The missing link: bicycle infrastructure networks and ridership in 74 US cities

Article *in* Transportation · November 2014

DOI: 10.1007/s11116-014-9538-1

2 authors:

Jessica E.T. Schoner
University of Minnesota Twin Cities

David Levinson
The University of Sydney

Transportation
DOI 10.1007/s11116-014-9538-1

The missing link: bicycle infrastructure networks and ridership in 74 US cities

Jessica E. Schoner · David M. Levinson

Abstract Cities promote strong bicycle networks to support and encourage bicycle commuting. However, the application of network science to bicycle facilities is not very well studied. Previous work has found relationships between the amount of bicycle infrastructure in a city and aggregate bicycle ridership, and between microscopic network structure and individual tripmaking patterns. This study fills the missing link between these two bodies of literature by developing a standard methodology for measuring bicycle facility network quality at the macroscopic level and testing its association with bicycle commuting. Bicycle infrastructure maps were collected for 74 Unites States cities and systematically analyzed to evaluate their network structure. Linear regression models revealed that connectivity and directness are important factors in predicting bicycle commuting after controlling for demographic variables and the size of the city. These findings provide a framework for transportation planners and policymakers to evaluate their local bicycle facility networks and set regional priorities that support nonmotorized travel behavior, and for continued research on the structure and quality of bicycle infrastructure and behavior.

Keywords Bicycling · Travel behavior · Networks

Introduction

Cities are increasingly promoting cycling as a valuable transportation alternative to driving, arguing that mode shift away from the private auto provides region-wide congestion,

J. E. Schoner (✉) · D. M. Levinson
Department of Civil, Environmental, and Geo-Engineering, University of Minnesota, 500 Pillsbury Drive SE, Minneapolis, MN 55455, USA
e-mail: schon082@umn.edu

D. M. Levinson
e-mail: dlevinson@umn.edu

Published online: 27 July 2014

environmental, and health benefits (Federal Highway Administration 2012). Between 1999 and 2011, total United States (US) federal and state government funding on bicycling and pedestrian infrastructure exceeded $7 billion. The US Federal Highway Administration, FHWA (2012) completed the Nonmotorized Transportation Pilot Program in 2012, which allocated $25 million to each of four pilot cities over 5 years to measure the impacts of new infrastructure on mode shift to bicycling and walking. Many projects are explicitly targeted at closing “gaps” in bike routes to form a more cohesive cycling network (Byers 2002). This idea of a “network” of bicycle routes connecting the region is an important indicator of the shift in transportation priorities from auto dominance to accommodation of nonmotorized modes. While bicycles are permitted to use most components of the road network in the US, bicycle-specific infrastructure provides safe, comfortable routes that many bicyclists prefer over sharing travel lanes with motorized vehicles. The network formed by bicycle-specific infrastructure in any given city, however, is not as expansive or complete as the underlying road network, so cyclists often have to detour to stay on dedicated facilities or else share the roadway with cars and trucks in order to complete their trip.

Numerous studies have identified relationships between rates of bicycling and provision of infrastructure (Nelson and Allen 1997; Dill and Carr 2003; Parkin et al. 2007; Buehler and Pucher 2011). Qualitative descriptions of bicycle facility network characteristics and planning priorities are prevalent in practice (American Association of State Highway and Transportation Officials 2012). However, the application of quantified network structure indicators to bicycle infrastructure design and research is nascent. This paper aims to fill this gap in the literature by developing a protocol for evaluating bicycle infrastructure network structure and testing its predictive power on bicycle commuting mode share. Understanding these relationships between bicycle commuting and bicycle network features will enable transportation and planning agencies to target investment in infrastructure components for optimum impact on existing riders and potential future bicyclists.

This paper analyzes bicycle facility networks from 74 mid- to large-sized cities in the US, mapped in Fig. 1 to identify, quantify, and evaluate the backbone network of dedicated bicycling infrastructure. Five network structure factors are constructed from a series of graph theory measures. Regression models are used to test the relationship between these factors and bicycle commuters per 10,000 workers, controlling for city population, land area, median income, household structure, college enrollment, and auto ownership. This paper is organized as follows: First, we discuss literature on the relationships between infrastructure and bicycling. Next, the “Methodology” section explains the data collection process, network structure measures, and factor analysis. The “Results” section describes findings from two linear regression models of bicycle commute share and the sensitivity of bicycle commuting to variables estimated in the model. Finally, we discuss implications for practice and opportunities for further study.

Literature review

Planners and researchers have long sought evidence about the relationship (if any) between certain types of infrastructure and inducing bicycling, with particular focus on the needs of beginner bicyclists. The evidence for bicycle facilities inducing mode shift is weak, but there is strong evidence for dedicated facilities shifting individuals’ route choice calculus, attracting bicyclists from unimproved routes (Pratt et al. 2012). Pucher et al. (1999)’s review of bicycling literature failed to find statistically rigorous studies that establish

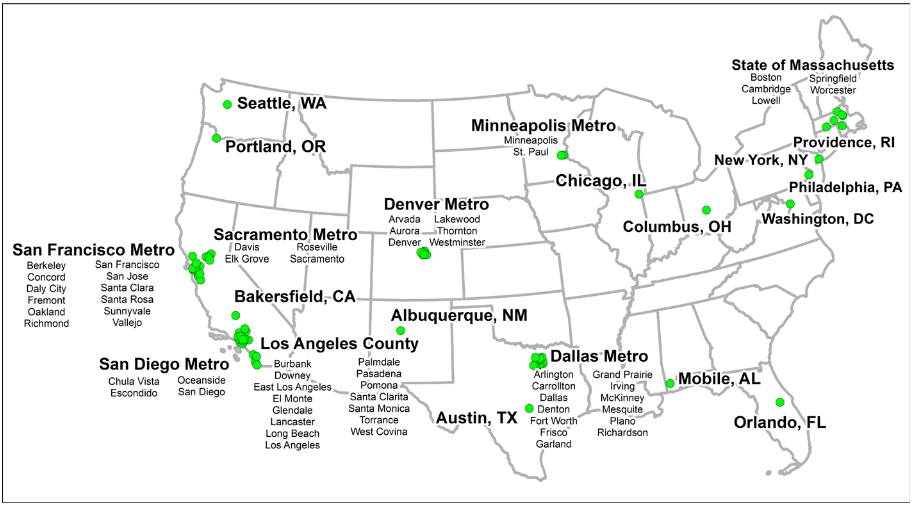


Fig. 1 74 US cities included in sample

causality between infrastructure development and induced cycling. The prevalence of extensive and highly connected cycling infrastructure networks in Europe correlates with high rates of bicycling, but the authors posited that the facility development might follow cyclists rather than incite cycling (Pucher et al. 1999; Buehler and Pucher 2012). A longitudinal study of infrastructure development and rates of commuting by bicycle in the Minneapolis–St. Paul Metropolitan Area reached almost the same conclusion (Krizek et al. 2009): most new facilities were installed in areas that already had much higher than average levels of cycling, with one exception. The University of Minnesota Transitway, a dedicated roadway exclusively for university-operated transit vehicles and bicycles connecting the Minneapolis and St. Paul campuses, made longer cycling commute trips viable, enabling people to switch to cycling for their work or school commutes (Krizek et al. 2009).

