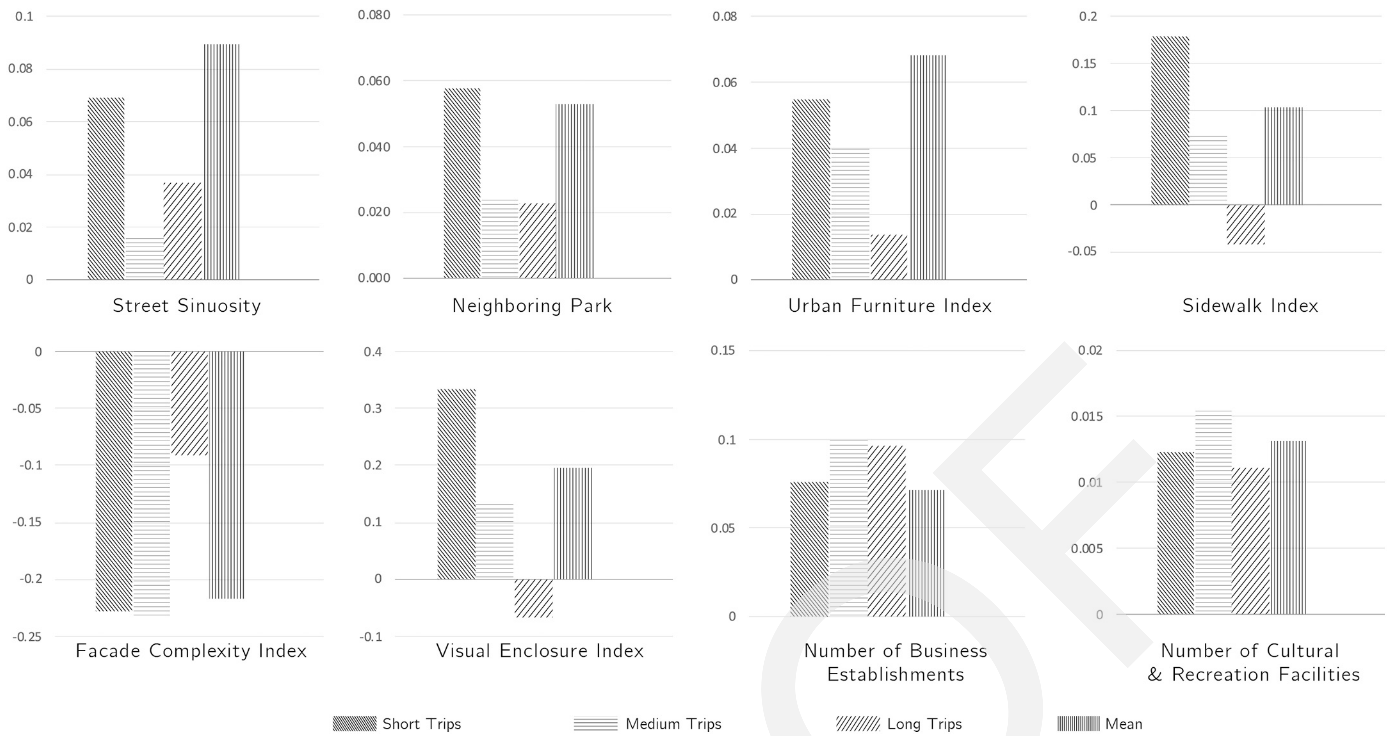

Fig. 7. Bar plots showing the difference between the frequency of each built environment characteristic in high and low desirability segments. Each panel reports the differences for trips of different lengths.

common patterns, in the rest of the paper, we will explore the determinants of the desirability index by pooling all trips together.

4.1. Regression evidence

To analyze the role of the different urban environment characteristics in explaining desirability, we estimate the following regression at the street segment level:

$$\text{Desirability Index}_i = \alpha + \beta \text{Built Environment}_i + \theta X_i + \varepsilon_i \quad (3)$$

The coefficient β captures the association between the attributes of the built environment ($\text{Built Environment}_i$) and the desirability index across street segments. To ensure that β is identified from a comparison of similar segments, we control for a vector X_i of additional segment characteristics measured in 2014. These controls include the average population of the Census Block containing each street segment, the length of each street segment (using a flexible combination of both linear and quadratic terms), and dummies for the type of street (heavy, moderate, and light).[19] In addition, all models are weighted by the number of trips passing through each street segment, so that the regressions are representative of the choices made by the average pedestrian. Finally, ε_i is the error term of the regression model. Standard errors are clustered at the Census Block level.

Table 2 presents the OLS estimates of equation 3. Each row reports the coefficients of the built environment characteristics on desirability. Column 1 presents results using the desirability index as the outcome and including public amenities, such as parks, urban furniture, and sidewalks as explanatory variables. Column 2 reports the results on businesses, cultural and recreation facilities, and access to services. Column 3 includes characteristics related to the street geometry, such as street sinuosity and slope. Finally, column 4 reports the results for the urban design measures, including facade complexity and visual enclosure.[20]

The estimates in column 4 show that access to public amenities such as parks and urban furniture is associated with more desirable street segments. The point estimate of 0.114 for neighboring parks implies that streets with a park increase their desirability by 11.4% relative to street segments with no park. That is, these street segments are visited 11.4% more often than what one would expect based only on their potential. Similarly, the 0.021 point estimate for urban furniture implies that a one standard deviation increase in the urban furniture index is associated with 2% higher street desirability. Finally, the point estimate of sidewalks implies that a one standard deviation increase in the sidewalk index is associated with 2.4% higher street desirability. Together, these results suggest that public amenities are important asset for streets to attract pedestrians.

Establishments exhibit positive but slightly less precise point estimates compared to public amenities (parks, sidewalks, and urban furniture). An additional business establishment increases the desirability of a street segment by 2%. The number of cultural and recreation facilities has a positive point estimate of 0.4%, but this coefficient is not precisely estimated.

Beyond public amenities and establishments, the geometric attributes of streets and their urban design also play a role in their desirability. We document that a one standard deviation increase in street sinuosity and visual enclosure is associated with 6.2% and 1.9% more street desirability, respectively. Finally, the −0.019 point estimate for facade complexity implies that a one standard deviation increase in facade complexity index is associated with a 1.9% lower desirability of streets. Table A2 presents results where we separately estimate equation 3 by trip type. Across trip types, the results are consistent with the findings in Table 2, except for the visual enclosure index.

