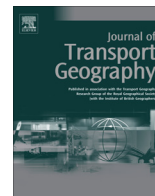




Contents lists available at ScienceDirect

Journal of Transport Geography

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

Catalysts and magnets: Built environment and bicycle commuting

Jessica E. Schoner^a, Jason Cao^{b,*}, David M. Levinson^a^a Department of Civil, Environmental, and Geo-Engineering, University of Minnesota, Twin Cities, United States^b Humphrey School of Public Affairs, University of Minnesota, Twin Cities, United States

ARTICLE INFO

Article history:

Received 29 May 2014

Revised 12 July 2015

Accepted 16 July 2015

Available online xxx

Keywords:

Bicycling

Bicycle commuting

Travel behavior

Mode choice

Self-selection

ABSTRACT

What effects do bicycle infrastructure and the built environment have on people's decisions to commute by bicycle? While many studies have considered this question, commonly employed methodologies fail to address the unique statistical challenge of modeling modes with small mode shares. Additionally, personal characteristics that are not adequately accounted for may lead to overestimation of built environment impacts.

This study addresses these two key issues by using an ordered probit Heckman selection model to jointly estimate participation in and frequency of commuting by bicycle, controlling for demographics, residential preferences, and travel attitudes. The findings suggest a strong influence of attitudinal factors, with modest contributions of bicycle accessibility. Bicycle lanes act as “magnets” to attract bicyclists to a neighborhood, rather than being the “catalyst” that encourages non-bikers to shift modes. The results have implications for planners and policymakers attempting to increase bicycling mode share via the strategic infrastructure development.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

The relationship between bicycling and the built environment, particularly dedicated bicycle lanes and trails, has captivated the attention of researchers and planners for decades. In a state of the practice and research needs paper, Porter et al. (1999) identified critical questions about the role of bicycle infrastructure: how to forecast use of new facilities, how to estimate mode shift due to building new facilities, and how these new facilities may affect mobility, congestion, and air quality. Despite many advances in the field, their questions about the impacts of infrastructure are still salient today.

Common strategies for researching and evaluating transportation projects fail to address the nuances of bicycling. The utility of bicycling, more so than any other mode, is strongly affected by weather phenomena and day-to-day variation in travel needs, such as hauling cargo or goods. As a consequence, many bicyclists are in fact multi-modal travelers (Heinen et al., 2010). Distinguishing between participation and frequency is critical for being able to model the impacts on bicycling (Heinen et al., 2010).

Because bicycling has such a small mode share, standard survey and data collection strategies, especially those that assume people tend to stick to a single mode throughout the week such as the

American Community Survey (ACS), underestimate its prominence. Many also employ research design strategies that skew the sample in favor of people who are already prone to bicycling, producing coefficients that are not accurate for modeling behavior among the general population.

What effects do bicycle infrastructure and the built environment have on people's decisions to commute by bicycle? Are some people more inclined to be “bikers” than others? This study explores the gap in research spanning both participation in bicycling and frequency of bicycle commuting, with aims of expanding the understanding of bicycling and the built environment. Existing survey data from Minneapolis, MN and a sample selection model are employed to jointly estimate participation in and frequency of bicycle commuting as a function of the built environment, controlling for demographics, residential preferences, and travel attitudes.

This research is significant because, while the magnitude and direction of the coefficients are consistent with other studies, the unique structure of the sample selection model provide deeper insight to the relationships between individual preferences and the built environment. When interpreted in this framework, it is easy to identify ways of harnessing the residential self-selection effect to increase rates of bicycling.

The extent to which bicycling infrastructure acts as a “catalyst” to induce mode shift among non-bicyclists to biking is unknown, given the difficulty of establishing causality in cross-sectional

* Corresponding author.

E-mail address: cao@umn.edu (J. Cao).

studies (Cao et al., 2009a). However, the evidence of a self-selection effect suggests that certain infrastructure types function as “magnets” for people who are already prone to bicycling for work, due to their demographic, residential preference, and travel attitude profiles. Combined with evidence from variables used to predict frequency after controlling for residential self-selection, the findings from this study can be used to locate new bicycling infrastructure strategically for providing housing choices for current and would-be bicyclists, and maximizing the number of bicycle trips they choose to make.

