Journal of Transport Geography 74 (2019) 191–200

Contents lists available at ScienceDirect

Journal of Transport Geography

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

Why do bicyclists take detours? A multilevel regression model using smartphone GPS data

Yujin Park*, Gulsah Akar

City and Regional Planning, The Ohio State University, 275 West Woodruff Ave, Columbus, OH 43210, USA

ARTICLE INFO

Keywords:
Bicycle detouring
Bicycle route preference
Environmental correlates
Multilevel mixed-effect generalized linear regression
Smartphone GPS application

ABSTRACT

Bicyclists often deviate from the shortest possible routes and take detours in search of more pleasant riding conditions. The extent of detours and the factors affecting bicyclists to ride excess distances have not yet been fully explored. This study aims to measure and analyze the detour extent of utilitarian bicycle trips and their relationships with the route-level environmental components using data collected from individual bicyclists' smartphone GPS in Columbus, Ohio. Comparing the chosen routes with their shortest counterparts, we calculate two detour indices (a distanced-based index and an area-based index) and provide a comparative analysis of built environment attributes for low, moderate, and high levels of detours. We then estimate multilevel mixed-effect generalized linear regression models to identify the contribution of built-environment characteristics to such detours while accounting for individual heterogeneity. We find that most bicycle trips (91.1%) include a detour and are 13.5% longer on average than their shortest alternatives with large variations. Detour degrees are higher for long-distance trips and for peak-period trips. We find that bicyclists choose routes with smaller shares of commercial and single-family land-uses and low levels of land-use diversity. Longer detours are positively associated with street greenery. We find that sparse bicycle facilities and high-speed limits are strong contributors to bicyclists' detour decisions, while multilevel mixed-effect linear regression models further present significant heterogeneity in bicyclists' responses to some environmental attributes. The area-based detour index performs better in explaining the relationships between land-use features and detour degrees.

1. Introduction

Bicycling has the potential to mitigate the negative environmental and health outcomes of our dependence on motorized modes and enhance urban vitality (Aultman-Hall et al., 1997; Buehler and Handy, 2008). As an active mode of transportation, bicycling is susceptible to the influence of external environments which may either foster pleasant rides or lead to detours to avoid potential nuisances or risks (Akar and Clifton, 2009; Broach and Bigazzi, 2017). Misra and Watkins (2017) argue that bicycle trips are different from vehicular trips because they are likely to be on routes that are optimal in safety and comfort, rather than on the shortest routes in travel time and/or distance. Empirical studies show that many bicyclists often choose suboptimal routes in terms of distance in search of better riding conditions, experiencing varying degrees of excess travel (Aultman-Hall et al., 1997; Krenn et al., 2014; Winters et al., 2010). This excess travel induces more effort per trip, undermining the efficiency and attractiveness of bicycling. Since travel efficiency affects bicycling choice and frequency of use in the long run, a better understanding of factors related to detours is important to improve the bicycling experience.

Many studies on bicycle routing behavior suggest that bicycle route choices are made through the joint consideration of efficiency, safety, and leisure (Broach et al., 2012; Casello and Usyukov, 2014; Chen et al., 2017; Hood et al., 2011; Sener et al., 2009; Wuerzer and Mason, 2015; Zimmerman et al., 2017). However, there is a lack of studies assessing detour extent resulting from such choices and its environmental correlates. Some studies measure excess travel distance and its correlations with several environmental determinants, but through exploratory analysis (Aultman-Hall et al., 1997; Duncan and Mummery, 2007; Krenn et al., 2014), or in a multiple travel mode context (e.g. car, transit, bicycle, and/or walking) (Dalton et al., 2015; Ta et al., 2016; Winters et al., 2010). Focusing only on bicycle trips, Misra and Watkins (2017) find that limited bicycle facilities and steep slopes along the shortest route make bicyclists seek alternative routes. Nonetheless, it remains unclear how the built environment and land-use attributes of both chosen and shortest routes affect detour extent after individual tendencies in route choice are taken into account.

In this paper, we examine environmental correlates affecting

* Corresponding author.
E-mail addresses: park.2329@osu.edu (Y. Park), akar.3@osu.edu (G. Akar).

https://doi.org/10.1016/j.jtrangeo.2018.11.013
Received 28 February 2018; Received in revised form 22 August 2018; Accepted 18 November 2018
Available online 04 December 2018
0966-6923/

bicyclists' detour decisions and the extent of these detours, with a focus on the differences between the chosen routes and the network-based shortest routes. We use the data collected through a smartphone GPS-based application, *CycleTracks*™,[1] to observe the bicycle trip trajectories of The Ohio State University community in Columbus, Ohio. Using two detour indices, Route Directness Index (RDI) and Normalized Degree of Deviation (NDD), we assess the degree of detours and their associations with the surrounding environmental and trip characteristics while accounting for the effect of individual heterogeneity.

2. Literature review

Researchers cite three reasons as to why excess travel occurs: poor planning, personal preference and misperception (Broach and Bigazzi, 2017; De Vos et al., 2016; Handy et al., 2005; Papinski and Scott, 2013; Ta et al., 2016). Ta et al. (2016) argue that the built environment has a major impact on the levels of detours.[2] They find that excess travel distances are closely related to poor street connectivity and low access to transport facilities, such as transit stations. Broach and Bigazzi (2017) suggest that networks with higher intersection densities can help mitigate the extent of deviation from the shortest path by offering a wider range of options for bicyclists while avoiding traffic-related pollution. Sener et al. (2009) provide a comprehensive review of factors associated with bicyclists' route choices. They classify these factors as physical (e.g. slope gradient, number of stop signs and signalized intersections), functional (e.g. traffic volume, speed limit), operational characteristics (e.g. travel distance or time) and bicycle amenities (e.g. bicycle facility type and continuity). The negative influences of roadway upslope, high traffic volumes, and making left turns at signalized intersections are found to be significant in many studies, although the magnitude of their impacts varies from study to study (Broach et al., 2012; Chen et al., 2017; Hood et al., 2011; Khatri et al., 2016; Zimmermann et al., 2017).

Bicycle facilities have a significant influence on route deviation. Using an online stated preference survey, Sener et al. (2009) find that bicyclists are willing to add on average 13 min to an original commute for the use of a separated bicycle path. In another stated preference survey, Tilahun et al. (2007) find that bicyclists would incur up to 20 min of additional travel time to ride on off-street bicycle paths. Although these studies suggest substantial diversions based on stated preference data, Menghini et al. (2010) find that bicyclists would add about 233 m to ride on a continuous bicycle path using revealed GPS data. More recently, Standen et al. (2017) implement an intercept survey for commuting and non-commuting bicycle riders, asking them directly how far they diverted from their old paths to use the new green cycleway in Sydney. They find that, on average, commuters and non-commuters diverted by 252 m and 551 m, respectively. Broach et al. (2012) show that commuter bicyclists' excess travel acceptance varies by facility type. They find that bicycle riders would accept a 16.0% increase for exclusive bike paths and a 10.8% for bike boulevards.

Surrounding land-uses and street features are also relevant to bicycle routing behaviors because many bicyclists seek diverse utilities during their trips, such as pleasant views and recreation (Plazier et al., 2017). Zacharias and Zhang (2016) examine whether certain streets with particular land-use characteristics are preferred than others by bicyclists. They find that bicyclists prefer streets with tree shades and are not likely to choose crowded streets due to restaurants and pedestrians. Similarly, Krenn et al. (2014) find that detour routes would have fewer restaurants and stores and better scenery as compared to direct (i.e. shortest) routes in their exploratory analysis. Chen et al. (2017) investigate the impacts of several land-use features on bicyclist route choices, including the level of land-use mix, lighting density, and the proportion of waters and parks. They find that lighting density has a robust and positive impact on route choices, while the proportion of waters and parks and land-use mix have varied impacts by trip purpose.