Facilities and individual behavior

Bike lanes and trails are still a recommended strategy for increasing cycling because studies have found a significant *correlation* between bicycling behavior and dedicated facilities, even if evidence for *causality* is limited. Wilkinson et al. (1994), in their FHWA manual on roadway design treatments for bicyclists, recommend providing bicycle facilities that both serve existing bicyclists and encourage new riders. They argue that a supply-driven approach of facilities targeted at basic cyclists (Group B) and children (Group C) will encourage mode shift to cycling and use of the facilities. Evidence about which network elements serve bicyclists best is mixed, though it is clear that dedicated infrastructure adds value for all types of bicyclists. Tilahun et al. (2007) found that stated preference survey respondents in the Minneapolis–St. Paul Metro Area valued a bike lane improvement as equivalent to saving about 16 min from a baseline 20-min trip. In Beijing, Zhao (2013) found that a 1 % increase in exclusive bike lanes corresponded to a 0.19 % increase in the probability of choosing to bicycle. Despite expressed willingness to detour to use dedicated facilities in stated preference surveys, bicycle commuters are shown to be highly sensitive to distance (Handy and Xing 2011). Therefore, building a network that provides direct connections with minimal detour is important. A network that requires substantial detour undermines its own utility by making trips longer for a population known to be particularly sensitive to distance. Wilkinson et al. (1994) recommended a network of bike lanes, separated paths, and bike boulevards, or quiet local streets parallel to major corridors with improvements to calm auto traffic and encourage bicycling, to connect Group B/C cyclists to their destinations (Wilkinson et al. 1994; American Association of State Highway and Transportation Officials 2012).

Some studies have found that advanced cyclists (Group A) value reduced travel time and directness, suggesting that their needs are significantly different from B/C cyclists and are sufficiently served with wide shoulders and consistent speed limit enforcement (Wilkinson et al. 1994; Stinson and Bhat 2009; American Association of State Highway and Transportation Officials 2012). However, other scholars have found evidence that even enthusiastic and confident bicyclists receive some benefit from bike lanes and trails. Dill (2009) used GPS devices to study route choice among regular cyclists and discovered that a disproportionate share of bicycle miles traveled occurred on dedicated facilities, suggesting that even experienced cyclists will trade efficiency for safety and comfort. Caulfield et al. (2012)'s stated preference route choice model in Dublin found that off-street trails and green lanes increased the chances of a route being chosen relative to a traditional bike lane, and this finding applied even after stratifying their sample by bicyclist confidence level (ranging from "completely confident" to "not at all confident"). These findings do not necessarily contradict observations that advanced cyclists have different needs. Some confident cyclists may feel that streets without bicycle improvements adequately serve their needs, but dedicated infrastructure provides a more comfortable route alternative when they choose to use it. In any case, it is clear that cyclists of all levels generally prefer and use bike lanes and trails over mixed traffic when it adequately connects them to their destinations.

Of course, the value of dedicated bicycle facilities can be highly context dependent. Several authors have noted that the presence of on-street parking reduces the utility of bike lanes, for example (Tilahun et al. 2007; Sanders 2014). Klobucar and Fricker (2007) used a bicycle compatibility index (BCI) based on traffic volumes and travel speeds to model the effects of distributing bike lanes on streets with high or low compatibility scores. Additionally, low traffic and low speed streets, particularly without street parking, can be just as useful or even more so than some types of dedicated infrastructure. The recently revised "Guide for the Development of Bicycle Facilities" published by American Association of State Highway and Transportation Officials (2012) describes bicycle network planning as a process involving comprehensive qualitative and quantitative assessment of existing conditions, needs, and feasibility. No single facility type serves all roadway contexts and user needs. Whatever the appropriate facility type, the Guide recommends that routes be direct with as few detours as possible, and bikeway density should be planned for maximum use and comfort (e.g., within one quarter mile of every resident). Complete streets policies are recommended to encourage consideration of bicyclists' needs during any road resurfacing or maintenance project (American Association of State Highway and Transportation Officials 2012).

Cross-regional studies

Cross-regional studies have found positive relationships between the size of a city's bicycle facility network and its bicycle commute share. Nelson and Allen (1997) found that each additional mile of bikeway per 100,000 residents was associated with a 0.069–0.075 % increase in commuters using bicycles across 18 US cities. Dill and Carr

(2003), using an expanded dataset of 50 cities, found each additional mile of facility per square mile of city area to be associated with a one percentage point increase in bike commuting.

Rietveld and Daniel (2004) performed a similar type of study modeling bicycle share of use in 103 cities in the Netherlands, though instead of measuring the quantity of bicycle facilities, they used several measures of infrastructure quality developed by the Dutch Cyclists' Union (Fietsersbond). They found the number of stops and hindrances per kilometer on any given trip were both negatively associated with bicycle share. If the travel time via bicycle for a trip was 10 % faster than by car, then bicycle use for that trip increased 3.4 %. They concluded that bicyclists are sensitive to the speed and directness of their routes, so city policy that makes bicycle trips easier and more efficient will increase rates of cycling.

Parkin et al. (2007)'s study of bicycle commuting in English and Welsh electoral wards found a positive, significant association between proportion of off-street bicycle routes and ridership, but they noted that the elasticity was small: only about 0.049, suggesting that a disproportionately large quantity of off-street routes would need to be built to see only a modest increase in bicycling. They also assert a bicycling saturation point of 43 %, though this was considerably higher than the maximum observed in their study. Buehler and Pucher (2011)'s findings from a study of 90 US cities corroborate this evidence: they found a 10 % increase in supply of bike lanes is associated with only a 3.1 % increase in the number of bike commuters per 10,000 residents. For off-street paths, the increase was only 2.5 %.