This paper is organized as follows: Section 2 reviews literature about how studies manage low numbers of bicyclists among the general population. Section 3 describes the survey administration, data, and modeling procedure. Section 4 presents findings from using an ordered probit Heckman selection model to predict participation in and frequency of bicycle commuting among urban residents. Finally, Section 5 concludes with a discussion of the implications of this research for practice.

2. Literature review

2.1. Complications of modeling bicycling

Bicyclists are distinctly “multi-modal” (Heinen et al., 2010). More so than driving and even transit, bicyclists are vulnerable to day-to-day changes in weather or varying travel needs. Having to make additional stops, carry groceries or other bulky items, or travel when it is dark all decrease the utility of bicycling. Many bicyclists, therefore, can be thought of as “part-time” bicyclists.

This phenomenon results in conventional survey questions underestimating bicycling. Surveys that ask about a single primary commute mode, such as the ACS, miss people who bike only 1–2 days per week, or only for non-work purposes. These questions tell us how many people are bicycling frequently, but do not tell

us how many people are biking infrequently and how many trips this translates to.

Surveys that ask what mode was used “yesterday” in theory should average out over the whole population to a representative value of the amount of bicycling being done, but this assumes bicyclists choose their biking days randomly and that sample sizes are large enough to reflect the ground truth of bicycling. With small sample sizes and such a low mode share, these types of questions have low chances of catching a part-time bicyclist on their biking days.

Much of the literature on bicycling employs binary logit models that predict who is a bicyclist in any capacity, and do not tell us how much bicycling is actually being done. Heinen et al. (2010) describes this problem in this way: “It is of interest to distinguish between (1) mode choice in general, that is to say, the bicycle is at least one of the modes used; and (2) daily choice, in terms of frequency. The latter is useful because many bicycle commuters choose not to cycle every day.”

2.2. Bicycling as a small mode share

Bicycling represents a relatively small mode share, particularly for commuting. In the United States, the ACS estimates that only 0.51% of commuters use a bicycle as their primary commuting mode. While the average is higher when focusing on central cities (0.95% in all Principal Cities, and 3.86% in the City of Minneapolis), the overall rates are still extremely low relative to driving, and even other so-called “alternate” modes such as transit.

A review of the literature on bicycling behavior and the built environment found five strategies for modeling bicycling, given the low mode share. The strategies include inclusion criteria, strategic over-sampling, hybrid inclusion criteria and strategic over-sampling, statistical distributions and no technique.

Table 1 summarizes the studies reviewed in each of these five categories. Some studies appear multiple times in the table

Table 1
Research design techniques for managing low mode share in selected studies of bicycling and the built environment.

Citation	Data source	Technique	Model
<i>No specific technique used – General population</i>			
Cao et al. (2009b)	Original Survey		SURE ^a – Bike/Walk Frequency
Krizek and Johnson (2006)	Regional Survey		Logit – Bike trip(s) in travel diary
Parkin et al. (2007)	Census		Logit – Bike commute share
<i>Inclusion criteria to select bicycling subset</i>			
Handy and Xing (2011)	Original Survey	Biked within past year	Logit – Primary bike commute
Wardman et al. (2007)	Census & Survey	Current/Potential Bicyclists	MNL ^f – Mode Choice
Winters et al. (2010)	Original Survey	Current/Potential Bicyclists	Multilevel Logistic – Bike (vs. car) trip
Xing et al. (2010)	Original Survey	Biked within past year	Logit – Utilitarian v. Recreation Biking
Xing et al. (2010)	Original Survey	Biked within past year	OLS ^b – Log-miles of Utilitarian Bike
<i>Strategic survey to oversample bicyclists</i>			
Akar and Clifton (2009)	Original Survey	University Affiliates	MNL – Mode Choice
Heinen et al. (2011b)	Original Survey	High biking cities	Logit – Bike commute
Hunt and Abraham (2007)	Original Survey	Bicyclists	Logit – SP experiment
Moudon et al. (2005)	Original Survey	Suitable geography	Logit – Biking at least weekly
Sener et al. (2009)	Original Survey	Bicyclists	OLogit ^c – Bike commute frequency
Thakuriah et al. (2012)	Original Survey	Bicyclists	Binary GMM ^d – Former captive car user
<i>Inclusion criteria & strategic survey</i>			
Heinen et al. (2011b)	Original Survey	High biking cities & Cyclists	Logit – FT vs. PT Bike Commute
Heinen et al. (2011a)	Original Survey	High biking cities & PT Cyclists	GEE/RCA Logit – Mode Choice
Rodríguez (2004)	Original Survey	City & University Campus	MNL, Nested, & HEV ^e – Mode Choice
<i>Statistical techniques</i>			
Buehler (2012)	Regional Survey	RELogit	RELogit – Bike commute