The impacts of environmental attributes on bicycle route choices, however, may often be uneven among individual bicyclists, as they can be mediated by unobserved personal preferences (Chen et al., 2017; Sener et al., 2009). The utility of transport choice consists of decision utility and experienced utility (De Vos et al., 2016; Handy et al., 2005). While decision utility refers to the objectively observable utility of users, such as travel time and cost, experienced utility reflects users' subjective travel experience and pleasure (De Vos et al., 2016; Handy et al., 2005). Handy et al. (2005) suggest that some travelers often choose to drive for a longer period of time intentionally to maximize the pleasure and joy of traveling. De Vos et al. (2016) find that the experienced utility of bicycling is closely related to land use preferences and personal cognitive evaluation. For example, Sener et al. (2009) find that some bicyclists are more safety-conscious, while others are time-conscious, and these personal tastes manifest themselves as different sensitivities to route attributes. Therefore, bicyclists' subjective preferences in routing behavior must be considered in modeling efforts.

Overall, the existing literature reveals the links between route characteristics, land-use, personal preferences, and bicycle routing behavior. However, little attention has been paid to the impacts of those environmental factors on bicycle travel efficiency outcomes. Assessing travel efficiency levels is important as they are directly and indirectly related to the feasibility of bicycle activities for urban travelers. This study adds to the existing literature by investigating the relationships of environmental components with bicycle detour levels, while accounting for the heterogeneity in individual bicyclists' responses towards those factors.

3. Data & methods

3.1. GPS data for bicycling trajectories

The data for this study come from the actual bicycle trajectories revealed by individual smartphone GPS (Global Positioning System). This high-resolution tracking technology offers significant potential to capture detailed movements of bicyclists, helping us better understand how the built environment influences routing behavior (Khatri et al., 2016). Origins, destinations, routes and time of travel, as well as basic personal information on bicyclists, were collected through a smartphone application, *CycleTracks*™, developed by San Francisco County Transportation Authority (SFCTA). We recruited study participants among the members of The Ohio State University (OSU) in Columbus, Ohio, including faculty, staff, graduate and undergraduate students, from August 21 to November 31, 2016. To promote our research and data collection efforts, we sent email invitations to a random sample of individuals. We asked survey participants to download the *CycleTracks* app and record their bicycle trips by turning the app on and off at the beginning and end of each bicycle trip. At the end of every trip, participants were requested to select the purpose of their trip. The options available were: Commute, School, Work-related, Errand & Shopping, Exercise, and Social.

We collected data on 1531 bicycle trips made by 78 bicyclists. Using the trip purposes revealed by the participants, we dropped leisure trips (i.e. Exercise and Social) for detour analysis because it is difficult to apply the concept of travel efficiency, which this paper aims to examine, to leisure trips. The ability to get to destinations efficiently by bicycle is important for utilitarian trips, but it may not necessarily be relevant for leisure trips. For instance, if the purpose is to exercise, one may intentionally choose a longer route. We also find that, in our data, the destinations of the leisure trips are almost always the same as their

[1] This application is developed by San Francisco County Transportation Authority (SFCTA).

[2] We use the term 'detour' in this study to broadly refer to the excess portion of travel distance as compared with the network-based minimum distance or to the situation where a bicyclist rides a longer distance than actually needed.

origins, without any additional stops. In this case, the shortest distance is zero, and a distance-based detour index cannot be computed. For these reasons, we only included utilitarian trips in our final analysis sample. It is possible that some individuals intended to introduce some level of exercise into their utilitarian trips. We assume that the influence of this personal preference can be accounted for by random coefficients for individuals in our multilevel mixed-effect regression models.

Data preparation consists of three steps: (i) GPS data cleaning, (ii) generation of a complete bicycle network, and (iii) map-matching of cleaned GPS traces to the complete network. We first cleaned up the noise and errors in the initial GPS data. We then removed those trips that were unlikely to be bicycle trips based on travel speed. Some of the trips include stops en route at locations such as shops. We split these trips into two separate trips in order to avoid overestimating the magnitude of the detour. We used OpenStreetMap (OSM) and updated these network data significantly to include all possible links that can be used for bicycling, such as park trails, alleys, and passageways. Finally, cleaned GPS traces were matched to the reconstructed network links. After removal of similar routes generated by the same rider, the final sample covered 439 unique utilitarian trips (i.e. commute, school, work-related, and errand & shopping) made by 73 bicyclists. Table 1 presents the descriptive statistics for this sample. Of the sample, 46.6% are daily bicyclists and 57.5% are male. Not all participants revealed their basic demographic information.

The City of Columbus makes a good case study area for bicycle travel efficiency research for three reasons. First, in contrast to the metropolitan cities targeted in previous studies–Portland (Broach et al., 2012); Waterloo (Casello and Usyukov, 2014); Seattle (Chen et al., 2017); San Francisco (Hood et al., 2011); Phoenix (Khatri et al., 2016); Zurich (Menghini et al., 2010); Beijing (Ta et al., 2016)–, Columbus is a medium-to-low density urban area with diverse urban settings, such as a downtown, suburban and urban communities, natural areas, and a large university campus (OSU). Second, the city has decent quality bike trails along river banks, good stretches of bicycle boulevards in residential areas, and other types of bicycle facilities along main roads. Lastly, the city has a predominantly flat topography, and its typical peak-hour traffic is rarely heavy except on highways, in contrast to other metropolitan areas. With these characteristics, the study area helps enrich the geographic diversity of literature.

Table 1
Sample characteristics.

Characteristics	Classification	N	%
Gender	Male	42	57.5
	Female	15	20.5
	Unknown	16	21.9
Age	18–25	21	28.8
	26–35	16	21.9
	36–45	6	8.2
	46–55	9	12.3
	56 and above	5	6.8
	Unknown	16	21.9
Bicycling frequency	Daily	34	46.6
	Several times per week	20	27.4
	Several times per month	3	4.1
	Less than once a month	0	0.0
	Unknown	16	21.9
Purpose per trip	Commute	286[a]	65.2
	School	89[a]	20.3
	Work-related	41[a]	9.3
	Errand & Shopping	23[a]	5.2
Total		73 (439[a])	100.0

[a] Number of trips

3.2. Dependent variables: detour indices

To measure the level of detour, we employ two measurement indices: Route Directness Index (RDI, %) and Normalized Degree of Deviation (NDD, m^2/m). RDI has been frequently used as an indicator of excess travel (Dalton et al., 2015; Misra and Watkins, 2017; Papinski and Scott, 2013; Standen et al., 2017; Ta et al., 2016; Zhu and Levinson, 2015). RDI is derived from the ratio of a chosen route distance to its shortest route distance:

$$RDI = \left(\frac{Chosen\ Route\ Distance}{Shortest\ Route\ Distance} - 1\right) \times 100\ (\%)$$

The lowest possible value is zero when the chosen route is the network-based shortest path. The RDI represents the increased amount of disutility of a chosen route by the percentage of excess distance in the route. The index is straightforward and easy to understand, but an increase in distance does not necessarily mean the two considered routes are spatially different. The spatial locations of chosen and shortest routes are of particular interest to geographers and urban planners because they imply preferences for certain environments and/or land-uses. Therefore, in addition to RDI, we employ another metric measuring spatial dissimilarity as well as the degree of deviation. In computational geometry, (dis)similarities of two or more trajectories of moving objects can be determined based on the distance between the spatiotemporal coordinates of those trajectories or the area between them (Chen et al., 2011; Hwang et al., 2005; Lin and Su, 2008). Borrowing this approach, we propose an area-based detour index and call it the Normalized Degree of Deviation (NDD). It is calculated as the area between a chosen path and the shortest path in a two-dimensional space, divided by the length of the shortest path. The area contained between two paths indicates how much a chosen route is geographically separated from the shortest path, therefore how far a bicyclist moves away from the shortest path. NDD is also closely associated with excess distance. As the area gets larger, a bicyclist rides farther away from the shortest distance path to complete a trip. The RDI and NDD are significantly correlated ($p = 0.68$) as presented in Table 2.

Fig. 1 displays three examples. The three chosen routes (A, B, and C) have similar RDI values (38.1–39.2%), but their NDDs differ ($275.3 m^2 \sim 653.4\ m^2$) because of their varying levels of detachment from the shortest paths. Our study uses both indexes to measure detour degrees and analyze the contributing factors.

3.3. Explanatory variables

To explain the variations in RDI and NDD for actual bicycle trips, we focus on five categories of explanatory variables: trip characteristics, land-use and streetscape, bicycle facilities, slope gradient, and roadway conditions.

Table 2
Detour extent of bicycle trips: paired-sample *t*-test and Pearson correlation.