The missing link in many studies is the quality of the network formed by the infrastructure: can a cyclist complete her desired trip using the bicycle network without significant detours or discontinuities that would require riding in unsafe or uncomfortable conditions? Most of the cross-regional studies conducted to date have focused on quantity or density of infrastructure, with no evaluation of how well connected the infrastructure is. Rietveld and Daniel (2004)'s study addressed network quality indirectly by measuring barriers along trips, but their variables did not directly measure connectivity.

Mekuria et al. (2012)'s analysis of "low stress" networks clearly demonstrates the importance of connectivity by evaluating the quality of routes holistically by their weakest link rather than an average or index. From this perspective, a route that is almost entirely comprised of off-street trails but requires crossing major streets at a few hostile intersections would be inaccessible for someone whose tolerance threshold is lower than those intersections would support. However, the focus is on microscopic street characteristics rather than a macroscopic summary of bicycle facilities.

Graph theory and bicycling

While it is not expected that a bicyclist complete 100 % of their trip exclusively on dedicated infrastructure, lanes and trails form a backbone network for bicycling in a city. Thus the utility of dedicated infrastructure is closely related to what level of connectivity it provides. Planning for isolated infrastructure segments without considering how these pieces of infrastructure connect to the broader street network undermines the potential utility of this infrastructure. Discontinuities in the bicycle network may have three potential consequences:

(1) Forcing the cyclist into mixed traffic
(2) Requiring lengthy detours to avoid mixed traffic
(3) Discouraging cycling altogether

Graph theory offers systematic methods for measuring network quality for comparison across cities and to see its effect on travel behavior. Garrison and Marble (1962) first introduced graph theory principles to transportation geography. Kansky (1963) presents the alpha index (α), beta index (β), and gamma index (γ) in his dissertation as ratios that describe the relationship between distinguishable elements of a graph.

Graph theory measures applied to transportation research are now fairly common (Xie and Levinson 2007; Derrible and Kennedy 2009; Rodrigue et al. 2009). Simple measures, such as street density or cul-de-sac density, are frequently used to characterize the built environment and model travel behavior. For example, Parthasarathi et al. (2013) found that increased street density was associated with longer perceived travel times for drivers. While one might hypothesize the relationship to be a bit more complex for nonmotorized road users, this application illuminates the possible mechanisms by which network connectivity may influence travel behavior.

Several studies have used graph theory measures to explain individuals' nonmotorized travel behavior. Berrigan et al. (2010) measured the link-node ratio and several graph theory indices of the local street grid within short buffers around survey respondents' home addresses. These measures factored into two main variables of network quality that predicted propensity and duration of active travel. Dill and Voros (2007) found significant differences between the connected node ratios people who biked during the previous summer and people who did not. Network quality and connectivity have been evaluated at a microscopic level by studying individual discontinuities in on-street bicycle facilities and long-term network development at gaps and critical points (Krizek and Roland 2005; Birk and Geller 2006; Barnes and Krizek 2005).

These studies significantly advanced our understanding of bicycle infrastructure quality, while previously described research (Nelson and Allen 1997; Dill and Carr 2003; Rietveld and Daniel 2004; Parkin et al. 2007; Buehler and Pucher 2011) set a precedent for modeling a population's use of bicycling as a function of infrastructure availability. This study fills the missing link between these two bodies of research by (1) adapting existing graph theory and transport geography measures to describe a city's bicycle infrastructure network, (2) measuring the quality of bicycle networks using a newly assembled collection of spatial data from 74 US cities, and (3) modeling bicycle commuting as a function of these measures, controlling for several common demographic correlates of cycling.

Methodology

Data collection and assembly

In this study, we collected spatial data and American Community Survey (ACS) household, demographic, and journey to work data for 74 mid- and large-sized cities in the US. The sampling frame consisted of mid-size and large US cities for which a spatial dataset of bicycle infrastructure was publicly available, and the data was organized in a usable format (ESRI Shapefile or Keyhole Markup Language).

Cities were only included if it was possible to distinguish separate bicycle facilities (lanes, side paths, paved trails) from mixed-traffic facilities (shared lane markings, signed bike routes, bike boulevards) as only the separate facilities were included in the analysis. As previously described, the appropriate facility type for any given corridor depends heavily on local context. Selecting only specific types of infrastructure for network analysis means that the quality of bike networks in cities that make contextually appropriate use of other types of infrastructure may be underestimated.

Several large metro areas provided data for all their member cities in one file. These shapefiles were disaggregated to the city level and then considered for analysis. Adjacent cities may share bicycle travel (e.g., a resident of City A commutes to City B for work), so these cases would violate the assumptions of independence in regression modeling. To reduce the impact of this phenomenon on the study, only cities with population greater than 100,000 were retained from disaggregated metro areas. This has two effects: sampling only a portion of cities reduces the chances that adjacent cities will be included, and using only larger cities increases the chance that any given commuter starts and ends their journey in the same city.

Many of the bicycle infrastructure network files had geometry inconsistencies arising from how they were created, such as dangling ends where two edges should meet at a single vertex. To reduce the possible impacts of errors and subtle digitization differences across cities, we cleaned the shapefiles using a model routine in ArcGIS to trim dangling ends less than 5 m, extend ends to fill gaps less than 5 m, remove excess vertices, and dissolve single network features that were represented with multiple segments.

The final sample included 74 US cities, covering 22 of the top 50 most populous. These cities are shown in Fig. 1. Where multiple cities from the same metro area are included, the city names are organized by the agency providing data.

Commuting behavior and demographics

We collected commuting and demographic data for all cities in the sample from the 2005 to 2009 ACS 5-year estimates. The US Census Bureau administers the ACS on a rolling basis and provides aggregated estimates over varying timeframes. 5-year estimates are more stable than the other alternatives because they have a larger sample and longer timeframe. The commuting data was converted into a rate of bike commuters per 10,000 commuters. Demographic control variables included percentage of households with children under 18, share of residents who are enrolled in college, auto ownership rates, median household income, total population, and number of workers. Auto ownership rates were constructed from an aggregate number of vehicles owned in the city and the number of households. College enrollment was included to control for possible outlier cases such as Davis, California.

Table 1 summarizes the commuting, demographic, and spatial variables in this study and their expected relationship with bicycle commuting. The very low rates of bicycle commuting are apparent here: the highest commute share in the study is 15.5 % (Davis), but all other cities have commute shares below 7 % (not shown). The sample included a diverse range of incomes, household structures, and vehicle ownership rates.