^a Seemingly Unrelated Regression Equations (SURE).

^b Ordinary Least Squares Regression (OLS).

^c Ordered Logistic Regression (OLogit).

^d Generalized Mixed Model (GMM).

^e Heteroscedastic Extreme Value Model (HEV).

^f Multinomial Logistic Regression (MNL).

because the paper includes several components employing one or more of the techniques.

2.2.1. Inclusion criteria

Inclusion criteria studies apply a filter to general population data to extract a subset sample that applies to their research question. Many, but not all, of the studies select people based on past bicycling behavior (e.g., having bicycled within the past year) or expressed willingness to bicycle. The effect is that these studies model bicycling behavior among a subset of people already expected to have some propensity to bicycle. The results may not be easily extrapolated to the general population.

The clearest example of this technique comes from [Wardman et al. \(2007\)](#). Within their large sample of census travel diary records and originally collected survey data, they found that about 60% of people indicate that they would “never contemplate switching to cycling”, so these participants were removed from the dataset. The authors explain, “A model is hardly needed to predict the behavior of such individuals and their actual choices or SP responses would provide little information for modeling purposes.” The authors also filter out trip records from individuals with commutes greater than 7.5 miles (about 12 km) because they were only interested in commute trips where bicycling is a viable option. The authors found that dedicated, separate infrastructure was positively and strongly associated with bicycle commuting. However, the results may be exaggerated by the removal of respondents who would not bicycle under any circumstance because their value of time by mode would be heavily skewed to favor time in the car.

[Winters et al. \(2010\)](#), [Handy and Xing \(2011\)](#), and [Xing et al. \(2010\)](#) all used a similar technique for screening out dedicated non-cyclists. [Winters et al. \(2010\)](#) considered anyone with access to a bicycle and who has cycled within the past year a “current cyclist”, and anyone who indicates willingness to cycle in the future as a “potential cyclist”. [Handy and Xing \(2011\)](#) and [Xing et al. \(2010\)](#) used a stricter definition, only including people who had bicycled within the past year.

[Xing et al. \(2010\)](#)'s and [Handy and Xing \(2011\)](#)'s studies use the same data about six small cities in California and Oregon and “past year” inclusion criterion. [Xing et al. \(2010\)](#)'s binary logit model of utilitarian bicycling found that physical environment measures, including shorter distances and more safe destinations, were associated with both a greater share of biking for transportation. While one would expect longer trips to correlate with more mileage biking, the authors explained that long trips are a barrier to making the trip via bicycle in the first place, so the increased distance is offset by reduced probability of making the trip and (presumably) frequency. This finding underscores the importance of understanding both whether people bicycle and how much (duration and/or frequency) they bicycle.

2.2.2. Strategic sampling

Another common strategy for bicycle research and evaluation is to sample deliberately to capture a greater than average proportion of cyclists. Oversampling strategies range from subtle, such as pre-selecting geographies expected to have higher than average rates of bicycling ([Akar and Clifton, 2009](#); [Moudon et al., 2005](#)), to deliberate, such as snowball sampling bicycle clubs and local bike shops ([Sener et al., 2009](#)) or bicyclist intercept surveys ([Hunt and Abraham, 2007](#)).