Overall, our regression results show that built environment characteristics explain 5% of the variation in desirability. This should be interpreted as a lower bound on the importance of the built environment

[19] When a street segment overlaps with more than one Census Block, we use a weighted average of the blocks' population. We determine the weights using the share of the segment that lies within each block.

[20] See Fig. A2 in the Appendix for binscatter plots between each built environment characteristic and the desirability index.

Table 2
Log Desirability: OLS estimates for all trips using OSM data weighted by number of trips.

	Dependent variable: Log Desirability			
	(1)	(2)	(3)	(4)
Neighboring Park	0.109***	0.110***	0.109***	0.114***
	(0.016)	(0.016)	(0.017)	(0.016)
Urban Furniture Index	0.021***	0.020***	0.020***	0.021***
	(0.004)	(0.004)	(0.004)	(0.004)
Sidewalk Index	0.031***	0.030***	0.030***	0.024***
	(0.008)	(0.007)	(0.008)	(0.008)
Access to Route Services		0.083***	0.079***	0.076***
		(0.029)	(0.028)	(0.028)
Number of Cultural & Recreation Facilities		0.013	0.015	0.004
		(0.020)	(0.020)	(0.020)
Number of Business Establishments		0.021***	0.023***	0.020***
		(0.006)	(0.006)	(0.006)
Street Sinuosity			0.059***	0.062***
			(0.009)	(0.009)
Street Slope			0.001	0.001
			(0.002)	(0.002)
Visual Enclosure Index				0.019***
				(0.006)
Facade Complexity Index				−0.019*
				(0.011)
Observations	24,179	24,179	24,179	24,179
R-squared	0.04	0.04	0.05	0.05
Covariates and weighting:				
Weighted by Number of Trips	✓	✓	✓	✓
Length of Street Segment	✓	✓	✓	✓
Quadratic Length of Street Segment	✓	✓	✓	✓
Street Type	✓	✓	✓	✓
Average Census Block Population	✓	✓	✓	✓

Notes: The Table reports the regression coefficients for the built environment characteristics. All models are weighted by the total number of observed trips in each street segment and control for the length of street segments (linear and quadratic), the type of street type (heavy, moderate, and light), and the average census block population. Column 1 presents the coefficients for public amenities. Column 2 shows the coefficients for business establishments, cultural and recreational facilities, and route services. Column 3 shows the coefficients for street sinuosity and slope. Column 4 shows the coefficients for facade complexity and visual enclosure. Robust standard errors clustered at the Census Block level reported in parentheses.
*** denotes a coefficient significant at the 1% level, ** at the 5% level, and * at the 10% level.

because it only accounts for the share of the variation captured by the built environment attributes included in this analysis.[21]

4.1.1. Spillover and network effects

Our previous results use street segments as the unit of analysis and attempt to explain their desirability based solely on their characteristics. However, since street segments are part of a network, the number of pedestrians passing by a given segment depends not only on this segments' characteristics but also on whether nearby segments have desirable features. This observation suggests the possibility of spillovers or network effects, wherein how frequented a segment is also depends on its neighbors' desirability.

To study this possibility empirically, we estimate the following variant of equation (3)

$$\text{DesirabilityIndex}_i = \alpha + \beta \text{BuiltEnvironment}_i + \gamma \text{DesirabilitySpillovers}_i + \theta X_i + \varepsilon_i, \quad (4)$$

where Desirability Spillovers$_i$ is the average desirability of segments that are adjacent to i.

The challenge in estimating this equation is that there is a mechanical *reflection problem*. Intuitively, if there is an unobserved feature of segment i that makes it more desirable, it will, by definition, make neighboring segments more desirable too. This introduces a positive correlation between ε_i and Desirability Spillovers$_i$ that confounds OLS estimates of this equation.

To address this econometric problem, we estimate equation 3 via 2SLS. In particular, we use predetermined characteristics of neighboring segments (their length, street type, slope, and sinuosity) as instruments for their desirability. This approach solves the reflection problem because these predetermined characteristics are not affected by other factors that might make segment *i* more desirable. Arguably, these instruments also satisfy the exclusion restriction since they are likely to affect neighboring pedestrian activity only through the desirability of neighbors. The 2SLS estimates identify spillovers by comparing segments with similar characteristics that have neighbors with more (or less) desirable features. For example, if the desirability index is higher for segments next to flat and short neighboring segments, the 2SLS estimate interprets this as a positive spillover.

Table A3 in the appendix conducts this analysis and shows that spillovers are an important determinant of the desirability of a segment. In particular, a 10% increase in the desirability of neighboring streets increases a segment's desirability by 4%. Moreover, accounting for spillovers increases the R-squared to 14%, suggesting a gain in explanatory power when considering the neighbors' characteristics. Accounting for spillovers does not change our qualitative results regarding the characteristics of the built environment that are associated with desirability. However, our estimates for the association between these characteristics and the desirability of a street segment become less precise, reflecting the fact that it is now harder to separately identify the role of spillovers from that of built environment characteristics that are highly clustered in space.

4.1.2. Variety and diversity of businesses

To further unpack the relationship between business establishments and the desirability of streets, in this section, we turn to explore what types of businesses are valued the most by pedestrians and if the variety of business is more important than their sheer number.

Column 1 of Table 3 reports the results for the number of business establishments. In column 2, we replace the number of establishments with the measure of business variety. In column 3, we include the number of different business establishment types, including foods and drinks, retail, services, and hotels. Finally, in column 4, we include dummy variables for each type of business establishment.

The estimates in column 2 show that the variety of business establishments matters more than the total number of business establishments reported in column 1. The point estimate of 0.038 for business variety implies that adding a new type of establishment increases street desirability by 3.8%. Column 3 shows that the coefficient for retail is positive and sizable. An additional retail establishment increases desirability by 6.3%. Finally, column 4 shows that having access to at least one food and drink establishment (as shown by the point estimate of 0.040 in column 4) is more important than the total number of food and drink establishments (as shown by the point estimate of 0.015 in column 3). In contrast, we find no difference in hotels across both models.