The ACS data has several notable caveats. The survey asks respondents by which mode they traveled to work the most in the past week, so the dependent variable only measures bicycle commuting. Bike commuting may not be representative of bicycling for other trip purposes. Bike commuters are more likely to be frequent, confident cyclists than someone who bikes to meet a friend for coffee once a week, so an observed relationship between infrastructure and bike commuting may not be useful for identifying what other types of cyclists need from infrastructure.

Bicyclists tend to be multimodal; they may bike some days, but drive or use transit on other days when the weather is bad or when they have to travel for meetings or appointments during the workday (Heinen et al. 2010). Part-time bike commuters may be

Table 1 Descriptive statistics for demographic/city variables

Variables	Average	Standard deviation	Min	Max
Bicyclists per 10,000 workers	116.0	210.6	5.1	1,547.5
Population	488,185	1,089,012	61,866	8,302,659
Land area (km^2)	218	243	17	1,214
Median income ($USD)	57,645	13,937	34,113	97,160
Percent households with kids (%)	35.9	9.4	15.9	56.3
Percent college students (%)	9.1	5	5	35.9
Vehicles per household	1.7	0.3	0.6	2.2

Data from ACS 2009 5-year estimates

underrepresented because the ACS only allows for one *single* mode in response to the question about commuting. The ACS year-round rolling sampling strategy introduces an incredible range of weather variability within and between cities. Finally, due to the sample size of the ACS and the relatively low numbers of bike commuters in each city, the standard error on these measures are very high; nonetheless, the ACS remains the only nationwide survey of travel behavior that can be used at this geographic resolution.

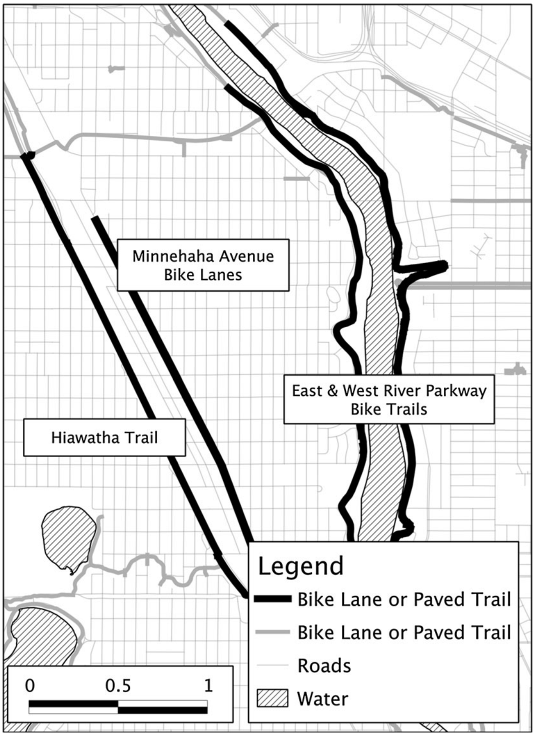
Network measures

Graph edges represent segments of bicycle infrastructure (bike lanes, trails, sidepaths). Vertices represent intersections and endpoints within the bicycle infrastructure network (e.g., where two trails cross). The dataset does not contain any connections between the bicycle network and the remaining road network. For each city, we measured the number of edges (e), number of vertices (v), and the total length of all the edges, or graph length (L).

We also measured an airline distance between the endpoints of each edge to calculate how direct each segment is. An overall directness measure was estimated using the cumulative difference between each edge's travel length and airline distance between its endpoints. The purpose of this measure is to gauge how much of a city's bike network is comprised of recreationally oriented paths that meander or circle back on themselves (e.g., paths winding through parks or circling around a lake), versus paths that provide an efficient utilitarian connection for commuting. Figure 2 shows four segments of infrastructure in Minneapolis's bicycle network to demonstrate direct versus indirect/recreational paths. The Hiawatha Trail and Minnehaha Avenue Bike Lanes are almost perfectly straight. The lengths of their edges are equal to the airline distances between their endpoints. The East and West River Parkway Trails have a lot of curves to take advantage of the natural scenery along the river, so the edge lengths are much longer than the airline distances between endpoints.

Many cities' bicycle networks are fragmented into sub-networks or subgraphs. Each distinct "island" of links within a city should be counted as a subgraph. However, some cities leave small gaps in the data where a bike lane crosses an intersection, even though the lanes on either side function as a single facility. These gaps were too large to resolve using the model routine described previously to fix very minor dangling ends or gaps. We generated a 50-m buffer around all links and treated any link that intersected another link's buffer as part of the same subgraph. Figure 3 shows a stylized example of this process. Although none of the four depicted segments intersect, they are contained within the same 50-m buffer and therefore defined as part of the same subgraph in this study. We then measured the number of subgraphs (p), length of the largest subgraph, and average

Fig. 2 Visual comparison of directness on four infrastructure segments in Minneapolis, MN



subgraph length. The relative size of the largest and average subgraphs were calculated by dividing by L. If the largest subgraph contains a high percentage of the bike network, this suggests one main network with smaller fragments. If the percent of the bike network contained by the average subgraph is relatively high, this suggests a more even distribution between two or more smaller but substantial subgraphs. Whether these scenarios are useful for bicyclists or a deterrent will depend on what types of road facilities connect the bike network fragments.

The α, γ, and β indices were calculated for each city based on Rodrigue et al. (2009)'s description of the formulas. The beta index (β) is a ratio of number of edges (e) to number of vertices (v), shown in Eq. 1.

$$\beta = \frac{e}{v}. \tag{1}$$

A collection of disconnected, non-intersecting edges has $\beta = 0.5$ (see Eq. 2).

$$\begin{aligned} v &= 2e \Rightarrow \\ \frac{e}{v} &= \frac{e}{2e} = \frac{1}{2}. \end{aligned} \tag{2}$$

A bicycle network with a low β suggests an increased chance that any given route requires leaving dedicated infrastructure to ride with mixed traffic. Adding links to a network, and thereby increasing the value of β, increases the complexity and the probability that any two given vertices have links between them. The gamma index (γ) is closely related: it

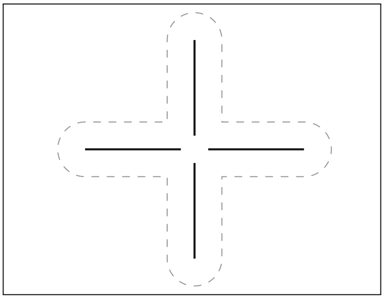


Fig. 3 Buffering process to identify subgraphs with small gaps

represents the number of observed edges (e) to the theoretical maximum number of edges, estimated using Eq. 3 (Rodrigue et al. 2009).

$$\gamma = \frac{e}{3(v-2)}. \tag{3}$$

γ values range from 0 to 1, with 0 indicating a cluster of vertices with no edges, and 1 indicating a fully connected graph. Higher values indicate greater internal connectivity and increased redundancy, though the excess redundancy associated with values of γ approaching 1 makes these types of networks impractical (Rodrigue et al. 2009). In general, a more highly connected bicycle network should provide more direct paths than a less connected network.