Trips (N = 439)	Distance (in meter)		Detour Extent	
	Shortest Possible Routes	Chosen Routes	RDI (%)	NDD (m^2/m)
Mean	3503.6[a]	4118.1[a]	13.5	226.0
Std. Dev.	2228.4	3045.6	16.4	279.2
Minimum	737.2	737.2	0.0	0.0
Maximum	12,029.9	19,063.6	116.1	1994.3
Pearson Bivariate Correlation[a]				
Shortest Dist.	–	0.97	0.35	0.63
Chosen Dist.	–	–	0.50	0.72
RDI (%)	–	–	–	0.68
NDD (m^2/m)	–	–	–	–

RDI = Route Directness Index, NDD = Normalized Degree of Deviation.
[a] p-value < .001 in paired-sample t-test and Pearson correlation.

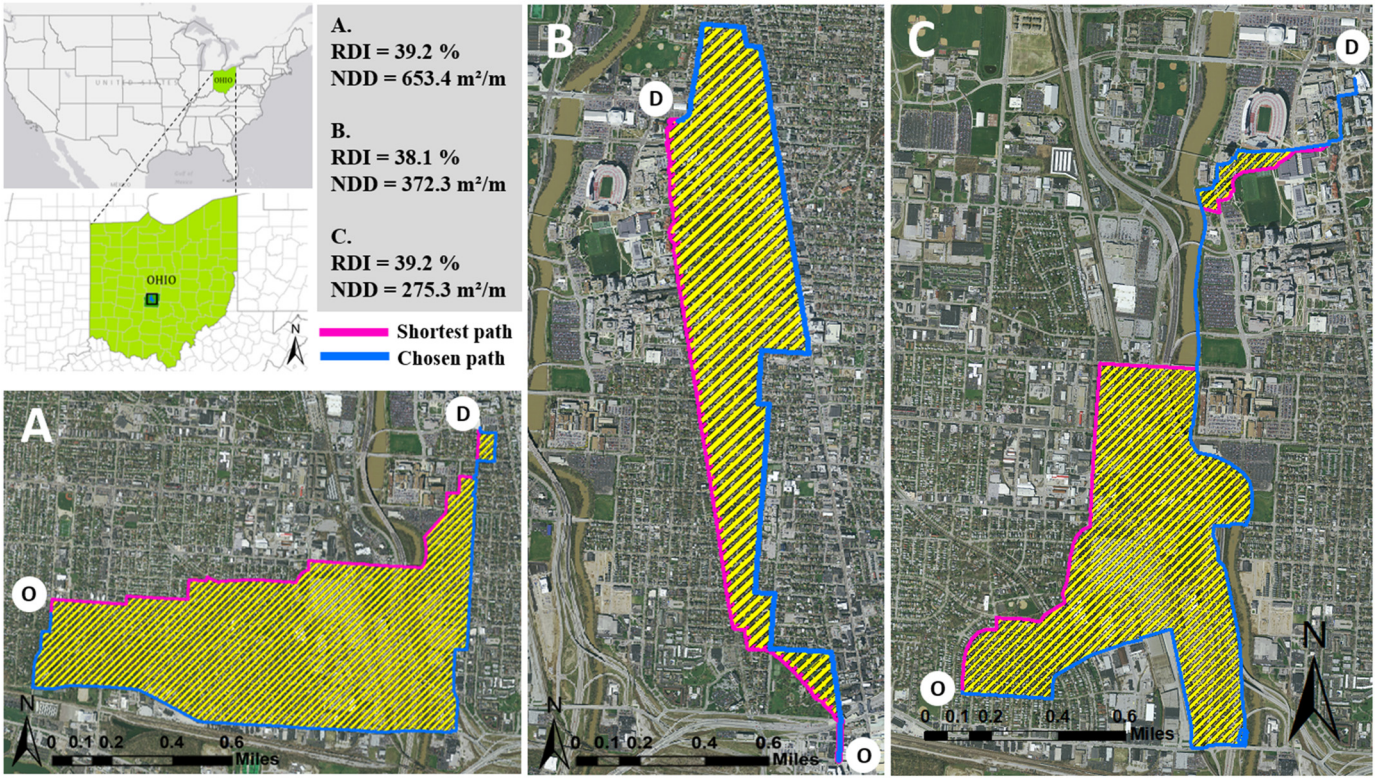

Fig. 1. Examples of chosen paths and shortest distance paths between three different origin and destination pairs in Columbus, Ohio, USA (O: origin, D: destination).

Trip characteristics include the natural logarithm of the shortest path distance, a dummy variable for peak hours (8 am to 9 am and 5 pm to 7 pm), and the number of left turns per kilometer. We evaluate the land-use characteristics using 50 m buffers. These characteristics include the share of commercial, industrial, single-family residential, multifamily residential, and public land-uses as well as the entropy index of the land-use mix. The formula for calculating the entropy index (E) is as follows:

$$E = -\left[\sum_{i=1}^{5} \frac{b_i}{a} ln\left(\frac{b_i}{a}\right)\right]/N$$

where a is the total square meter of land for all five land uses present in buffer, b_i is the area for each land-use type observed in buffer, and N is the number of different land uses present in buffer. For streetscape, we focus on the share of street vegetation using 15 m buffers. To identify the amount of vegetation, we use color near-infrared (CNIR) aerial imagery produced by the US Department of Agriculture (August 2016, 1-m ground sampling distance) (Fig. 2). We classify this raster image using the supervised maximum likelihood classification. In the classification, the pixels in the image are assigned into four land-cover types based on their spectral information: built-up surface, vegetation, bare soil, and water. After error adjustment and quality assessment, we compute the share of vegetation pixels within the 15 m buffer area of each route.

We consider five bicycle facility types: bike paths, bike boulevards, bike lanes, bike routes, and bike sharrows.[3] We calculate the share of each bicycle facility on a route. We also consider the roadway hierarchy ranging from 3 to 8 (secondary arterials, tertiary arterials, collectors, local streets, residential streets, and walkways). We calculate the percentage of the 40–45 mph speed limit for each route, and the numbers of signalized intersections per kilometer. These data were obtained from the Mid-Ohio Regional Planning Commission (MORPC). Distance-weighted average slope gradient and the share of route segments with > 15% upslope are calculated on the basis of US Geological Survey Digital Topographic Model (DTM). Roadway traffic conditions are identified using Google Maps historical typical traffic flows (1: fast, 2: moderate, 3: heavy, 4: very slow). We get this information for a sample of 3500 points across the study area network and then calculate the average traffic for each route.

3.4. Analytical strategy and modeling

We estimate multilevel mixed-effect generalized linear regression models to explain the extent of detours between chosen routes and the shortest routes, as measured by RDI and NDD. We adopt this hierarchical and nested model structure to account for the heterogeneity across individuals in routing preferences. While the level of excess travel is influenced by surrounding built environment and trip characteristics (level 1), it may be in part steered by the personal preferences and habits of a bicyclist (level 2) (Goulias, 2002). A multilevel mixed-effect linear model takes potential intra-individual error correlation into account by estimating random effects for each individual (McCulloch and Searle, 2005). A multilevel regression model with dependent variable Y with a vector of explanatory variables **X** can be written as (Goulias, 2002):

$$Y_{ij} = \beta_{0ij} + \beta_{1ij}X_{ij} + \varepsilon_{ij}$$

[3] The Mid-Ohio Regional Planning Commission (MORPC) classifies bicycle facilities into five types: A bike path is a separated off-street path or bike trail. A bike boulevard is an exclusive bikeway in residential neighborhoods designed to give priority to bicyclists. A bike lane is a portion of a roadway designated by striping and pavement markings for bicycling. A bike route is designated by signs along roadways to indicate their appropriateness for bicycle travel within a normal width vehicular travel lane. A bike sharrow is a variation on shared lanes where arrows or chevrons serve to alert motorists to expect bicyclists.