5. Conclusion

This paper leverages the revealed preference of thousands of pedestrians deviating from their shortest path to construct a desirability index for streets. This measure captures the scenic and experience value provided by

[21] Given the high level of disaggregation of the data, we find an R^2 of 5% to be meaningful. There are many unobserved idiosyncratic factor affecting the flow of pedestrians through a single street segment. Many of these factors would average out in a more aggregate analysis, while here they are part of the unexplained variation.

Table 3
Desirability and variety of business establishments: OLS estimates.

	Dependent variable: Log Desirability			
	(1)	(2)	(3)	(4)
Number of business establishments	0.016**	−0.014		
	(0.006)	(0.016)		
Variety of establishments		0.038**		
		(0.018)		
Number of food and drinks			0.015**	
			(0.008)	
Number of retail establishments			0.063**	
			(0.027)	
Number of service establishments			−0.031*	
			(0.016)	
Number of hotels			−0.066	
			(0.050)	
Dummy of food and drinks				0.040***
				(0.014)
Dummy of retail establishments				0.075**
				(0.034)
Dummy of service establishments				−0.054***
				(0.019)
Dummy of hotels				−0.070
				(0.049)
Observations	28,363	28,363	28,363	28,363
R-squared	0.02	0.02	0.02	0.02
Covariates and weighting:				
Length of street segment	✓	✓	✓	✓
Quadratic length of street segment	✓	✓	✓	✓
Street type	✓	✓	✓	✓
Average census block population	✓	✓	✓	✓
Weighted by number of trips	✓	✓	✓	✓

Notes: The Table reports the regression coefficients for the number and variety of businesses on a street segment's desirability. All models control for the street segment length (linear and quadratic), the street type (light, moderate, heavy), the average census block population, and are weighted by the total number of observed trips in each street segment. Column 1 reports the coefficients for the total number of businesses. Column 2 reports the coefficients for business variety. Column 3 reports the coefficients for the number of different types of business establishments. Finally, column 4 reports the coefficients for dummies for each type of business establishment. Robust standard errors clustered at the Census Block level reported in parentheses.
*** denotes a coefficient significant at the 1% level, ** at the 5% level, and * at the 10% level.

different parts of the city. Beyond its conceptual value, the desirability index offers practical value for urban planners by allowing them to identify areas with high potential to attract pedestrians. Conversely, it can also be used to pinpoint areas with low desirability that would be suitable candidates for street improvements and other interventions.

The first part of the paper describes how we construct the desirability index and then shows its application to Boston. We show that people systematically deviate from the shortest path to frequent highly desirable streets. This pattern is consistent across different types of trips, including trips taking place at different times of the day (morning, noon, afternoon, and night trips) and trips of different lengths (defined by the total length of the shortest path). This finding suggests that although in principle, people may value the features of the built environment differently depending on the time of the day and the length of the trip, the desirability index captures a common set of features that are desirable to most pedestrians in our sample. Second, we show that desirable street segments are widely dispersed throughout the study area. From this observation, we learn that there are streets across all neighborhoods whose characteristics appeal to pedestrians. Moreover, these local differences in the built environment are commonplace, and pedestrians across all neighborhoods systematically deviate from their shortest path to experience them. The geographic dispersion of desirable streets also highlights a methodological advantage of using deviations from the shortest path rather than the count of pedestrians to describe the desirability of streets. Since the construction of the desirability index exploits very local differences in the environment instead of differences across large geographic areas, it allows us to characterize desirable streets relative to their nearby context instead of comparing broader differences between neighborhoods—for example, by comparing streets located at the center versus the periphery. Another advantage of this approach is that it allows us to trace how specific streets change in time. This approach can be useful, for example, to revitalize distressed streets or detect areas affected by blight or other negative phases of urban change.

In the second part of the paper, we analyze a diverse set of built environment characteristics that have been shown to affect the pedestrian experience. In particular, we construct a set of measures that aim to capture the provision of amenities in each street segment, as well as some of its urban design features. To do so, we combine detailed data on businesses, amenities and the street layout pulled from OpenStreetMaps with urban design characteristics of streets collected from Google Street View imagery.

By comparing street segments with varying levels of desirability and different built environment characteristics, we document that the desirability index is systematically associated with the provision of public amenities. More desirable streets are characterized by having better access to parks, sidewalks, and a higher presence of urban furniture. These results provide complementary evidence underscoring the importance of pedestrian infrastructure for supporting pedestrian activity (Saelens & Handy, 2008a, 2008b), and also strengthens evidence on the importance of parks for physical activity and leisure (Coombes, Jones, & Hillsdon, 2010; Lee & Maheswaran, 2011).

The paper also provides evidence on the urban design characteristics that are associated with more desirable streets. In particular, desirable streets tend to be more visually enclosed, sinuous, and tend to have more homogeneous facades. Our results on visual enclosure align with normative urban design proposals suggesting that great walking environments should feel like a grand corridor with buildings that define and enclose space (Jacobs & Appleyard, 1987; Speck, 2012). The fact that desirable streets are more sinuous aligns with the view that irregular street forms can easily accommodate focal locations for landmark buildings and reduce traffic due to their curvature. Finally, our findings on desirable streets being more homogeneous in terms of their facade present empirical support for the notion that architectural homogeneity is positively valued by pedestrians, which aligns with existing evidence on residential property markets (Lindenthal, 2020), evidence pointing to commonalities in architectural preferences (Nasar, 1992), and more broadly with studies underscoring the importance of urban design characteristics and aesthetics in determining behavior (Kopec, 2012; Lynch, 1960).

Finally, our results point to desirable streets having more business establishments. Although the number of business establishments is an important determinant of desirability, the variety of businesses is more important in predicting it. In addition, we find that among all the included businesses, retail is consistently associated with higher desirability of streets. These results are in line with theories that suggest that the density and variety of land uses can encourage frequent opportunities for interaction (Jacobs, 1961; Talen, 1999), can raise the vitality of places (Yue et al., 2017), and also aligns with evidence suggesting that retail shops are key to increase pedestrian travel (Lund, 2003).