Cycles are connected chains of edges with the same starting and ending point, and the number of cycles (u) is estimated using Eq. 4, where p is the number of subgraphs.

$$u = e - v + p. \tag{4}$$

The alpha index (α) is the ratio of u to the theoretical maximum number of cycles, shown in Eq. 5.

$$\alpha = \frac{u}{2v-5}. \tag{5}$$

Higher values of u correspond to a higher level of complexity and development within the network. Simple networks or tree-like networks have no cycles. Like γ, α values range from 0 to 1 and values approaching 1 are highly unlikely due to excessive redundancies. Unlike γ, this measure is independent from the number of nodes and therefore should be less size-dependent (Rodrigue et al. 2009).

Table 2 provides descriptive statistics for all the network measures included in this study. It should be noted that the 50-m buffer used to define subgraphs affected the estimates for number of cycles (u) and α index calculations because the number of subgraphs measured this way is artificially low relative to the number of vertices. The minimum values of the α index and number of cycles (u) are both negative, which is theoretically impossible according to the formulas, but feasible given the subgraph estimation procedure used in this study.

The numbers of edges and vertices show the wide range of cities included in this study. Four cities in this sample have only a single link in their bicycle network ($e = 1$, $v = 2$), while the largest city (New York) has over 1,000 edges. The bicycle networks represented in this sample are fragmented into an average of 32 subgraphs per network, and 26 % of vertices are intersections (vs. end points).

Table 2 Descriptive statistics for bicycle facility network variables

Variables	Average	SD	Min	Max
Length (km)	311.16	450.24	1.54	2,204.22
Number of edges	190.58	246.98	1.00	1,034.00
Number of vertices	202.34	239.18	2.00	1,028.00
Number of subgraphs	31.54	34.75	1.00	168.00
Network density (km/km^2)*	1.74	2.67	0.03	18.67
Edge density (#/km^2)*	1.49	3.19	0.00	24.50
Subgraph density (#/km^2)*	0.22	0.26	0.00	1.45
Vertex density (#/km^2)*	1.55	2.80	0.01	20.24
Proportion intersections	0.26	0.17	0.00	0.64
Direct length (km)	275.70	408.57	1.38	1,987.12
Average length ratio	0.91	0.09	0.42	1.00
Minimum length ratio	0.28	0.35	0.00	1.00
Overall length ratio	0.86	0.11	0.42	1.00
Average subgraph length (km)	11.84	21.29	0.77	155.12
Average subgraph percent	0.15	0.24	0.01	1.00
Max subgraph length (km)	144.60	249.30	0.99	1,621.75
Max subgraph percent	0.49	0.26	0.13	1.00
Number of cycles	19.78	46.78	−63.00	188.00
α Index	0.03	0.05	−0.07	0.17
γ Index	0.28	0.06	0.17	0.43
β Index	0.81	0.20	0.50	1.28

* Density measures use city's total land area as the denominator

The average density of bike lanes and trails among these 74 cities is 1.74 km/km^2 of land in the city, with a range from 0.03 to 18.67 km/km^2 This sample has considerably denser networks than Dill and Carr (2003)'s cross-regional study of bicycle commuting (0.38 km/km^2). Among the 21 cities included in both cities, these data suggest that bicycle facility density has increased by as much as 210 % (from 0.51 km/km^2 to 1.09) in the decade between data collection periods, depending on variation in data sources, accuracy, and measurement technique.

Factor analysis

Many of the network variables were derived from the same set of spatial measurements, so they have a high degree of correlation with each other. To avoid multicollinearity issues in the regression model, principal component analysis with varimax rotation was used to identify underlying qualities from the list of 21 measures. The resulting five factors with eigenvalues greater than 1 characterize each network by its (1) size, (2) connectivity, (3) density, (4) fragmentation, and (5) directness.

Table 3 shows how variables load onto these factors. The size factor is mostly composed of variables measuring the number of network components (number of edges, vertices, subgraphs), as well as measures of their overall length. Connectivity contains the α, γ, and β indices and percent of vertices that are intersections. The density factor characterizes both length of facility per unit area as well as number of components (edge, vertex, and subgraph density). Fragmentation represents the percent of the network contained in the largest and average subgraphs and the length of the average subgraph. Finally, the directness factor captures all ratio measures between the length of the path and the length of a straight line connecting the path's endpoints.

Factors are normalized to have a mean of 0 and standard deviation of 1. Each city's factor scores are also relative to the sample used to construct the factors. For example, if a city has a connectivity factor value of 0, that city can be considered average *within* the sample. A city with a connectivity factor of -1 would be one standard deviation lower than average, relative to all the cities included. Figure 4 demonstrates these differences for the network size factor by mapping cities with factor scores closest to -1, 0, and $+1$ on the same scale.

Results

Table 4 shows the results of two linear regressions. In Model 1, the five network factors are used to model bicycle commuters per 10,000 commuters. The adjusted R^2 for this model is 0.509, suggesting that network structure measures explain about 50 % of the variation in rates of bicycle commuting between cities in the sample. The connectivity and density factors are positive and significant. A one unit increase in the density factor (or one standard deviation, for normalized factor scores) corresponds to about 150 additional bicycle commuters per 10,000 commuters in the city, all else equal. Connectivity has a weaker relationship, with a one unit change associated with an increase of 37 bicyclists. The standard errors for all coefficients in Model 1 are equal because all explanatory variables in this model are normalized to have a mean of 0 and standard deviation of 1.

Model 2 controls for city size and demographic and economic characteristics. The adjusted R^2 for this model is 0.804, which represents a considerable improvement over the network only model. In this model, density and connectivity are still positive and significant, now joined by fragmentation and directness. Among the control variables, percent college students is the only significant one, with each additional percentage point increase in population enrolled in college corresponding to about 22 additional bicyclists per 10,000 commuters. The Average VIF scores for both models are quite low, suggesting that factor analysis addresses any possible multicollinearity issues between similarly constructed network measures. We repeated the analysis using robust standard errors to account for heteroscedasticity arising from the diverse sample of city sizes and populations, but this did not change the significance of any variables at the thresholds indicated.