Like the inclusion criteria studies, over-sampling runs the risk of measuring effects on a concentrated population of people already prone to bicycling. Additionally, the outreach method for contacting bicyclists heavily biases the types of bicyclists who respond. Specifically, bike club members are more likely to fit the “fearless” category in [Geller \(2006\)](#)'s framework. These cyclists

may be less affected by built environment characteristics than occasional cyclists because they are comfortable bicycling in mixed traffic. For modeling the effects of dedicated bike infrastructure on mode shift or bicycling behavior among infrequent cyclists, this strategy may not be appropriate.

[Moudon et al. \(2005\)](#) and [Sener et al. \(2009\)](#) model frequency of bike commuting, though their oversampling strategies and model structures differ substantially. [Moudon et al. \(2005\)](#) pre-selects geographies using Geographic Information System (GIS) that are expected to be conducive to bicycling. They then administered a general population telephone survey via random digit dialing within these geographies. The geography selection component facilitates capturing a higher number of bicyclists than average, while the administration ensures that within those geographies, the sample is relatively representative. Where [Moudon et al.'s](#) sampling strategy was subtle, [Sener et al. \(2009\)](#) took the opposite approach. The researchers administered the survey using a snowballing technique, sending an online link to bicycling clubs, posting the link in local bike shops, and purchasing ads in local papers. The effort captured bicyclists in over 100 cities in Texas (USA), but clearly is not a representative sample of the general population.

One final strategy for oversampling is the intercept survey. Survey administrators reach out specifically to bicyclists while they are biking, either by stopping them along a trail or corridor, or by attaching a paper survey to parked bicycles. [Hunt and Abraham \(2007\)](#), [Hunt \(2009\)](#), and [Thakuriah et al. \(2012\)](#) all make use of this strategy ([Hunt, 2009](#); [Hunt and Abraham, 2007](#); [Thakuriah et al., 2012](#)).

2.2.3. Combined inclusion criteria and strategic sampling

Some studies made use of both of the aforementioned strategies. They first administered a survey using strategic sampling, and then screened their participants using inclusion criteria to focus on a subset of bicyclists. [Heinen et al. \(2011b,a\)](#) sample employees in cities with high rates of bicycling, as described above ([Heinen et al., 2011a,b](#)). However, in some of their analyses, they also filter out non-cyclists. [Heinen et al. \(2011b\)](#) constructs one binary logit model of whether the participant bike commutes Full-time (FT) or Part-time (PT). [Heinen et al. \(2011a\)](#) followed up with PT cyclists from the original study periodically over the course of a year to survey them about how they commuted on that particular day. This set of models (Binary Logit, Generalized Estimating Equation (GEE), and Random Coefficient Analysis (RCA)) attempt to explain day-to-day factors that affect a PT cyclist's choice of mode. From this study, the authors conclude that bicyclists are distinctly multi-modal.

Like [Akar and Clifton \(2009\)](#), [Rodríguez \(2004\)](#) surveys affiliates of a university campus to oversample bicyclists. They further filter their results by selecting only respondents within certain municipal boundaries that are considered “close enough” to be bicycling distance.

2.2.4. Statistical techniques

While relatively uncommon among the literature on bicycling and the built environment, statistical techniques can be used to address the relatively low mode share for bicycling among general population surveys. [Buehler \(2012\)](#) models bicycling for any given commute trip in a large regional travel diary survey using Rare Events Logistic Regression (RELogit). For these individual commute trips, the authors found strong associations with bicycle facilities provided at work, including bike parking. Free car parking was negatively associated with making the commute trip by bicycle.

While specific statistical techniques are used infrequently for modeling the low mode shares associated with bicycling, examples from related fields suggest that these techniques are an opportunity for bicycling research.

Handy et al. (2006) and Schoner and Cao (2014) use Negative Binomial regression to estimate frequency of utilitarian and recreational walking. The Negative Binomial model relaxes the Poisson distribution's strong assumption of variance equal to the mean by adding a separate dispersion parameter. Large numbers of people who do not bike appear in a dataset as excess "zeroes", which in effect appears as overdispersion.

Selection models use a logit or probit function to predict a binary indicator (e.g., participation) jointly with another measure (e.g., frequency). Cao et al. (2008) employ the Heckman Selection model to predict participation in trip-chaining behavior on the evening commute and, given participation, the number of stops comprising that trip chain.