There are several limitations in this study that could be addressed in future research. First, due to a non-disclosure agreement, we have very little information about the individuals whose devices we observe in our data. Not having access to demographic data, for example, limits the generalizability of our findings. Second, the information we do have comes from an unknown selection process based on when devices' phone applications request and log location data. Although the data generating process is uncertain, the granularity of our data (pings at 1 second intervals) is frequent enough to confidently provide a detailed picture of where people are at different types of the day. Third, because the GPS data is collected as people go about their daily life, it tends to overweight commuting patterns. This limits how much we can say about the type of people being attracted to specific streets and the types of activities they engage with once they are there. Finally, the desirability measure we propose is a local measure. This means that although it can be calculated for entire cities in practice, the desirability of a street should always be interpreted with respect to its immediate context. The local nature of the measure limits the extent to which one can draw useful comparisons about the average desirability of large geographical areas or cities.

Appendix A.

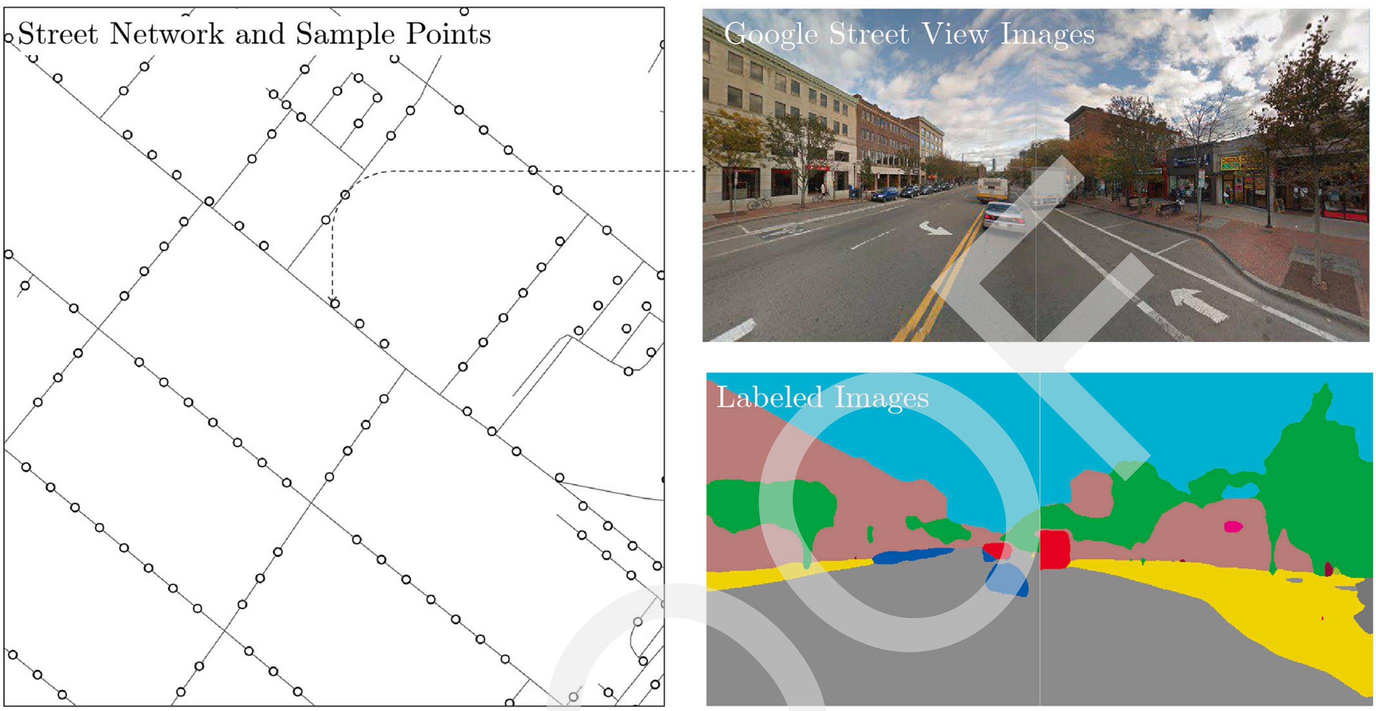

Fig. A1. Figure showing the sampled points from the OSM streets network used to query the images from Google Street View (left) and a labeled image to illustrate the classification algorithm's output (right).