Fig. 2. Illustration of the method of capturing the amount of street greenery (left: aerial imagery in natural color, center: classified land-cover map with green pixels indicating vegetation, right: the amount of vegetation within the 15 m buffer area). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

where Y_{ij} is the dependent variable of a trip i made by a person j (with i = 1, 2, …, number of trips per person j, j = 1, 2, …., number of persons in the sample). The vector **X** consists of trip-level variables. The error term ε_{ij} is the usual regression error term. The intercept β_{0ij} of the model has a mean γ_0 and the variation around this mean among individuals is depicted by the random variable u_{0j} with $E(u_{0j}) = 0$ and $var.(u_{0j}) = \sigma_{u0}{}^2$ written as:

$$\beta_{0ij} = \gamma_0 + u_{0j}$$

The coefficient β_{1ij} associated with variable X_{ij} can also be assumed as random, where the coefficient has a mean γ_1 and its variation among individuals is described by a random variable u_{1j} with $E(u_{1j}) = 0$ and $var.(u_{1j}) = \sigma_{u1}{}^2$:

$$\beta_{1ij} = \gamma_1 + u_{1j}$$

We estimate two multilevel mixed-effect generalized linear models. Models 1 and 2 analyze RDI and NDD as dependent variables, respectively, with random effects applied to intercepts and coefficients at the individual level.

The explanatory variables are taken as the differences between the characteristics of the chosen path and those of the shortest path, instead of using all chosen or all shortest route characteristics. For example, in case of the slope variable, we use:

> Difference in the average slope = The average slope gradient (%) of the chosen route – the average slope gradient (%) of the shortest matching route

If we consider only the characteristics of the chosen routes or the shortest paths, the resulting outcomes would be misleading (Aultman-Hall et al., 1997; Misra & Watkins, 2016). For instance, low land-use mix around a chosen route may not be meaningful unless the value is significantly lower than that of the shortest path. Taking differences in route characteristics between the chosen and shortest routes helps clarify the reasons as to why riders choose detours over their shortest counterparts.

4. Discussion of results

4.1. Detour degrees

Among the 439 chosen routes, 8.9% (N = 39) are the same as the shortest paths in terms of distance and location, which means that 91.1% of the chosen routes involve detours. Table 2 shows that the chosen routes deviate from the shortest paths significantly in terms of RDI and NDD. As for the RDI, the chosen routes are 13.5% longer on average than the shortest paths. The average distance of the chosen routes is 4118 m, while that of the shortest routes is 3504 m. As for the NDD, the chosen routes deviate from the shortest routes by 226.0 m^2/m on average. The standard deviations of the RDI (16.4%) and NDD (279.2m^2/m) show that there are large variations among trips in terms of their detour levels. These detour indices are positively correlated with the travel distances, indicating that longer trips tend to deviate farther from their shortest paths. The correlations are stronger for NDD (0.63 for shortest distances and 0.72 for chosen distances) than for RDI (0.35 and 0.50).

4.2. Exploratory comparative analysis of built environment conditions

We compare mean route attributes of chosen and shortest routes to explore significant differences between them. Unlike prior studies, which compare variable means between all chosen and shortest routes (Dalton et al., 2015; Krenn et al., 2014; Winters et al., 2010), we divide the 439 routes into three strata in ascending order of detour level (either RDI or NDD):

- Small deviation: 1st stratum (N = 146, RDI mean = 2.0% or NDD mean = 31.1 m^2/m);
- Moderate deviation: 2nd stratum (N = 146, RDI mean = 8.6% or NDD mean = 133.7 m^2/m);
- Large deviation: 3rd stratum (N = 147, RDI mean = 29.9% or NDD mean = 521.2 m^2/m).

Table 3
Mean environmental attributes of GPS-revealed chosen routes and the shortest possible routes for each stratum: paired sample *t*-test results.

Characteristic	Stratum	RDI (Route Directness Index)			NDD (Normalized Degree of Deviation)		
		Shortest routes	Chosen routes	p-value	Shortest routes	Chosen routes	p-value
Land Use (%): Single-Family Residential (50 m)	1	15.44	14.92	0.31	**12.31**	**11.93**	**0.05**
	2	**18.03**	**16.34**	**0.04**	**16.49**	**14.51**	**0.01**
	3	**15.72**	**12.51**	**0.01**	**20.27**	**17.24**	**0.02**
Land Use (%): Multifamily Residential (50 m)	1	**9.75**	**10.31**	**0.10**	**8.60**	**9.17**	**0.01**
	2	**10.93**	**12.66**	**0.01**	**11.05**	**12.15**	**0.05**
	3	**10.45**	**7.93**	**< 0.00**	10.37	10.68	0.69
Land Use (%): Commercial (50 m)	1	14.69	15.12	0.34	13.20	13.21	0.98
	2	**14.86**	**16.34**	**0.05**	13.87	13.64	0.71
	3	**21.00**	**12.84**	**< 0.00**	**23.27**	**17.39**	**< 0.00**
Land Use (%): Industrial (50 m)	1	1.88	1.67	0.25	1.12	1.16	0.74
	2	**2.82**	**2.07**	**0.05**	2.42	2.06	0.20
	3	**2.99**	**3.99**	**< 0.00**	4.10	4.45	0.40
Land Use Mix Index (50 m) (range: 0–1)	1	0.68	0.70	0.11	**0.68**	**0.70**	**0.05**
	2	**0.80**	**0.78**	**0.02**	**0.77**	**0.74**	**0.01**
	3	**0.77**	**0.70**	**< 0.00**	**0.80**	**0.74**	**< 0.00**
Street Greenness (%) (15 m)	1	**33.17**	**34.89**	**< 0.00**	**34.23**	**35.95**	**< 0.00**
	2	38.65	38.19	0.60	37.29	37.92	0.48
	3	**32.46**	**44.44**	**< 0.00**	**32.82**	**43.50**	**< 0.00**
Bicycle Facility (%): Bike Path and Boulevard	1	**10.68**	**12.82**	**0.01**	**11.48**	**14.02**	**< 0.00**
	2	**11.29**	**20.48**	**< 0.00**	**9.37**	**19.25**	**< 0.00**
	3	**7.83**	**44.03**	**< 0.00**	**9.00**	**43.49**	**< 0.00**
Bicycle Facility (%): Bike Lane	1	**1.62**	**2.80**	**< 0.00**	0.71	0.70	0.21
	2	**1.37**	**4.34**	**< 0.00**	**1.99**	**3.95**	**< 0.00**
	3	**0.22**	**3.47**	**< 0.00**	**0.52**	**5.89**	**< 0.00**
Bicycle Facility (%): Bike Route	1	**0.74**	**1.11**	**0.10**	**0.75**	**1.30**	**0.04**
	2	**1.11**	**2.08**	**0.07**	**1.47**	**2.81**	**0.03**
	3	**1.78**	**3.52**	**< 0.00**	**1.40**	**2.56**	**0.02**
Bicycle Facility (%): Bike Sharrow	1	**26.62**	**34.45**	**< 0.00**	**29.05**	**37.33**	**< 0.00**
	2	**16.71**	**31.77**	**< 0.00**	**13.59**	**30.19**	**< 0.00**
	3	**25.35**	**17.63**	**0.01**	**26.02**	**16.71**	**< 0.00**
Number of Left Turns per 1 km	1	**0.52**	**0.50**	**0.02**	**0.59**	**0.56**	**0.05**
	2	**0.53**	**0.46**	**< 0.00**	**0.54**	**0.47**	**< 0.00**
	3	**0.44**	**0.38**	**< 0.00**	**0.38**	**0.31**	**< 0.00**

Note) 1-stratum (n = 146, small detour), 2-stratum (n = 146, moderate detour), and 3-stratum (n = 147, long detour).
Numbers in bold are significant at the 90% confidence level.

In the 1st stratum, chosen and shortest routes are mostly similar, while in the 3rd stratum, chosen and shortest routes are largely dissimilar in terms of their distance and spatial location. This distinction is made to emphasize the statistical differences between chosen and shortest routes in the 3rd and also 2nd strata, rather than in the 1st stratum where chosen and shortest routes mostly overlap each other and therefore few differences are found. We also make this distinction in the expectation of behavioral differences between long detours and shorter detours. We compare route attributes of chosen and shortest routes within each stratum using paired sample *t*-tests.