Elasticity

Table 5 shows sensitivity of bicycle commuting to changes in network factor scores using the regression results in Model 2. Since factor scores are relative to the cities within the sample and normalized to have a mean of 0 and a standard deviation of 1, it is impossible to represent a 1 % increase from $\mu = 0$ meaningfully. Elasticities are calculated based on adding 0.01, 0.25, 0.5, and 1.0 to the mean factor score, representing 0.01, 0.25, 0.5, and 1 standard deviations. Bicycle commuting is most sensitive to changes in network density. A change in the density factor score from 0 to 1 (or one standard deviation) corresponds to a 77 % increase in rates of bicycle commuting. Increasing a city's connectivity score or fragmentation score from 0 to 1 is associated with about a 20 % increase in biking, and the same change in directness corresponds to a 27 % change in the dependent variable.

Table 3 Factor loadings

Variables	Factor 1 Size	Factor 2 Connectivity	Factor 3 Density	Factor 4 Fragmentation	Factor 5 Directness
Length (km)	0.965				
Direct length (km)	0.961				
Number of vertices	0.851				
Edges	0.814				
Number of subgraphs	0.790				
Max subgraph length (km)	0.793				
α Index		0.895			
γ Index		0.915			
Percent intersections		0.908			
β Index		0.910			
Number of cycles	0.536	0.627			
Average length ratio					0.928
Overall length ratio					0.937
Average subgraph percent				0.720	
Max subgraph percent				0.881	
Minimum length ratio					0.511
Average subgraph length (km)				0.608	
Network density (km/km^2)			0.808		
Vertices per km^2			0.959		
Edges per km^2			0.921		
Subgraphs per km^2			0.734		

Loadings smaller than 0.5 are suppressed

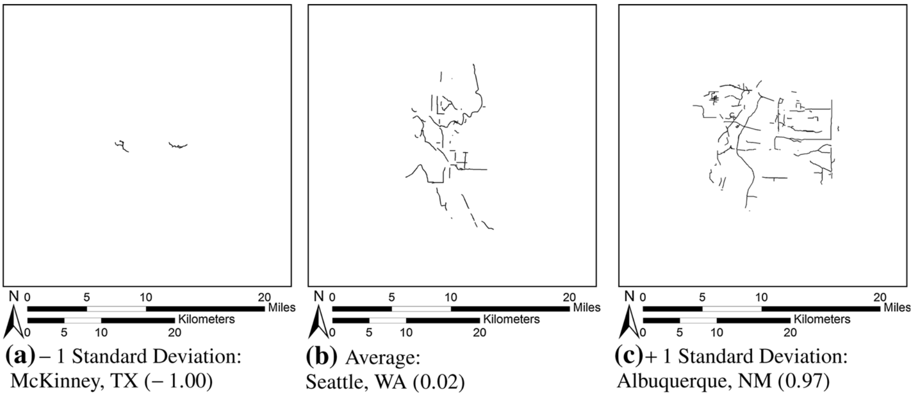


Fig. 4 Comparison of network size factor scores

Table 6 shows similar sensitivity calculations for city size and demographic variables. College enrollment is the only significant size or demographic variable in the model, and bicycle commuting is more sensitive to changes in college enrollment than all other

Table 4 Regression results

Variables	Model 1		Model 2	
	Coefficient	SE	Coefficient	SE
Factor 1: size	−2.184	17.272	−19.352	16.109
Factor 2: connectivity	37.351**	17.272	22.426*	12.054
Factor 3: density	149.956***	17.272	89.846***	15.245
Factor 4: fragmentation	6.715	17.272	23.881*	12.720
Factor 5: directness	10.955	17.272	31.701**	11.990
Area (km^2)			0.113	0.073
Population (1,000)			0.006	0.018
Median income ($1,000)			−1.407	1.072
Percent HH with kids			−2.975	2.111
Percent college students			22.398***	3.232
Vehicles per HH			1.920	69.564
Constant	116.045***	17.155	68.507	98.296
R^2	0.543		0.833	
Adjusted R^2	0.509		0.804	
Average VIF	1.000		2.300	

*, ** or *** significance at $p < 0.1$, $p < 0.05$, or $p < 0.01$

Table 5 Elasticity of bicycle commuting to network factors

Change variables:	+1 %	+0.01	+0.25	+0.5	+1 Standard deviation
Factor 1: size	0.00	−0.17	−4.17	−8.34	−16.68
Factor 2: connectivity*	0.00	0.19	4.83	9.66	19.33
Factor 3: density***	0.00	0.77	19.36	38.71	77.42
Factor 4: fragmentation**	0.00	0.21	5.14	10.29	20.58
Factor 5: directness**	0.00	0.21	5.14	10.29	27.32

Elasticities computed with all other variables evaluated at means. Factors are normalized to $\mu = 0$ and $\sigma = 1$

*, ** , *** Variable was significant in Model 2 at $p < 0.1$, $p < 0.05$, or $p < 0.01$

network, size, and demographic variables alike. A 1 % increase in college enrollment corresponds to a 1.76 % increase in bike commuting. Increasing college enrollment by one standard deviation (five percentage points, in this case) is associated with a 97 % increase.

Discussion

The regression model results show that a city's bicycle commuting rate is associated with several network structure measures, even after controlling for the city's size, population, median income, household structure, college enrollment, and vehicle ownership. The density factor had the strongest coefficient and elasticity: a 1-standard deviation increase in network density corresponded to a larger increase in bicycle commuting than 1-standard deviation increases in connectivity, fragmentation, and directness combined. Although the

Table 6 Elasticity of bicycle commuting to city size and demographics

Change in variable:	+1 %	+25 %	+50 %	+100 %	+1 Standard deviation
Population (1,000)	0.03	0.66	1.32	2.63	5.87
Area (km^2)	0.21	5.29	10.58	21.17	23.67
Median income ($1,000)	−0.70	−17.48	−34.96	−69.92	−16.90
Percent HH with kids	−0.92	−23.00	−46.01	−92.02	−24.11
Percent college students***	1.76	44.09	88.17	176.34	96.73
Vehicles per HH	0.03	0.69	1.38	2.76	0.53

Elasticities computed with all other variables evaluated at means

*, **, *** Variable was significant in the regression model at the $p<0.1$, $p<0.05$, or $p<0.01$ level

network factor scores are relative to this particular sample, the results from this study still demonstrate the relative importance of network characteristics for biking.