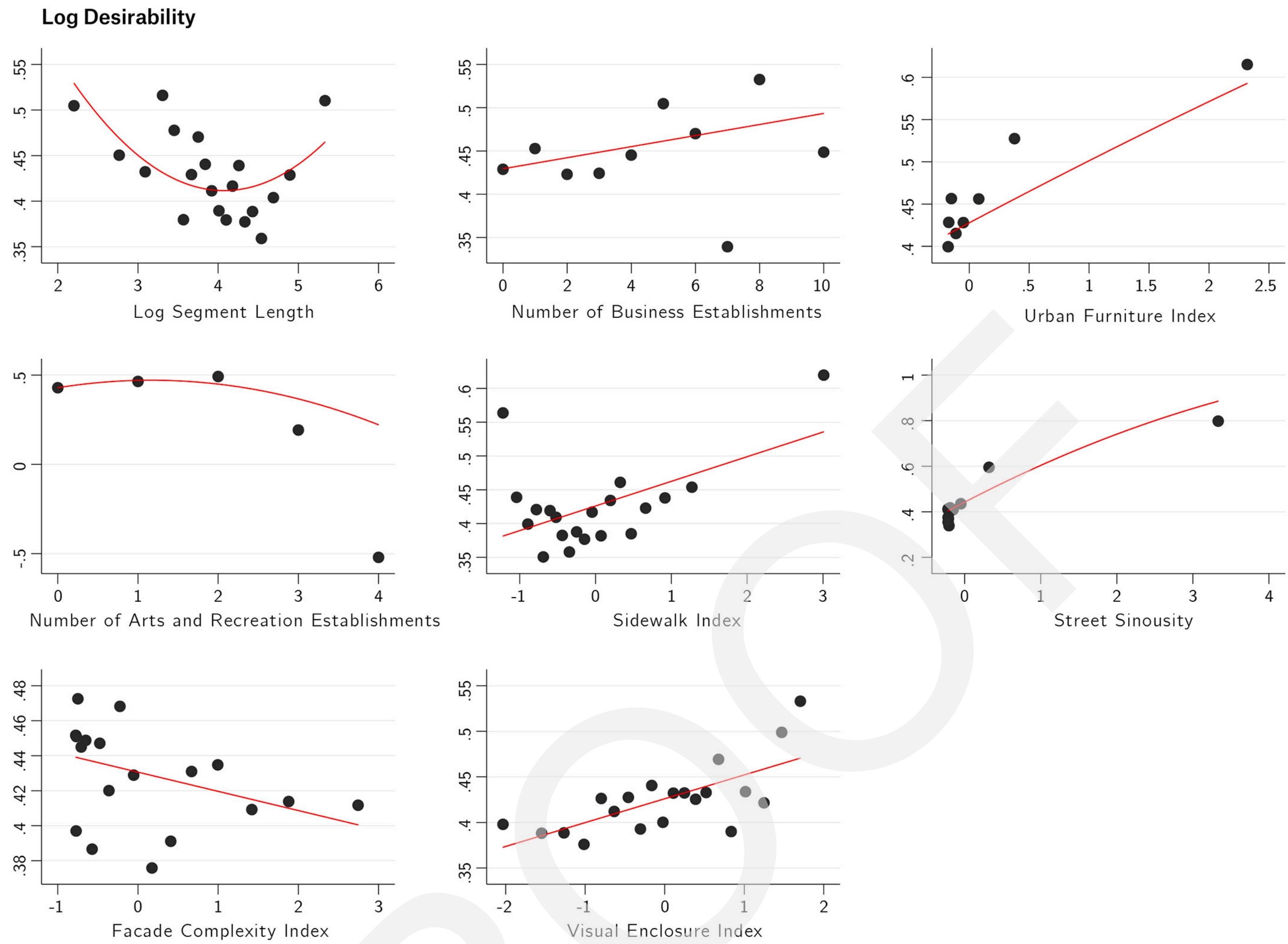

Fig. A2. Binscatter plots of built environment characteristics and the desirability index used in Table 2.

Table A1
OSM category classification.

Groups used in empirical analysis	Original OSM tags
Cultural & recreational facilities	arts centre, artwork, attraction, community centre, kindergarten, library, museum, playground, sports centre, stadium, swimming pool, theatre, arts centre ....
Business establishments	bakery, bank, bar, beauty shop, beverages, bicycle shop, bookshop, butcher, cafe, car dealership, car wash, cinema, clothes, computer shop, convenience, dentist, department store, doctors, fast food, florist, food court, fuel, furniture shop, gift shop, greengrocer, hairdresser, hospital, hostel, hotel, jeweller, laundry, mall, mobile phone shop, motel, optician, outdoor shop, pharmacy, pub, restaurant, shoe shop, sports shop, stationery, shop, supermarket, toy shop, travel agent, veterinary, video shop..........
Route amenities	atm, bicycle rental, drinking water, fountain, parking bicycle, post box, telephone, toilet, tourist info, vending any, vending machine..........
Food and drinks	pub, bar, beverages, cafe, fast food, food court, restaurant..........
Retail	beauty shop, bicycle shop, bookshop, butcher, car dealership, clothes, computer shop, convenience, department store, florist, fuel, furniture shop, gift shop, greengrocer, jeweller, mall, mobile phone shop, outdoor shop, shoe shop, sports shop, stationery, shop, supermarket, toy shop, video shop..........

(*continued on next page*)

Table A1 (*continued*)

Groups used in empirical analysis	Original OSM tags
Services	bank, car wash, cinema, dentist, doctors, hairdresser, laundry, optician, pharmacy, travel agent, veterinary, hospital..........
Neighboring park	parks..........

Table A2
Log desirability: OLS estimates by trip length using OSM data weighted by number of trips.

	Dependent variable: Log Desirability		
	Short Trips	Medium Trips	Long Trips
	Less than 1 km	1–3 km	3 km or more
	(1)	(2)	(3)
Neighboring park	0.076***	0.137***	0.193***
	(0.013)	(0.026)	(0.047)
Urban furniture index	0.013***	0.023***	0.042***
	(0.005)	(0.005)	(0.007)
Sidewalk index	0.034***	0.035***	0.006
	(0.007)	(0.013)	(0.013)
Access to route services	0.072***	0.109***	0.283***
	(0.024)	(0.027)	(0.063)
Number of cultural & recreation facilities	0.016	−0.006	−0.022
	(0.017)	(0.026)	(0.041)
Number of business establishments	0.025***	0.032***	0.059***
	(0.006)	(0.007)	(0.017)
Street sinuosity	0.056***	0.048***	0.030***
	(0.009)	(0.008)	(0.008)
Street slope	0.001	0.002	−0.003
	(0.001)	(0.005)	(0.007)
Facade complexity index	−0.018	−0.043***	−0.048***
	(0.012)	(0.013)	(0.012)
Visual enclosure index	0.041***	0.021***	−0.041***
	(0.005)	(0.008)	(0.015)
Observations	19,980	18,530	12,236
R-squared	0.05	0.05	0.06
Covariates and weighting:			
Weighted by number of trips	✓	✓	✓
Length of street Segment	✓	✓	✓
Quadratic length of street segment	✓	✓	✓
Street type	✓	✓	✓
Average Census Block population	✓	✓	✓

Notes: The Table reports the estimates of eq. 1 for different samples using the specifications from column 4 in Table 2. Column 1 shows estimates of column 3 for short trips of less than 1 km (85%). Column 2 shows estimates of column 3 for medium trips between 1 and 3 km (14%). Column 3 shows estimates of column 3 for long trips of 3 km or more (1%). Robust standard errors clustered at the Census Block level reported in parentheses.
*** denotes a coefficient significant at the 1% level, ** at the 5% level, and * at the 10% level.

Table A3
Log desirability: IV estimates for all trips using OSM data weighted by number of trips.