3. Methodology

This study aims to test three hypotheses: (1) Bicycle commuting participation and frequency are associated with proximity to dedicated bicycle infrastructure; (2) bicycle commuting participation and frequency are associated with different factors; (3) jointly estimating bicycle commuting participation and frequency is an improvement over modeling either separately. Heckman Selection models are used to correct for non-random sampling where the dependent variable is only observed among a subset of the sample. Many people would never consider bicycling as an option for commuting, so the frequency of bicycle commuting is only observed among people who are inclined to be bicyclists. An ordered probit model with Heckman sample selection was used to jointly estimate bicycle commuting participation and frequency. The data and measures used for the model are described next.

3.1. Survey administration

The data for this study came from a self-administered ten-page survey mailed in May 2011 to households in five corridors in the Twin Cities as part of a study on the effects of Light Rail Transit (LRT) and associated built environment on travel behavior (Cao and Schoner, 2014). These corridors were selected with the help of local planners because they had similar demographic trends. Three of the corridors are located in South Minneapolis: Nicollet Avenue, Bloomington Avenue, and Hiawatha Avenue from Lake Street to 50th Street. The two remaining corridors were in suburban communities outside the City of Minneapolis: Coon Rapids, 12 miles north of downtown Minneapolis, and Bloomington, 17 miles south of downtown.

For each corridor, we purchased a database of residents from AccuData Integrated Marketing (<http://www.accudata.com>), a commercial data provider. From this database, we drew random samples of about 2000 residents from the Hiawatha corridor and about 1000 residents from each of Nicollet, Bloomington, Coon Rapids, and Burnsville corridors. The database over-sampled longer-term residents as part of a related study, with about half of the residents moving into their homes prior to 2004 (Cao and Schoner, 2014).

The survey was pretested by students and staff members at the University of Minnesota and neighbors and friends of the investigators. Survey content was revised based on the feedback from pre-testers. The survey and two reminder postcards (1 and 2 weeks later) were mailed in May 2011. Ten \$50 gift cards were provided to randomly-selected respondents as the incentive for the survey. The original database consisted of 6017 addresses but only 5884 were valid. The number of responses totaled 1303, equivalent to a 22.2% response rate based on the valid addresses only. This is considered reasonable for a survey of this length, since the

response rate for a survey administered to the general population is typically 10–40% (Sommer and Sommer, 1997).

This study focuses specifically on residents in the three urban corridors. All three urban corridors exhibit traditional urban development patterns: a well-connected street grid, high levels of transit service (LRT in the Hiawatha corridor and bus in Nicollet and Bloomington), a variety of land uses and housing types, and similar built environment context, as shown in Fig. 1. All three corridors contain bicycle infrastructure.

Since this study focuses specifically on bicycling for the journey-to-work commute trip, survey participants who indicated that they are non-employed students or not working (e.g., retired, homemaker, or unemployed) were removed from the sample.

Table 2 compares characteristics of survey respondents and the working sample with the 2011 ACS. Overall, homeowners and households with children are overrepresented among survey respondents due to the oversampling of residents who have lived in their homes since before 2004. Respondents have a higher than average income and less likely to live in a zero-vehicle household. These are typical results for voluntary self-administered surveys.

3.2. Commuting

The dependent variables in this study are an ordinal measure of bicycle commuting frequency and a binary measure of bicycle commuting participation. Survey respondents were asked, "In a typical week with good weather, how many days do you use each of the following as your primary means of transportation between home and work/school?". Available modes included teleworking, driving alone, carpooling, transit, walking, biking, and other. For each mode, they were presented with six ordinal categories: (1) "Never", (2) "Less than once per month", (3) "1–3 days per month", (4) "Once per week", (5) "2–3 days per week", and (6) "4–5 days per week". Bicycle commute participation was defined as any respondent who indicated that they commute by bike at least occasionally (categories 2 through 6). Participants were asked to report additional details about their commute, including the distance in both miles and minutes and whether their employer/school provides free parking.