The results presented in Table 3 show that there are significant environmental differences between chosen and shortest routes, in terms of land-use and land-use mix, street greenery, bicycle facility types, street classes, and roadway physical attributes. Surrounding land-use patterns along chosen routes, specifically single-family residential, commercial land use, land-use mix and street greenness, present different characteristics from the shortest paths. The routes in the third stratum (large deviation) display higher contrasts between chosen and shortest routes than other strata. The percentage of single-family residential land use significantly decreases from shortest to chosen routes, especially in the case of long detours (3rd stratum). The magnitude of this contrast is less significant or minimal in case of small (1st stratum) to moderate detours (2nd stratum). The chosen routes are likely to have lower levels of land-use mix (i.e. less diversity in land use), and this contrast is more noticeable in longer detours. Street greenness increases from shortest to chosen routes, especially following large detours. No notable change is found for moderate detours. The share of commercial land uses shows a slightly different pattern. Long detour routes pass through areas with significantly lower proportions of commercial areas than do their shortest counterparts, while shorter detour routes pass through areas with slightly higher proportions of commercial land uses than do their shortest counterparts. (See Table 3.)

The results on specific bicycle facility types, street classes, and roadway traffic also vary across the strata, but they are intriguing in terms of the tradeoffs between route directness and comfortable rides. Long detour routes are more likely to be on off-street bicycle paths as compared with small to moderate detour routes. About 44% of long detours are conducted on bicycle paths and boulevards, while bicycle sharrows are only partially used (16.7% for NDD stratum to 17.6% for RDI stratum). In contrast, short detour routes are largely on roads with bicycle sharrows (34.5% to 37.3%), while only 12.8% to 14.0% on bicycle paths and boulevards. Although bicycle sharrows are only painted markings on roads for lane sharing, it seems that this relatively direct and dense facility is likely to be preferred for shorter distances rather than off-street bicycle paths whose availability is limited. Similarly, small and moderate detour routes are more likely to be on secondary/tertiary and collector class roads, while long detour routes are more associated with lower-class streets, such as passageways and walkways. As a result, small and moderate detour routes are likely to face more vehicular traffic and signalized intersections than long detour routes, because they include major roadways more than large detour routes do. The share of commercial land use can be explained in the same vein. The increase in the percentage of commercial land uses for chosen routes in the 2nd stratum may be associated with the higher use of secondary/tertiary and collector-level roadways and/or bicycle sharrows, because commercial buildings tend to be located near major roadways and these roads are likely to have bicycle sharrows for the protection of bicyclists. In the 3rd stratum, the share of commercial

Table 3
Mean environmental attributes of GPS-revealed chosen routes and the shortest possible routes for each stratum: paired sample *t*-test results (Cont.)

Characteristic	Stratum	RDI (Route Directness Index)			NDD (Normalized Degree of Deviation)		
		Shortest routes	Chosen routes	p-value	Shortest routes	Chosen routes	p-value
Distance-weighted Average Street Class (range:3–8)	1	**6.49**	**6.31**	**< 0.00**	**6.65**	**6.53**	**< 0.00**
	2	**6.52**	**6.17**	**< 0.00**	**6.67**	**6.36**	**< 0.00**
	3	**6.25**	**6.82**	**< 0.00**	**5.95**	**6.41**	**< 0.00**
Secondary (Broach and Bigazzi, 2017) and tertiary (Broach et al., 2012) road class (%)	1	**17.27**	**19.83**	**0.02**	**14.68**	**15.65**	**0.02**
	2	**19.15**	**24.19**	**0.02**	**16.40**	**20.72**	**< 0.00**
	3	**24.54**	**13.95**	**< 0.00**	**29.61**	**21.53**	**0.01**
Local collector road class (Buehler and Dill, 2016; Buehler and Handy, 2008) (%)	1	**12.22**	**14.65**	**0.02**	**10.43**	**12.81**	**0.01**
	2	**8.14**	**14.06**	**< 0.00**	**7.35**	**12.18**	**< 0.00**
	3	9.06	9.13	0.96	11.60	12.85	0.44
Residential road class (Casello and Usyukov, 2014) (%)	1	40.15	39.22	0.49	39.71	40.97	0.26
	2	**40.35**	**33.85**	**< 0.00**	39.76	36.94	0.18
	3	**37.88**	**31.24**	**< 0.00**	**38.92**	**26.65**	**< 0.00**
Service and walkway road class (Chen et al., 2011) (%)	1	**30.36**	**26.30**	**< 0.00**	**35.18**	**30.56**	**< 0.00**
	2	**32.36**	**27.90**	**0.01**	**36.48**	**30.16**	**< 0.00**
	3	**28.52**	**45.68**	**< 0.00**	**19.87**	**38.97**	**< 0.00**
Links with Speed Limit between 40 and 45 mph (%)	1	0.83	0.74	0.71	0.52	0.54	0.81
	2	**2.26**	**0.87**	**0.01**	2.05	1.45	0.24
	3	**3.52**	**1.34**	**< 0.00**	**3.98**	**0.96**	**< 0.00**
Links with Upslope > 15% (%)	1	0.49	0.42	0.25	0.55	0.49	0.40
	2	**0.30**	**0.22**	**0.09**	**0.23**	**0.14**	**0.03**
	3	**0.25**	**0.18**	**0.08**	**0.26**	**0.19**	**0.06**
Distance-weighted Average Slope (gradient, %)	1	0.03	0.04	0.51	0.01	0.02	0.26
	2	−0.01	0.01	0.17	0.02	0.03	0.38
	3	0.04	0.03	0.26	0.03	0.03	0.45
Number of Signalized Intersections per 1 km	1	**0.42**	**0.44**	**0.01**	**0.43**	**0.45**	**0.02**
	2	**0.36**	**0.41**	**0.03**	**0.31**	**0.39**	**< 0.00**
	3	**0.36**	**0.27**	**< 0.00**	**0.39**	**0.28**	**< 0.00**
Distanced-weighted Average Traffic (range: 0–4)	1	**1.04**	**1.12**	**< 0.00**	**0.99**	**1.06**	**< 0.00**
	2	**0.97**	**1.03**	**0.01**	**0.94**	**1.05**	**< 0.00**
	3	**1.04**	**0.85**	**< 0.00**	**1.11**	**0.88**	**< 0.00**

Note) 1-stratum (n = 146, small detour), 2-stratum (*n* = 146, moderate detour), and 3-stratum (*n* = 147, long detour).
Numbers in bold are significant at the 90% confidence level.

land uses significantly decreases.

Unlike the varied findings on surrounding land-uses, bicycle facilities, street classes, and traffic, we find that some route characteristics, such as high-speed limits, roadway upslope, and left turns, are consistently avoided along the chosen routes. The percentages of these significant obstacles, such as speed limits of 40 to 45 mph, steep upslopes of > 15% gradient, and frequent left turn, significantly decrease from shortest to chosen routes in almost all strata. No significant difference is found in terms of the average slope gradients.

Overall, the exploratory analyses reveal that there are two groups of variables. While there are variables consistently preferred (e.g. bicycle paths, bicycle lanes, street greenery) or avoided (e.g. high-speed limit, upslopes of > 15% gradient, left turn at intersections, single-family residential use) along the chosen routes, there are variables showing the tradeoffs between route directness and comfortable rides (e.g. bicycle sharrows, commercial land use, secondary/tertiary roadways, local collector roadways, walkway/service roads, and average traffic). Prior studies on bicycle route choice acknowledge these tradeoffs. People who are sensitive to travel distances tend to choose more direct routes at the cost of safety or comfort, while those who seek pleasurable rides are more willing to accept longer routes with bicycle-supportive facilities or less traffic (Sener et al., 2009).

4.3. Multilevel regression analysis

We estimate two multilevel mixed-effect generalized regression models, with RDI and NDD as dependent variables, respectively, as functions of trip and route characteristics. All explanatory variables represent the differences between the values of the chosen routes and those of the shortest routes. For example, if the chosen route contains two left turns per kilometer and the shortest route has three, the corresponding variable is equal to −1 (Aultman-Hall et al., 1997; Broach and Bigazzi, 2017), implying that the rider makes one less left turn per kilometer by detouring. The coefficients consist of two types: fixed and random effects. The fixed effect is the mean effect of a variable, while the random effect is the variance of a fixed effect across individuals.