	Dependent variable: Log Desirability			
	(1)	(2)	(3)	(4)
Desirability spillovers	0.407***	0.387***	0.321**	0.321***
	(0.134)	(0.131)	(0.127)	(0.112)
Neighboring park	0.060**	0.063***	0.070***	0.073***
	(0.024)	(0.025)	(0.026)	(0.023)
Urban furniture index	0.018***	0.018***	0.018***	0.018***
	(0.004)	(0.004)	(0.004)	(0.004)
Sidewalk index	0.024***	0.023***	0.024***	0.022***
	(0.008)	(0.007)	(0.008)	(0.008)
Access to route services		0.055**	0.056**	0.054**
		(0.026)	(0.026)	(0.026)
Number of cultural & recreation facilities		−0.007	−0.001	−0.005
		(0.020)	(0.020)	(0.019)
Number of business establishments		0.014**	0.017**	0.015**
		(0.007)	(0.007)	(0.007)
Street sinuosity			0.054***	0.055***
			(0.009)	(0.009)
Street slope			0.001	0.001
			(0.002)	(0.002)
Visual enclosure index				0.004

(*continued on next page*)

Table A3 (*continued*)

	Dependent variable: Log Desirability			
	(1)	(2)	(3)	(4)
				(0.007)
Facade complexity index				−0.010
				(0.007)
First-stage F stat	53.53	52.58	51.00	59.18
Observations	24,134	24,134	24,134	24,134
R-squared	0.14	0.14	0.14	0.14
Covariates and weighting:				
Weighted by number of trips	✓	✓	✓	✓
Length of street segment	✓	✓	✓	✓
Quadratic length of street segment	✓	✓	✓	✓
Street type	✓	✓	✓	✓
Average Census Block population	✓	✓	✓	✓

Notes: The Table reports the regression coefficients for the built environment characteristics. All models are weighted by the total number of observed trips in each street segment and control for the length of street segments (linear and quadratic), the type of street type (heavy, moderate, and light), and the average census block population. Column 1 presents the coefficients for the built environment characteristics related to the visual characteristics of streets. Column 2 shows the coefficients business establishments and recreation amenities. Column 3 shows the coefficients for the straightness and slope of streets. Column 4 shows the coefficients for facade complexity and visual enclosure. Robust standard errors clustered at the Census Block level reported in parentheses.

*** denotes a coefficient significant at the 1% level, ** at the 5% level, and * at the 10% level.

References

Abdelghany, A., Abdelghany, K., Mahmassani, H., Al-Ahmadi, H., & Alhalabi, W. (2010). Modeling the evacuation of large-scale crowded pedestrian facilities. *Transportation Research Record, 2198*, 152–160.

Beidler, K. J. (2007). *Sense of place and new urbanism: Towards a holistic understanding of place and form*. Environmental Design and Planning. Doctor of.

Ben-Akiva, M., & Bierlaire, M. (2006). Discrete choice models with applications to departure time and route choice. *Handbook of Transportation Science, 7*–37.

Ben-Akiva, M. E., Lerman, S. R., & Lerman, S. R. (1985). *Discrete choice analysis: Theory and application to travel demand* (vol. 9). MIT press.

Brossard, T., & Wieber, J.-C. (1984). Le paysage: trois définitions, un mode d'analyse et de cartographie. *L'Espace géographique*, 5–12.

Calthorpe, P. (1993). *The next American metropolis: Ecology, community, and the American dream*. Princeton architectural press.

Carmona, M., Heath, T., Tiesdell, S., & Oc, T. (2010). *Public places, urban spaces: The dimensions of urban design*. Routledge.

Cervero, R., & Radisch, C. (1996). Travel choices in pedestrian versus automobile oriented neighborhoods. *Transport Policy, 3*(3), 127–141.

Coombes, E., Jones, A. P., & Hillsdon, M. (2010). The relationship of physical activity and overweight to objectively measured green space accessibility and use. *Social Science & Medicine (1982), 70*(6), 816–822.

Daganzo, C. F., & Sheffi, Y. (1977). On stochastic models of traffic assignment. *Transportation Science, 11*(3), 253–274.

Debord, G. (1958). *Theory of the dérive*. Atlantic Books.

Duany, A., Plater-Zyberk, E., Krieger, A., & Lennertz, W. R. (1991). *Towns and town-making principles* (Rizzoli).

Ewing, R., & Handy, S. (2009). Measuring the unmeasurable: Urban design qualities related to walkability. *Journal of Urban Design, 14*(1), 65–84.

Ewing, R., Handy, S., Brownson, R. C., Clemente, O., & Winston, E. (2006). Identifying and measuring Urban Design qualities related to walkability. *Journal of Physical Activity and Health, 3*(s1), S223–S240.

Foltête, J.-C., & Piombini, A. (2010). Deviations in pedestrian itineraries in urban areas: A method to assess the role of environmental factors. *Environment and Planning. B, Planning & Design, 37*, 723–739.

Forrest, T. L., & Pearson, D. F. (2005). Comparison of trip determination methods in household travel surveys enhanced by a global positioning system. *Transportation Research Record, 1917*, 63–71.

Forsyth, A., & Southworth, M. (2008). Guest editorial: Cities afoot - pedestrians, walkability and urban design. *Journal of Urban Design, 13*(1), 1–3.

Gehl, J., & Matan, A. (2009). Two perspectives on public spaces. *Building Research and Information, 37*(1), 106–109.

Giles-Corti, B., & Donovan, R. J. (2003). Relative influences of individual, social environmental, and physical environmental correlates of walking. *American Journal of Public Health, 93*(9), 1583–1589.

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., … Donovan, R. J. (2005). Increasing walking: How important is distance to, attractiveness, and size of public open space? *American Journal of Preventive Medicine, 28*(2 SUPPL. 2), 169–176.

Golledge, R. G. (1995). *Path selection and route preference in human navigation: A progress report*. PhD thesis. University of California at Santa Barbar.

Golledge, R. G. (1999). *Wayfinding behavior, cognitive mapping and other spatial processes* (p. 428). The Johns Hopkins University Press.

González, M. C., Hidalgo, C. A., & Barabási, A. L. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779–782.

Greenberg, E., Natapov, A., & Fisher-Gewirtzman, D. (2020). A physical effort-based model for pedestrian movement in topographic urban environments. *Journal of Urban Design, 25*(1), 86–107.