The measures that loaded onto the density factor are density of facilities (km/km^2), density of edges and vertices, and density of subgraphs. These findings suggest that cities hoping to maximize the impacts of their bicycle infrastructure investments should first consider densifying their bicycle network before expanding its breadth.

Fragmentation has the potential to affect bicyclists in a number of ways. A series of careless, one-off projects could leave bicyclists without adequate routes across the city, but strategically installing bike lanes or trails to augment a comfortable residential street network could expand cyclists' mobility opportunities. The fragmentation factor is composed of the share of bike network contained in the largest subgraph and average subgraph, and the length of the average subgraph (Table 3). The coefficient on the fragmentation factor is positive and significant in Model 2 (Table 4). This suggests that having either an even distribution of reasonably sized fragments (large share of bike network in average subgraph) or one dominant section (large share of bike network in largest subgraph) facilitates bicycling, but excessive fragmentation with small fragments should be avoided.

The γ and β indices were the strongest loadings on the connectivity factor. Additionally, the strongest loading in the fragmentation factor was the percent of the network contained in the largest subgraph. If an agency has the funds to build one new bike lane, they could capitalize on the significance of the connectivity, fragmentation, and density factors simultaneously by making the new link intersect with parts of the existing network within the largest subgraph or linking the largest subgraph to adjacent ones. The mere fact of adding facility densifies the network, while targeting its placement to connect to existing links and expand the largest subgraph improves the connectivity and fragmentation scores, respectively.

The results also showed that college enrollment was a strong predictor for bicycle commuting, with an even stronger elasticity than the network measures. While this is not surprising, it does highlight how bicycle commuting may be more useful or practical for certain segments of the population than others. This has several possible policy implications. First, it is clear that college towns will always have an edge for rates of bicycle commuting. Building a new college or university may attract a new population of bicycling residents. More practically, agencies could consider whether infrastructure investments around a college would have a smaller or larger marginal effect than investments in other parts of the city. If a college has weak rates of bicycle commuting, targeting network development between campus and neighborhoods where students tend to live may have a bigger impact than an equivalent investment in another part of the network.

Limitations and areas for future study

As discussed throughout this study, data issues are always a concern when studying bicycling. Bicycling has a small commute mode share and the characteristics of the mode make its users more vulnerable to distance, climate, and weather. Typical sampling strategies for the ACS do not capture the full range of bicycling that occurs in any given city. Additionally, bicycle infrastructure data is sparse and lacks the industry-wide standard for creation, maintenance, storage, and distribution that we have for most road networks. Building a sample from cities that publish their bicycle network data biases the sample toward cities that value bicycling.

Studying aggregate travel behavior does not necessarily imply individual effects or outcomes. The graph theory measures employed here characterize the bicycle network's structure, but they do not account for the built environment context surrounding the bicycle network. At the individual level, the overall bike network structure might not matter much, as long as the commute trip distance is appropriate for bicycling and the individual has a route they feel comfortable with connecting their home and work. One way to address this in future research is through measuring accessibility to jobs or other destinations that the network provides. Between-city variation in policy, weather, topography, and other contexts also may influence results. Additionally, aggregate studies are vulnerable to deceptively high adjusted R^2 values because aggregation smooths out individual variability. Therefore the strength of our model may be overstated.

While insignificant in this model, some demographic factors such as gender and households with children are negatively associated with bicycling in the literature. Weak or negative relationships between demographic groups and bicycling should not be used as an excuse to completely divest from certain neighborhoods. Further study is needed to fully explain infrastructure and other barriers to cycling, and to identify what *other* kinds of bicycle trips beyond commuting could be made accessible through strategically building out the bicycle network.

This study treated on-street and off-street dedicated infrastructure (bike lanes and paved trails) interchangeably for purposes of measuring network structure because they both provide a designated space for bikes, versus newer or experimental treatments like shared lane markings or bike boulevards. However, existing research shows that bicyclists value these facility types differently. Tilahun et al. (2007)'s stated preference survey findings suggest that cyclists value the presence of a bike as equivalent to saving about 16 min of travel time, while off-road improvements add less value than a quiet street without on-street parking (5 and 9 min, respectively). Sanders (2014)'s results suggest that bicyclist comfort level in a bike lane varies based buffers, barriers, and on-street parking. Klobucar and Fricker (2007) demonstrated the use of BCI to model the effects of building new dedicated bike infrastructure on compatible and incompatible streets, which points to the information we miss by not evaluating the streets on which the bike lanes are built. Mixed traffic facilities such as shared lane markings, signed bike routes, and bike boulevards were completely excluded because they vary so widely in quality and definition. This study sheds light on the desperate need for standardized data collection and management practices for bicycle infrastructure networks and nonmotorized travel behavior. Given more standardized data, future study should consider hierarchies of infrastructure types within bicycle networks and complementary street networks and what effects these have on bicycling.

Acknowledgments The authors are grateful for the feedback from three anonymous reviewers whose feedback greatly improved the quality of the paper.

References

American Association of State Highway and Transportatsion Officials: Guide for the Development of Bicycle Facilities, 4th edn. American Association of State Highway and Transportation Officials, Washington, DC. https://bookstore.transportation.org/collection_detail.aspx?ID=116 (2012). Accessed 15 Feb 2014

Barnes, G., Krizek, K.: Estimating bicycling demand. Transp. Res. Rec. **1939**, 45–51 (2005). doi:10.3141/1939-06

Berrigan, D., Pickle, L.W., Dill, J.: Associations between street connectivity and active transportation. Int. J. Health Geogr. **9**(20) (2010). doi:10.1186/1476-072X-9-20

Birk, M., Geller, R.: Bridging the gaps: how quality and quantity of a connected bikeway network correlates with increasing bicycle use. In: TRB 85th Annual Meeting Compendium of Papers number 06-0667. http://trid.trb.org/view.aspx?id=776458 (2006). Accessed 14 Aug 2013

Buehler, R., Pucher, J.: Cycling to work in 90 large American cities: new evidence on the role of bike paths and lanes. Transportation **39**(2), 409–432 (2011). doi:10.1007/s11116-011-9355-8

Buehler, R., Pucher, J.: Walking and cycling in Western Europe and the United States: trends, policies, and lessons. TR News **280**, 34–42. http://trid.trb.org/view.aspx?id=1143635 (2012). Accessed 12 Nov 2013