4.3.1. Trip characteristics

The estimation results presented in Table 4 show that the deviation from the shortest path increases with trip distances. Both RDI and NDD models confirm that longer trips are more likely to have deviations than shorter trips. Long trips usually have higher probabilities of encountering intermittent and traffic-related obstacles, thus may lead bicyclists to take longer detours. As bicyclists facing longer travel distances will experience these conditions for longer durations, they may seek comfortable riding conditions more than short-distance riders. It is also possible that bicyclists riding longer distances are likely to deviate simply because there is more opportunity to deviate on longer trips than shorter trips, such as street intersections. Having a variety of alternative routes for deviation may increase the likelihood of taking a different route that provides better conditions than the shortest path. However, having more alternative routes may also have the effect of reducing large deviations because well-connected streets with dense intersections provide a larger pool of route options for riders to take. In this case, it might be easier to find a small or moderate detour route with satisfying conditions in the vicinity of the shortest path (Broach and Bigazzi, 2017). We tested this relationship by including intersection density within 100m buffers around the shortest paths as a proxy for the availability of deviation opportunities. We found no significant effects pertinent to this relationship. Therefore, we do not present the models including this variable here for brevity. In order to better clarify the effect of travel distance on excess travel, future research may need to control for the number (and maybe quality) of viable alternative routes between a pair of origin and destination.

Table 4
Results of multilevel mixed-effect linear regression models with individual-level random coefficients (dependent variables: RDI and NDD, respectively).

Explanatory variables: *difference (=Chosen route – shortest route)*	RDI (%) (Route Directness Index)			NDD (m^2/m) (Normalized degree of deviation)		
	Coef.	*t*-stat	p-value	Coef.	*t*-stat	p-value
Trip characteristics						
ln(Length of the Shortest Path)[a]	**0.114**	**2.47**	**0.01**	**0.264**	**5.01**	**< 0.00**
Peak time trip (peak = 1, non-peak = 0)[a]	0.083	1.16	0.25	**0.130**	**2.34**	**0.02**
Number of Left Turns per 1 km	−0.064	−1.06	0.30	−0.094	−1.72	0.11
Land Use (%, 50 m buffer)						
Commercial	**−0.165**	**−2.53**	**0.01**	**−0.093**	**−1.72**	**0.09**
Industrial	0.058	1.21	0.23	0.031	0.73	0.46
Single-Family Residential	−0.096	−1.39	0.17	**−0.148**	**−1.95**	**0.06**
Multifamily Residential	0.016	0.28	0.78	**0.163**	**3.28**	**< 0.00**
Land Use Mix Index (range: 0–1)	−0.025	−0.51	0.61	**−0.091**	**−2.36**	**0.02**
Street Greenery (%, 15 m buffer)	0.079	1.09	0.28	**0.136**	**2.24**	**0.03**
Bicycle Facilities (%)						
Bike Path & Boulevard (%)	**0.274**	**3.13**	**< 0.00**	**0.141**	**1.86**	**0.07**
Bike Lane (%)	**0.072**	**1.82**	**0.07**	**0.123**	**3.70**	**< 0.00**
Bike Route (%)	0.045	1.23	0.22	0.009	0.31	0.76
Bike Sharrow (%)	**0.151**	**2.31**	**0.02**	**0.112**	**2.04**	**0.04**
Roadway Slope						
Dist.-weighted Average Slope (gradient, %)	−0.109	−0.97	0.34	−0.020	−0.66	0.51
Segments with Upslope > 15% (%)	0.002	0.06	0.95	0.029	1.13	0.26
Roadway Conditions and Classes						
Dist.-weighted Average Road Class (range: 3–8)[1]	**0.180**	**2.19**	**0.03**	**0.181**	**1.90**	**0.06**
Share of Segments with Speed Limit 40–45 mph	**−0.109**	**−2.36**	**0.02**	**−0.170**	**−4.20**	**< 0.00**
Number of Signalized Intersections per 1 km	0.067	1.04	0.30	0.060	1.14	0.26
Dist.-weighted Average Traffic (range: 1–4)[1]	0.057	0.96	0.34	−0.014	−0.27	0.79
Intercept	**−0.173**	**−3.25**	**< 0.00**	**−0.239**	**−4.80**	**< 0.00**
Variance(residual)[2]	**0.366**	**12.25**	**< 0.00**	**0.200**	**11.37**	**< 0.00**
Variance(subjects)	0.010	0.67	0.50	0.011	0.66	0.51
Variance(ln(length of the shortest path)	–	–	–	**0.036**	**1.67**	**0.09**
Variance(number of left turns per 1 km)	**0.040**	**1.74**	**0.08**	**0.060**	**2.15**	**0.03**
Variance(single-family residential)	**0.052**	**1.91**	**0.06**	**0.102**	**2.27**	**0.02**
Variance(bike path & boulevard)	**0.101**	**2.29**	**0.02**	**0.075**	**2.32**	**0.02**
Variance(dist.-weighted average road class)	–	–	–	**0.129**	**2.14**	**0.03**
Variance(dist.-weighted average slope)	**0.313**	**2.51**	**0.01**	–	–	–
−2 Restricted Log Likelihood	1012.99			818.69		
Number of Observations (trip)	439			439		
Number of Observations (bicyclist)	73			73		

Note) The coefficients are standardized. Significant coefficients at the 90% confidence level are in bold.
[a] not differenced values
[1] For roadway classes, '3' means secondary arterial, while '8' indicates footway. For typical traffic, '1' means little traffic and '4' means heavy traffic on weekday rush hours.
[2] Variance estimates come with Wald Z test statistics, instead of t-statistics.

The NDD model shows that time of travel has a significant impact on detour levels. The extent of detour is likely to increase during peak periods (8 AM–9 AM or 5 PM–7 PM), possibly because of heavier vehicular traffic on major roadways and intersections. Bicyclists may be more likely to seek quiet routes to avoid stress related to large vehicular volumes during these periods (Sener et al., 2009; Ta et al., 2016).

Longer detours are likely to involve fewer left turns per kilometer compared with the shortest paths, although this relationship is only weakly significant for NDD ($p = 0.11$). Left turns usually become more cumbersome when combined with high traffic volumes, signalized intersections and lack of bikeways (Broach et al., 2012). We find that the estimated random parameters of the left-turn variable are significant, indicating that individual tendencies towards fewer left turns vary significantly.

4.3.2. Land use

Surrounding land-use patterns have a significant impact on detour extent. Specifically, detour routes are characterized by significantly fewer commercial activities as compared to the shortest paths, with other factors controlled for in the models. Commercial land-use is associated with mixed traffic conditions, including irregular vehicle movements and pedestrian flows waiting for or using services. This mixed street environment may negatively influence bicyclists' perception of safety and comforts (Zacharias and Zhang, 2016), therefore contributing to detours.

Other land-use types also significantly correlate with detour extent, but only for NDD, and not for RDI. This may be due to their different focus: route deviation versus incremental distances. As NDD reflects the dissimilarity of the spatial locations of the chosen and shortest routes, it may be more correlated with differences in land-use patterns than RDI. The results show that longer detour routes involve significantly less single-family residential land-use and lower land-use mix as compared to the shortest paths. Single-family residential areas are often characterized by lower street connectivity and limited bicycle lanes. Although riding in such areas may be safer due to lower vehicular traffic volumes (Stinson and Bhat, 2003), going through unfamiliar single-family residential blocks might be unappealing for utilitarian bicyclists. The relative increase of multifamily dwellings in detour routes may indicate a preference towards riding in denser neighborhoods rather than in single-family residential blocks. Further research is needed to identify the attributes of these land uses that may affect bicycle routing preferences. Land-use mix decreases with long detours, because individuals often take detours to avoid certain land-uses (i.e. commercial and single-family residential uses) and to ride on continuous bicycle facilities. The level of street greenness is also associated with detours in terms of NDD. Longer detour trips are surrounded by

more vegetation, at least in summer and early autumn, when most of the data were collected. Recent studies pay attention to the positive effects of greenways and street vegetation as they relate to active transport modes (Dill et al., 2017; Chen et al., 2017). As greenery provides natural views and shades and often serves as a buffer from other activities, it may be preferred by detouring bicyclists.