Guo, Z., & Loo, B. P. (2013). Pedestrian environment and route choice: Evidence from new York City and Hong Kong. *Journal of Transport Geography, 28*, 124–136.

Hess, P. M., Moudon, A. V., Snyder, M. C., & Stanilov, K. (1999). Site design and pedestrian travel. *Transportation Research Record, 1674*, 9–19.

Hill, M. R. (1982). *Spatial structure and decision-making aspects of pedestrian route selection through an urban environment*.

Hillier, B., Penn, A., Hanson, J., Grajewski, T., & Xu, J. (1993). Natural movement: Or, configuration and attraction in urban pedestrian movement. *Environment and Planning. B, Planning & Design, 20*(1), 29–66.

Jacobs, J. (1961). *The death and life of great American cities*. New York: Random House, Inc.

Jacobs, A. (1993). *Great streets*. MIT Press.

Jacobs, A., & Appleyard, D. (1987). Toward an Urban Design manifesto. *Journal of the American Planning Association, 53*(1), 112–120.

Jiang, S., Ferreira, J., & González, M. C. (2012). Clustering daily patterns of human activities in the city. *Data Mining and Knowledge Discovery, 25*(3), 478–510.

Kelly, C. M., Wilson, J. S., Baker, E. A., Miller, D. K., & Schootman, M. (2013). Using Google street view to audit the built environment: Inter-rater reliability results. *Annals of Behavioral Medicine, 45*(SUPPL.1), 108–112.

Kopec, D. A. (2012). *Environmental psychology for design*. Fairchild Books.

Krier, L. (1984). Urban components. *Architectural Design, 54*(7/8), 43–49.

Law, S., & Traunmueller, M. (2017). Off the shortest path: Betweenness on street network level to study pedestrian movement. In *2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC)* (pp. 971–976).

Lee, A. C., & Maheswaran, R. (2011). The health benefits of urban green spaces: A review of the evidence. *Journal of Public Health, 33*(2), 212–222.

Li, H., Thrash, T., Hölscher, C., & Schinazi, V. R. (2019). The effect of crowdedness on human wayfinding and locomotion in a multi-level virtual shopping mall. *Journal of Environmental Psychology, 65*(June).

Lindenthal, T. (2020). Beauty in the Eye of the Home-Owner: Aesthetic Zoning and Residential Property Values. *Real Estate Economics, 48*, 530–555.

Lund, H. (2003). Testing the claims of new urbanism: Local access, pedestrian travel, and neighboring behaviors. *Journal of the American Planning Association, 69*(4), 414–429.

Lynch, K. (1960). *The image of the city*. MIT Press.

Malleson, N., Vanky, A., Hashemian, B., Santi, P., Verma, S. K., Courtney, T. K., & Ratti, C. (2018). The characteristics of asymmetric pedestrian behavior: A preliminary study using passive smartphone location data. *Transactions in GIS, 22*(2), 616–634.

Manley, E., & Cheng, T. (2018). Exploring the role of spatial cognition in predicting urban traffic flow through agent-based modelling. *Transportation Research Part A: Policy and Practice, 109*(January), 14–23.

Manley, E. J., Addison, J. D., & Cheng, T. (2015). Shortest path or anchor-based route choice: A large-scale empirical analysis of minicab routing in London. *Journal of Transport Geography, 43*, 123–139.

Meyer, M. D. (2016). Chapter 21. In *Recreational areas* (pp. 975–1012). John Wiley & Sons, Ltd.

Montgomery, J. (1995). Urban vitality and the culture of cities. *Planning Practice and Research, 10*(2), 101–110.

Moudon, A. V., Hess, P. M., Snyder, M. C., & Stanilov, K. (1997). Effects of site design on pedestrian travel in mixed-use, medium-density environments. *Transportation Research Record, 1578*, 48–55.