The lower half of Table 2 focuses specifically on the journey to work mode for survey respondents and the general population. The ACS asks respondents to indicate their single primary mode of transportation for the commute trip (McKenzie and Rapino, 2011), assuming that people use only a single mode each and every day. This assumption is problematic for all commuters, and particularly for bicyclists because they experience barriers such as weather events or the need to transport goods that change from day to day (Heinen et al., 2011b).

3.3. Independent variables

Table 3 presents the hypothesized relationship and descriptive statistics for all independent variables considered in this study, and the following sections explain how each variable is measured. Variables were tested in both portions of the Heckman selection model, and iteratively removed in a backwards stepwise fashion based on lowest Z-score until only variables significant at $p < 0.1$ remained in either the participation or frequency equation.

3.3.1. Built environment

Built environment characteristics were measured in two ways: (1) through a set of survey questions asking respondents to indicate how true each of 29 neighborhood characteristics was of their neighborhood, and (2) using a GIS to objectively measure the infrastructure and land use around their homes.

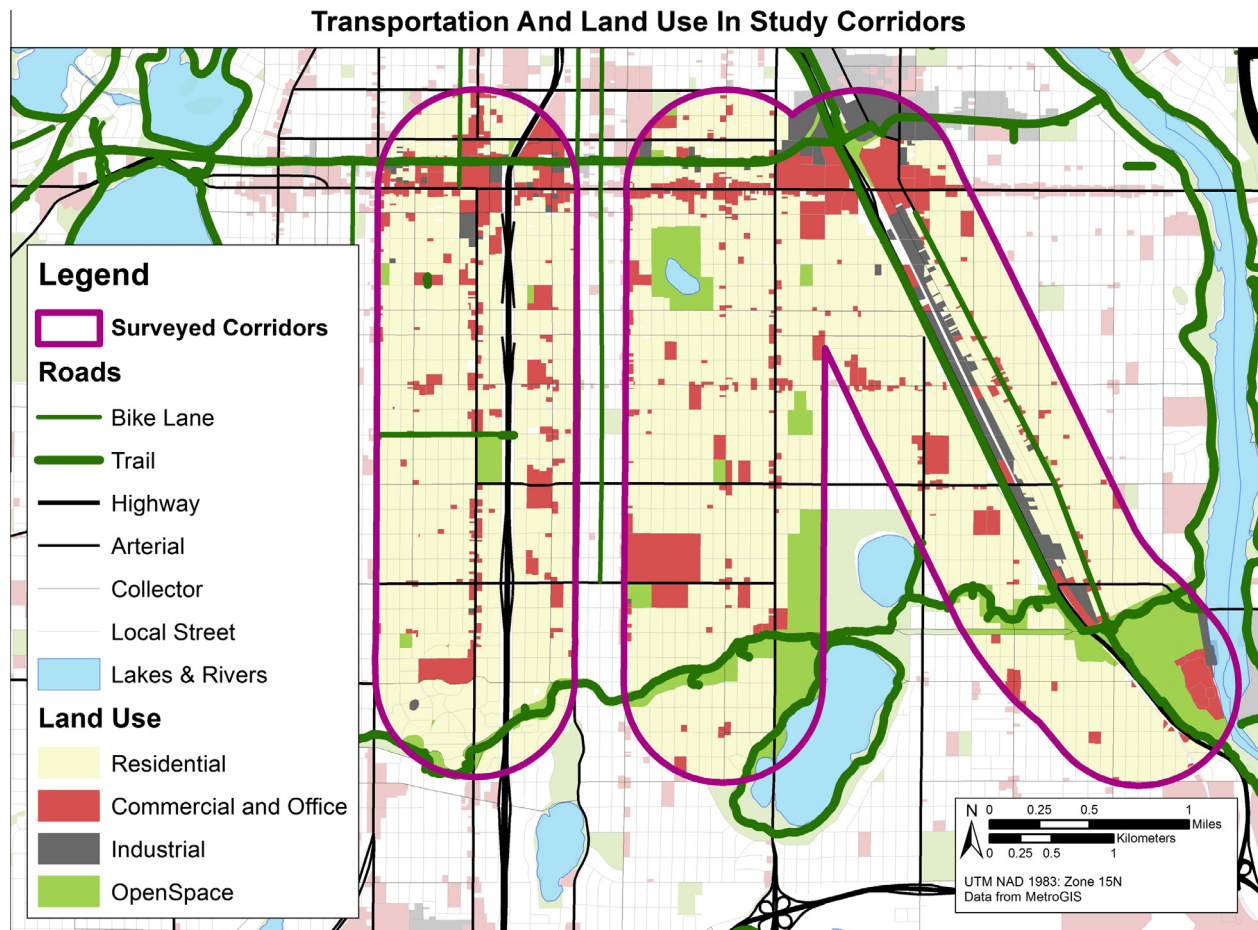


Fig. 1. Street network, bicycle infrastructure, and land use in three urban study corridors.

Table 2
Demographics & commute mode split for respondents and general population.

	2011 ACS		2011 Survey
	City	Tracts	Sample
Population	381,833	113,614	614
Pct. Female	49.8%	50.0%	50.8%
Avg. HH Size	2.3	2.3	2.3
Pct. HH with Kids	24.2%	28.8%	26.1%
Pct. Owner Occupied	50.4%	57.5%	86.1%
Median Income	\$ 47,478	\$ 50,231	\$ 75,561
Pct. Fulltime	58.1%	62.6%	85.0%
Pct. Part time	23.9%	21.1%	15.0%
Pct. Not Working	17.9%	16.3%	0%
Avg. Vehicles/HH	1.7	1.7	1.7
Pct. Zero-Vehicle HH	8.8%	8.1%	4.0%
<i>Percent primarily commuting by^a</i>			
SOV	61.4%	60.9%	63.7%
Carpool	8.5%	11%	4.2%
Transit	14%	16.1%	14.4%
Bike	3.9%	4.2%	8.5%
Walk	6.4%	2.6%	3.8%
Other	0.9%	0.9%	1.3%
Telecommute	4.9%	4.4%	5.6%