4.3.3. Bicycle facilities

The shares of bicycle facilities are positively correlated to the increase in the magnitude of detour, except for bike route. Off-street bike paths and bike boulevards provide safe and pleasant riding environments, and therefore individuals accept riding longer distances to take advantage of these facilities. Because their availability is spatially limited, one needs to trade in extra distances to approach these facilities. Another explanation is that some riders intentionally ride on off-street bikeways for a longer distance than actually needed, seeking benefits other than fulfilling a utilitarian trip itself, such as exercises or riding for fun (Broach and Bigazzi, 2017). While the percentage share of bicycle paths and boulevards has a significant fixed effect on excess travel, this effect may become higher or mitigated by individual bicyclists' tendencies, as reflected by the significant variance of the fixed effect. Wergin and Buehler (2017) provides a case related to this finding that regular users of bike-share systems prefer to ride on bicycle trails significantly longer than occasional users. In addition to bicycle paths and boulevards, our results also indicate that bicycle lanes and sharrows contribute to deviations from the shortest paths. Although bike routes are by definition generally of better quality than bike sharrows (designated signs versus painted markings on roads), it seems that they do not offer enough advantages for bicyclists to take a detour, possibly because of their limited number and continuity as compared to bike sharrows.

4.3.4. Roadway slope

Our results indicate that the average slope and the percentage of segments with more than a 15% slope gradient have no significant impact on the magnitude of detour. We also tested the effects of variations in roadway slopes and found no statistically significant relationship. This may be partly due to the geographic context of the study area: a predominantly flat topography, with small frequency and magnitude of upslopes. Some prior studies find significant associations between slopes and bicycle route preferences, especially in study areas with hilly topographies, such as Portland, Seattle, and Atlanta (Broach et al., 2012; Chen et al., 2017; Misra and Watkins, 2017). Slopes of > 15% may cause a detour when faced en route, as we found in the exploratory analysis, but they may not correlate to the overall extent of detour. Instead, our results reveal that there is significant heterogeneity among bicyclists in terms of tolerance towards roadway slopes, as reflected by the significant random effect for the average slope variable in the RDI model. A hilly section may constitute a significant barrier for some individuals, for example, for women and for commuting trips (Hood et al., 2011), but minor or moderate gradients may not necessarily lead to detours for multipurpose bicyclists (Sener et al., 2009).

4.3.5. Roadway conditions and classes

Roadway conditions have varied impacts on detour extent. Roadways with high-speed limits (40 to 45 mph) lead to larger deviations, whereas average peak-hour traffic does not have any significant relationship with detour behavior. We analyzed the effect of the percentage of heavy traffic volumes and found no statistically significant relationship. The literature presents mixed findings regarding the impact of traffic volumes on bicycle route choices. Sener et al. (2009) find a stated aversion to high traffic volumes, and Broach et al. (2012) find that bicyclists are not likely to choose heavy traffic routes with no bicycle facilities around. In contrast, Wang et al. (2016) find that the level of traffic stress is not always useful in explaining commuter bicyclist behavior because its effect is uneven among bicyclists. We find that bicyclists are likely to take detours to evade speeding vehicles next to them, while traffic volume by itself is not related to deviation from the shortest paths.

Average roadway classes have significant impacts on detours. Larger deviations are likely to occur on lower class roadways. This is consistent with prior studies, where bicyclists prefer local streets to arterial roads when other factors are taken into account (Broach et al., 2012; Stinson and Bhat, 2003). Lower class passages and service ways that cut through a large lot can be good complements to the existing street network that help improve bicycle routing efficiency. The preference towards lower class streets along the way, however, may vary significantly across individual bicyclists, as indicated by its significant random parameter in the case of NDD.

Overall, the regression results suggest that NDD is a better tool as compared with RDI from statistical and practical viewpoints. The NDD model attains significant improvement in the restricted log-likelihood value (818.693) as compared with the RDI model (1012.991). The NDD index has significant associations with more explanatory variables representing peak-time travel, the amount of street greenery, and the shares of different land use types (i.e. single-family residential use, multi-family residential use, and land-use mix). Most of the significant variables in the RDI model remain significant in the NDD model. These observations indicate the relevance of using the NDD for assessing the linkages between route characteristics and the degree of excess travel.

5. Conclusions

The present study used individual bicyclists' GPS trajectories to identify the extent of their trip deviations from the shortest paths and the factors associated with these detours. We conducted a comparative analysis of riding environments for small-, moderate-, and large-detour routes to explore significant correlates to detour behavior. We then estimated multilevel mixed-effect linear regression models to explain the variation in detour degrees, as measured by two indices (Route Directness Index and Normalized Degree of Deviation), as a function of route-level environmental factors.

We find that 91.1% of the sampled 439 bicycle trips include a detour and are on average 13.5% longer than their shortest counterparts. Bicyclists who travel longer distances during peak periods are more likely to take detours. This relatively low bicycling efficiency in peak-hour commute may weaken the potential of bicycling as an alternative to motorized modes. Deviations from the shortest paths are closely related to the availability of bicycle facilities. Our study suggests that an increase in bicycle facilities may increase travel efficiency. Policy interventions, such as strategic installment of bicycle infrastructure on major commuter roads, would help reduce detours, especially during peak periods. We find that in order to keep individuals on their shortest paths, providing bicycle lanes or sharrows may be good alternatives when providing a separated bicycle path is not a feasible option.

Land-use patterns are closely associated with bicycle riding environments. Our study suggests that bicyclists take detours in order to avoid single-family residential blocks. Lower street connectivity and lack of bicycle infrastructure may make it hard for bicyclists to navigate efficiently through these blocks. Our results also suggest that bicyclists avoid areas with commercial land-use. To encourage bicycling, mixed-use developments should be accompanied by bicycle facilities or other protective amenities on busy streets, such as street furniture separating pedestrians and bicyclists. Consistent with existing studies (Chen et al., 2017; Krenn et al., 2014), our study finds a positive association between street greenery and route preferences. Synergistic use of street trees and bicycle facilities may contribute to better riding experiences. Higher street densities complemented by local streets with low to moderate speed limits may also help keep individuals on their shortest paths.

There are substantial variations in the impacts of certain route attributes on detour levels due to unobserved personal traits. For instance, the multilevel mixed-effect models provide additional insights

that there is significant heterogeneity in bicyclists' preferences towards bicycle paths and lower class streets. The tendency to take detours to avoid hills or left turns also varies significantly among individuals. Buehler and Dill (2016) note that the demographic and socioeconomic status of individuals may mediate the impact of the built environment on bicycle routing behavior. Future research may experiment with individual characteristics information to help understand the nature of individual heterogeneity. To conduct more rigorous inference of the causal effect of the environment, information on individual perceptions and longitudinal survey design may be needed (Handy et al., 2014).

This study could not account for the influence of temporary events or changes in route conditions (such as road construction, temporary closures due to crashes, etc.) at the time of the trip, which might have affected individuals' route choices. This study focused on a subset of bicyclists (i.e. university bicyclists), who may not be representative of urban bicyclists. Despite these limitations, the results still extend our knowledge of bicycle routing behavior, and the methodology developed can be applied elsewhere.

Acknowledgments

The authors would like to thank the two anonymous reviewers for their constructive comments and suggestions. This research was partially supported by the NEXTRANS Center, Purdue University under Grant No. DTRT12-G-UTC05 of the U.S. Department of Transportation.

References

Akar, G., Clifton, K., 2009. Influence of individual perceptions and bicycle infrastructure on decision to bike. Trans. Res Rec. J. Trans. Res. Board 2140, 165–172. https://doi.org/10.3141/2140-18.

Aultman-Hall, L., Fred, L., Baetz, B., 1997. Analysis of bicycle commuter routes using geographic information systems: Implications for bicycle planning. Transp. Res. Rec. 1578, 102–110. https://doi.org/10.3141/1578-13.

Broach, J., Bigazzi, A.Y., 2017. Existence and use of low-pollution route options for observed bicycling trips. Trans. Res. Rec. J. Trans. Res. Board 2662, 152–159. https://doi.org/10.3141/2662-17.

Broach, J., Dill, J., Gliebe, J., 2012. Where do cyclists ride? A route choice model developed with revealed preference GPS data. Transp. Res. A Policy Pract. 46 (10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005.

Buehler, R., Dill, J., 2016. Bikeway networks: a review of effects on cycling. Transp. Rev. 36 (1), 9–27. https://doi.org/10.1080/01441647.2015.1069908.