Naik, N., Raskar, R., & Hidalgo, C. A. (2016). Cities are physical too: Using computer vision to measure the quality and impact of urban appearance. *American Economic Review, 106*(5), 128–132.
Nasar, J. L. (1992). *Environmental aesthetics: Theory, research, and applications*. Cambridge: Cambridge Univ. Press.
Nasar, J. L. (1994). Urban Design aesthetics: The evaluative qualities of building exteriors. *Environment and Behavior, 26*(3), 377–401.
Nasar, J. L., & Jones, K. M. (1997). Landscapes of fear and stress. *Environment and Behavior, 29*(3), 291–323.
Natapov, A., & Fisher-Gewirtzman, D. (2016). Visibility of urban activities and pedestrian routes: An experiment in a virtual environment. *Computers, Environment and Urban Systems, 58*, 60–70.
Olszewski, P., & Wibowo, S. S. (2005). Using equivalent walking distance to assess pedestrian accessibility to transit stations in Singapore. *Transportation Research Record, 1927*, 38–45.
Ordonez, V., & Berg, T. L. (2014). Learning high-level judgments of urban perception. In D. Fleet, T. Pajdla, B. Schiele, & T. Tuytelaars (Eds.)*, vol. 8694*. *Computer vision ECCV* (pp. 494–510). Cham: Springer International Publishing.
Pettersson, R., & Zillinger, M. (2011). Time and space in event behaviour: Tracking visitors by GPS. *Tourism Geographies, 13*(1), 1–20.
Pikora, T. J., Giles-Corti, B., Knuiman, M. W., Bull, F. C., Jamrozik, K., & Donovan, R. (2006). Neighborhood environmental factors correlated with walking near home: Using SPACES. *Medicine & Science in Sports & Exercise, 38*(4).
Porta, S., Crucitti, P., & Latora, V. (2008). Multiple centrality assessment in Parma: A network analysis of paths and open spaces. *Urban Design International, 13*(1), 41–50.
Purciel, M., & Marrone, E. (2006). *Observational validation of urban design measures for New York City. Technical report* (Columbia, New York).
Quercia, D., O'Hare, N., & Cramer, H. (2014). Aesthetic capital: What makes London look beautiful, quiet, and happy ?. In *Proceedings of the 17th ACM conference on computer supported cooperative work & social computing* (pp. 945–955).
Ratti, C., Frenchman, D., Pulselli, R. M., & Williams, S. (2006). Mobile landscapes: Using location data from cell phones for urban analysis. *Environment and Planning. B, Planning & Design, 33*(5), 727–748.
Saelens, B. E., & Handy, S. L. (2008a). Built environment correlates of walking: A review. *Medicine & Science in Sports & Exercise, 40*(Supplement), S550–S566.
Saelens, B. E., & Handy, S. L. (2008b). Built environment correlates of walking: A review. *Medicine and Science in Sports and Exercise*, 40.
Saiz, A., Salazar, A., & Bernard, J. (2018). Crowdsourcing architectural beauty: Online photo frequency predicts building aesthetic ratings. *PLoS One, 13*(7), 1–15.
Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the City: Mapping the inequality of urban perception. *PLoS One, 8*(7).
Sallis, J. F., Floyd, M. F., Rodríguez, D. A., & Saelens, B. E. (2012). Role of built environments in physical activity, obesity, and cardiovascular disease. *Circulation, 125*(5), 729–737.
Seneviratne, P. N., & Morrall, J. F. (1985). Analysis of factors affecting the choice of route of pedestrians. *Transportation Planning and Technology, 10*(2), 147–159.
Sevtsuk, A. (2014). Location and agglomeration: The distribution of retail and food businesses in dense urban environments. *Journal of Planning Education and Research, 34*(4), 374–393.
Sevtsuk, A. (2020). *Street commerce. The hidden structure of retail location patterns and vibrant sidewalks*. Philadelphia: Penn Press.
Sevtsuk, A., & Ratti, C. (2010). Does urban mobility have a daily routine? Learning from the aggregate data of mobile networks. *Journal of Urban Technology, 17*(1), 41–60.
Sevtsuk, A., Kalvo, R., & Ekmekci, O. (2016). Pedestrian accessibility in grid layouts: The role of block, plot and street dimensions. *Urban Morphology, 20*, 89–106.
Shaftoe, H. (2012). *Convivial urban spaces: Creating effective public places* (Earthscan).
Shen, Y., & Karimi, K. (2016). Urban function connectivity: Characterisation of functional urban streets with social media check-in data. *Cities, 55*, 9–21.
Shoval, N. (2008). Tracking technologies and urban analysis. *Cities, 25*(1), 21–28.
Southworth, Michael, & Owens, Peter M. (1993). The Evolving Metropolis: Studies of Community, Neighborhood, and Street Form at the Urban Edge. *Journal of the American Planning Association, 59*(3), 271–287. https://doi.org/10.1080/01944369308975880.
Speck, J. (2012). *Walkable city: how downtown can save {America}, one step at a time* (1st ed ed.). New York: Farrar, Straus and Giroux.
Talen, E. (1999). Sense of community and neighbourhood form : An assessment of the social doctrine of new urbanism. *Urban Studies, 36*(8), 1361–1379.
Talen, E. (2013). *Charter of the new urbanism: congress for the new urbanism* (2nd ed.). New York: McGraw Hill Education.
Talen, E., & Ellis, C. (2002). Beyond relativism: Reclaiming the search for good city form. *Journal of Planning Education and Research, 22*(1), 36–49.
Talen, E., & Koschinsky, J. (2013). The walkable neighborhood: A literature review. *International Journal of Sustainable Land Use and Urban Planning, 1*(1), 42–63.
Trudeau, D. (2013). A typology of new urbanism neighborhoods. *Journal of Urbanism, 6*(2), 113–138.
Vanky, A. P., Verma, S. K., Courtney, T. K., Santi, P., & Ratti, C. (2017). Effect of weather on pedestrian trip count and duration: City-scale evaluations using mobile phone application data. *Preventive Medicine Reports, 8*, 30–37.
Verlander, N., & Heydecker, B. (1997). Pedestrian route choice: An empirical study. In *Ptrc conference* (pp. 1–5). Brunel University (September).
Willis, A., Gjersoe, N., Havard, C., Kerridge, J., & Kukla, R. (2004). Human movement behaviour in urban spaces: Implications for the design and modelling of effective pedestrian environments. *Environment and Planning. B, Planning & Design, 31*(6), 805–828.
Wolff, M. (1973). Notes on the behaviour of pedestrians. *People in Places: The Sociology of the Familiar, 35*.
Ye, Y., Zeng, W., Shen, Q., Zhang, X., & Lu, Y. (2019). The visual quality of streets: A human-centred continuous measurement based on machine learning algorithms and street view images. *Environment and Planning B: Urban Analytics and City Science, 46*(8), 1439–1457.
Yin, L. (2017). Street level urban design qualities for walkability: Combining 2D and 3D GIS measures. *Computers, Environment and Urban Systems, 64*, 288–296.
Yin, L., & Wang, Z. (2016). Measuring visual enclosure for street walkability: Using machine learning algorithms and Google street view imagery. *Applied Geography, 76*, 147–153.
Yue, Y., Zhuang, Y., Yeh, A. G. O., Xie, J.-Y., Ma, C.-L., & Li, Q.-Q. (2017). Measurements of {POI}-based mixed use and their relationships with neighbourhood vibrancy. *International Journal of Geographical Information Science, 31*(4), 658–675.
Zacharias, J. (1997a). The impact of layout and visual stimuli on the itineraries and perceptions of pedestrians in a public market. *Environment and Planning. B, Planning & Design, 24*(1), 23–35.
Zacharias, J. (1997b). Virtual path decisions of visitors to the historic core of Montpellier, France. In *Proceedings of the international conference on behaviour-environment studies for the 21st century*.
Zhang, F., Zhou, B., Liu, L., Liu, Y., Fung, H. H., Lin, H., & Ratti, C. (2018). Measuring human perceptions of a large-scale urban region using machine learning. *Landscape and Urban Planning, 180*(August), 148–160.
Zhu, S., & Levinson, D. (2015). Do people use the shortest path? An empirical test of wardrop's first principle. *PLoS One, 10*(8), 1–18.
Zhu, D., Zhang, F., Wang, S., Wang, Y., Cheng, X., Huang, Z., & Liu, Y. (2020). Understanding place characteristics in geographic contexts through graph convolutional neural networks. *Annals of the American Association of Geographers, 110*(2), 408–420.