^a Results displayed here do not sum to 100% due to estimation procedure used to make survey results more comparable to ACS.

Respondents rated how true each of 29 characteristics, such as “Large back yards” and “Easy access to transit stop/station”, was of their current neighborhood. The ordinal scale ranged from (1) “Not at all true” to (4) “Entirely true”. The two primary characteristics included in this study are “Good bicycle routes beyond the

neighborhood” and “Close to where I work”. Two additional characteristics, “Low crime rate within neighborhood” and “Low level of car traffic on neighborhood streets,” were tested but found to be insignificant.

To evaluate the objectively-measured built environment and its impacts on travel behavior, we constructed network distance buffers around each participant’s homes at about 400-, 800-, and 1600-m ($\frac{1}{4}$ -, $\frac{1}{2}$ -, and 1-mile) distances. The network distance buffer includes only areas that the respondent could actually walk to.

Bicycle facilities around each respondent’s home were measured by summing the total distance of bike lanes in meters within each respondent’s network distance buffers.

An additional measure of bicycle cumulative opportunity accessibility summed the number of jobs accessible by bicycle within 10 min using (OpenTripPlanner Development Team, 2013), assuming an average bicycling speed of 16.1 km per hour (10 miles per hour). Data inputs were 2010 Longitudinal Employer-Household Dynamics (LEHD) job estimates by census block and an OpenStreetMap (OSM) street network shapefile.

3.3.2. Residential preferences

Respondents rated how important each of the 29 characteristics was when they were last looking for a place to live. The ordinal scale ranged from (1) “Not at all important” to (4) “Extremely important”. Respondents could also choose “I never considered it”. Three variables were considered for the analysis due to their expected relationship with cycling: residential preference for good bicycle routes, close proximity to work, and living on a cul-de-sac. Although judgment about “good” bicycle routes is subjective from

Table 3
Variables with descriptive statistics.