Buehler, T., Handy, S., 2008. Fifty years of bicycle policy in Davis, California. Trans. Res. Rec. J. Trans. Res. Board 2074, 52–57. https://doi.org/10.3141/2074-07.

Casello, J., Usyukov, V., 2014. Modeling cyclists' route choice based on GPS data. Trans. Res. Rec. J. Trans. Res. Board 2430, 155–161. https://doi.org/10.3141/2430-16.

Chen, J., Shaw, S., Yu, H., Lu, F., Chai, Y., Jia, Q., 2011. Exploratory data analysis of activity diary data: a space–time GIS approach. J. Transp. Geogr. 19 (3), 394–404. https://doi.org/10.1016/j.jtrangeo.2010.11.002.

Chen, P., Shen, Q., Childress, S., 2017. A GPS data-based analysis of built environment influences on bicyclist route preferences. Int. J. Sustain. Trans. 24, 1–14. https://doi.org/10.1080/15568318.2017.1349222.

Dalton, A., Jones, A., Panter, J., Ogilvie, D., 2015. Are GIS-modelled routes a useful proxy for the actual routes followed by commuters? J Transp Health 2, 219–229. https://doi.org/10.1016/j.jth.2014.10.001.

De Vos, J., Mokhtarian, P., Schwanen, T., Van Acker, V., Witlox, F., 2016. Travel mode choice and travel satisfaction: bridging the gap between decision utility and experienced utility. Transportation 43 (5), 771–796. https://doi.org/10.1007/s11116-015-9619-9.

Dill, J., Neal, M., Luhr, G., Adkins, A., 2017. Perceptions of green streets and their influence on walking (no. 17-06718). In: Transportation Research Board Annual Meeting, . http://docs.trb.org/prp/17-06718.pdf.

Duncan, M., Mummery, W., 2007. GIS or GPS? A comparison of two methods for assessing route taken during active transport. Am. J. Prev. Med. 33 (1), 51–53. https://doi.org/10.1016/j.amepre.2007.02.042.

Goulias, K., 2002. Multilevel analysis of daily time use and time allocation to activity types accounting for complex covariance structures using correlated random effects. Transportation 29 (1), 31–48. https://doi.org/10.1023/A:1012973004152.

Handy, S., Weston, L., Mokhtarian, P., 2005. Driving by choice or necessity? Transp. Res. A Policy Pract. 39 (2), 183–203. https://doi.org/10.1016/j.tra.2004.09.002.

Handy, S., Van Wee, B., Kroesen, M., 2014. Promoting cycling for transport: research needs and challenges. Transp. Rev. 34 (1), 4–24. https://doi.org/10.1080/01441647.2013.860204.

Hood, J., Sall, E., Charlton, B., 2011. A GPS-based bicycle route choice model for San Francisco, California. Trans. Lett. 3 (1), 63–75. https://doi.org/10.3328/TL.2011.03.01.63-75.

Hwang, J., Kang, H., Li, K., 2005. Spatio-temporal similarity analysis between trajectories on road networks. In: Akoka (Ed.), Perspectives in Conceptual Modeling. ER 2005. Lecture Notes in Computer Science. Vol. 3770. Springer, Berlin, Heidelberg, pp. 280–289. https://doi.org/10.1007/11568346_30.

Khatri, R., Cherry, C., Nambisan, S., Han, L., 2016. Modeling route choice of utilitarian bikeshare users with GPS data. Trans. Res. Rec. J. Trans. Res. Board 2587, 141–149. https://doi.org/10.3141/2587-17.

Krenn, P., Oja, P., Titze, S., 2014. Route choices of transport bicyclists: a comparison of actually used and shortest routes. Int. J. Behav. Nutr. Phys. Act. 11 (1), 31–37. https://doi.org/10.1186/1479-5868-11-31.

Lin, B., Su, J., 2008. One way distance: for shape based similarity search of moving object trajectories. GeoInformatica 12 (2), 117–142. https://doi.org/10.1007/s10707-007-0027-y.

McCulloch, C., Searle, S., 2005. Linear mixed models (LMMs). In: Shewhart, W., Wilks, S. (Eds.), Generalized, Linear, and Mixed Models. Wiley, pp. 156–186. https://doi.org/10.1002/0471722073.ch6. Chapter 6.

Menghini, G., Carrasco, N., Schüssler, N., Axhausen, K., 2010. Route choice of cyclists in Zurich. Transp. Res. A Policy Pract. 44 (9), 754–765. https://doi.org/10.1016/j.tra.2010.07.008.

Misra, A., Watkins, K., 2017. Modeling cyclists' willingness to deviate from shortest path using revealed preference data (no. 17-03556). In: Transportation Research Board Annual Meeting, . https://trid.trb.org/view.aspx?id=1438418.

Papinski, D., Scott, D., 2013. Route choice efficiency: an investigation of home-to-work trips using GPS data. Environ Plan A 45 (2), 263–275. https://doi.org/10.1068/a44545.

Plazier, P., Weitkamp, G., van den Berg, A., 2017. "Cycling was never so easy!" An analysis of e-bike commuters' motives, travel behaviour and experiences using GPS-tracking and interviews. J. Transp. Geogr. 65, 25–34. https://doi.org/10.1016/j.jtrangeo.2017.09.017.

Sener, I., Eluru, N., Bhat, C.R., 2009. An analysis of bicycle route choice preferences in Texas, US. Transportation 36 (5), 511–539. https://doi.org/10.1007/s11116-009-9201-4.

Standen, C., Crane, M., Collins, A., Greaves, S., Rissel, C., 2017. Determinants of mode and route change following the opening of a new cycleway in Sydney, Australia. J Transp. Health 4, 255–266. https://doi.org/10.1016/j.jth.2016.10.004.

Stinson, M., Bhat, C., 2003. Commuter bicyclist route choice: Analysis using a stated preference survey. Trans. Res Rec. J Trans. Res. Board 1828, 107–115. https://doi.org/10.3141/1828-13.

Ta, N., Zhao, Y., Chai, Y., 2016. Built environment, peak hours and route choice efficiency: an investigation of commuting efficiency using GPS data. J. Transp. Geogr. 57, 161–170. https://doi.org/10.1016/j.trangeo.2016.10.005.

Tilahun, N., Levinson, D., Krizek, K., 2007. Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. A Policy Pract. 41 (4), 287–301. https://doi.org/10.1016/j.tra.2006.09.007.

Wang, H., Palm, M., Chen, C., Vogt, R., Wang, Y., 2016. Does bicycle network level of traffic stress (LTS) explain bicycle travel behavior? Mixed results from an Oregon case study. J. Transp. Geogr. 57, 8–18. https://doi.org/10.1016/j.jtrangeo.2016.08.016.

Wergin, J., Buehler, R., 2017. Where do bikeshare bikes actually go? Analysis of capital bikeshare trips with GPS data. Trans. Res. Rec. J. Trans. Res. Board 2662, 12–21. https://doi.org/10.3141/2662-02.

Winters, M., Teschke, K., Grant, M., Setton, E.M., Brauer, M., 2010. How far out of the way will we travel? Built environment influences on route selection for bicycle and car travel. Trans. Res. Rec. J. Trans. Res. Board 2190, 1–10. https://doi.org/10.3141/2190-01.

Wuerzer, T., Mason, S., 2015. Cycling willingness: investigating distance as a dependent variable in cycling behavior among college students. Appl. Geogr. 60, 95–106. https://doi.org/10.1016/j.apgeog.2015.03.009.

Zacharias, J., Zhang, R., 2016. Revealed bicyclist route preferences and street conditions. Trans. Res. Rec. J. Trans. Res. Board 2587, 17–22. https://doi.org/10.3141/2587-03.

Zhu, S., Levinson, D., 2015. Do people use the shortest path? An empirical test of Wardrop's first principle. PLoS One 10 (8), e134322. https://doi.org/10.1371/journal.pone.0134322.

Zimmermann, M., Mai, T., Frejinger, E., 2017. Bike route choice modeling using GPS data without choice sets of paths. Trans. Res. Part C Emerg. Technol. 75, 183–196. https://doi.org/10.1016/j.trc.2016.12.009.