Byers, R.H.: Hennepin County bicycle system gap study. Technical Report, Hennepin County. http://www.hennepin.us/files/HennepinUS/Transportation/BikewaysandTrails/BicyclePlanning/BikeGapStudyFull.pdf (March 2002). Accessed 7 Aug 2013

Caulfield, B., Brick, E., McCarthy, O.T.: Determining bicycle infrastructure preferences: a case study of Dublin. Transp. Res. D **17**(5), 413–417 (2012). doi:10.1016/j.trd.2012.04.001

Derrible, S., Kennedy, C.: Network analysis of world subway systems using updated graph theory. Transp. Res. Rec. **2112**, 17–25 (2009). doi:10.3141/2112-03

Dill, J.: Bicycling for transportation and health: the role of infrastructure. J. Public Health Policy **30**(Suppl 1), S95–S110 (2009). doi:10.1057/jphp.2008.56

Dill, J., Carr, T.: Bicycle commuting and facilities in major US cities: if you build them, commuters will use them. Transp. Res. Rec. **1828**, 116–123. doi:10.3141/1828-14 (2003)

Dill, J., Voros, K.: Factors affecting bicycling demand: initial survey findings from the Portland, Oregon Region. Transp. Res. Rec. **2031**, 9–17 (2007). doi:10.3141/2031-02

Federal Highway Administration: Report to the U.S. Congress on the Outcomes of the Nonmotorized Transportation Pilot Program SAFETEA-LU Section 1807. Technical Report, April. Submitted by the Federal Highway Administration With the Assistance of the U.S. Department of Transportation's Volpe National Transportation Systems Center, Washington, DC (2012)

Garrison, W.L., Marble, D.F.: The Structure of Transportation Networks. Technical Report. Northwestern University, Evanston (1962)

Handy, S.L., Xing, Y.: Factors correlated with bicycle commuting: a study in six small US cities. Int. J. Sustain. Transp. **5**(2), 91–110 (2011). doi:10.1080/15568310903514789

Heinen, E., van Wee, B., Maat, K.: Commuting by bicycle: an overview of the literature. Transp. Rev. **30**(1), 59–96 (2010). doi:10.1080/01441640903187001

Kansky, K.J.: Structure of transportation networks. Doctoral Dissertation, University of Chicago (1963)

Klobucar, M.S., Fricker, J.D.: Network evaluation tool to improve real and perceived bicycle safety. Transp. Res. Rec. **2031**, 25–33 (2007). doi:10.3141/2031-04

Krizek, K.J., Roland, R.W.: What is at the end of the road? Understanding discontinuities of on-street bicycle lanes in urban settings. Transp. Res. D **10**(1), 55–68 (2005). doi:10.1016/j.trd.2004.09.005

Krizek, K., Barnes, G., Thompson, K.: Analyzing the effect of bicycle facilities on commute mode share over time. J. Urban Plan. Dev. **135**(2), 66–73 (2009). doi:10.1061/(ASCE)0733-9488(2009)135:2(66)

Mekuria, M.C., Furth, P.G., Nixon, H.: Low-Stress Bicycling and Network Connectivity. Technical Report. Mineta Transportation Institute, San Jose (2012)

Nelson, A., Allen, D.: If you build them, commuters will use them: association between bicycle facilities and bicycle commuting. Transp. Res. Rec. **1528**, 79–83 (1997). doi:10.3141/1578-10

Parkin, J., Wardman, M., Page, M.: Estimation of the determinants of bicycle mode share for the journey to work using census data. Transportation **35**(1), 93–109 (2007). doi:10.1007/s11116-007-9137-5

Parthasarathi, P., Levinson, D., Hochmair, H.: Network structure and travel time perception. PloS One **8**(10), e77718 (2013). doi:10.1371/journal.pone.0077718

Pratt, R.H., Levinson, H.S., Turner, S.M., Jeng, C.Y., Nabors, D.: Chap 16: pedestrian and bicycle facilities. In: TCRP Report 95: Traveler Response to Transportation System Changes. Transportation Research Board, Washington, DC. http://trid.trb.org/view.aspx?id=171542 (2012). Accessed 12 Nov 2013

Pucher, J., Komanoff, C., Schimek, P.: Bicycling renaissance in North America? Recent trends and alternative policies to promote bicycling. Transp. Res. A **33**(7/8), 625–654 (1999). doi:10.1016/S0965-8564(99)00010-5

Rietveld, P., Daniel, V.: Determinants of bicycle use: do municipal policies matter? Transp. Res. A **38**(7), 531–550 (2004). doi:10.1016/j.tra.2004.05.003

Rodrigue, J.P., Comtois, C., Slack, B.: The Geography of Transport Systems, 2nd edn. Routledge, New York (2009)

Sanders, R.L.: Roadway design preferences among drivers and bicyclists in the bay area. In: TRB 93rd Annual Meeting Compendium of Papers number 14-5454, 2014

Stinson, M.A., Bhat, C.R.: A comparison of the route preferences of experienced and inexperienced bicycle commuters. In: TRB 84th Annual Meeting Compendium of Papers number 05-1434, 2005

Tilahun, N.Y., Levinson, D.M., Krizek, K.J.: Trails, lanes, or traffic: valuing bicycle facilities with an adaptive stated preference survey. Transp. Res. A **41**(4), 287–301 (2007). doi:10.1016/j.tra.2006.09.007

Wilkinson, W., Clarke, A., Epperson, B., Knoblauch, R.: Selecting Roadway Design Treatments to Accommodate Bicycles. Technical Report, January. USDOT Federal Highway Administration, Great Falls (1994)

Xie, F., Levinson, D.: Measuring the structure of road networks. Geogr. Anal. **39**(3), 336–356 (2007). doi:10.1111/j.1538-4632.2007.00707.x

Zhao, P.: The impact of the built environment on bicycle commuting: evidence from Beijing. Urban Stud. (2013). doi:10.1177/0042098013494423

Jessica E. Schoner is a Doctoral Student in the University of Minnesota Department of Civil, Environmental, and Geo-Engineering. Her current research focuses on mutual causality between bicycling and infrastructure, and travel behavior of bicyclists and bike share users.

David M. Levinson is a Professor in the University of Minnesota Department of Civil, Environmental, and Geo-Engineering and Director of the Nexus Research Group. His current research focuses on understanding the process of network growth, evaluating transportation technology and policy, and modeling travel behavior.