Variable	Description	Hypothesis	Mean	(S.D.)
Y	Number of days in a typical month with good weather that respondent commutes by bicycle overall	Dependent variable	1.90	(4.58)
Y_b	Days per month commuting by bicycle for bicyclists		7.26	(6.43)
A	Jobs accessible by bike within 10 min (1000's)	+	0.61	(0.52)
C	Number of children under 12 in household	-	0.43	(0.83)
D	Respondent has college degree or higher	+	0.74	(0.44)
E_d	Commute distance in km	-	13.37	(13.17)
E_p	Employer provides free parking	-	0.69	(0.46)
E_t	Respondent works part time	-	0.15	(0.36)
F_b	Pro-biking Factor	+	0.29	(1.06)
F_d	Pro-driving Factor	-	-0.12	(1.15)
F_u	Pro-travel Factor	+	-0.03	(1.26)
G	Respondent's age in years	-	45.14	(12.64)
H	Land use entropy within 1600 m	+	0.35	(0.11)
I_b	Residential preference for "Good bicycle routes"	+	3.03	(1.07)
I_c	Residential preference for "Living unit on cul-de-sac"	-	1.36	(0.74)
I_w	Residential preference for "Close to where I work"	+	3.16	(0.88)
K	Income (\$1000)	-	7.56	(3.34)
L	Respondent has a limitation that makes biking difficult	-	0.05	(0.21)
N_l	km of bike lane within 1600 m	+	3.37	(2.45)
N_t	km of bike trail within 1600 m	+	3.84	(2.79)
N_{4w}	Intersection density within 400 m	+	0.19	(0.03)
P_b	Perception: Good bicycle routes beyond the neighborhood	+	3.66	(0.62)
P_c	Perception: Living unit on cul-de-sac rather than through street	-	1.13	(0.52)
P_w	Perception: Close to where I work	+	2.87	(1.06)
V	Fewer than 1 car per driver in household	+	0.18	(0.38)
W	Respondent is female Transit	-	0.51	(0.50)

the respondent's perspective, Fig. 1 provides context for bicycle infrastructure availability near residents.

3.3.3. Travel attitudes

To measure attitudes regarding travel, the survey asked respondents whether they agreed or disagreed with a series of 21 statements on a 5-point scale from "strongly disagree" (1) to "strongly agree" (5). Factor analysis was then used to extract the fundamental dimensions spanned by these 21 items, since some of the items are highly correlated. Seven underlying dimensions were identified: pro-drive, pro-walk, pro-bike, pro-transit, safety of car, status of car and pro-travel (Cao, 2015). Pro-bike (F_b), pro-travel (F_u) and pro-drive (F_d) were selected for use in both models to control for travel attitudes.

4. Results

Table 4 shows the results from modeling participation and frequency of commuting by bicycle using ordered probit regression with Heckman selection. First, we explain how to read the table. Step 1 includes only demographic and workplace characteristics, and the subsequent models add residential preference, travel attitude, and built environment (including bicycle lanes) variables. The two components of the Heckman selection model are shown

separately, with frequency on top and participation/selection below.

4.1. Model fit

The fourth section of the table shows the ρ parameter from the Heckman Selection model. ρ is insignificant in all four model steps, which means that we can model the participation and frequency separately. That is, our models do not support the third hypothesis.

There is no directly comparable measure in probit or ordered probit regression to the R^2 used in OLS, which represents the percent of variation in the dependent variable that can be explained by the independent variables. The McFadden's pseudo- R^2 takes on values from 0 to 1, but it represents the relative improvement of this model's log likelihood over that of a null or constant-only model. The McFadden's pseudo- R^2 values range from 0.065 in the first step (demographics and workplace) to 0.337 in the final step (built environment). The largest increase in pseudo- R^2 occurs when travel attitude variables are added in the third step, suggesting that these are the most influential components of the model.

4.2. Demographics and workplace characteristics

The coefficients on demographic and commuting variables are all intuitive, given existing literature on bicyclists, but their final placement in the participation or frequency equations after the backwards stepwise modeling procedure provide additional insight.

Longer commutes (E_d) decrease the probability that someone will commute by bicycle, but are not associated with frequency. It suggests the possibility that individuals have a preferred distance beyond which they would not consider bicycling, but distances within that threshold do not affect how often they choose to bike. Age is negatively associated with probability of participation, consistent with literature about cycling. Being a part-time employee (versus full-time) is positively associated with participation in bicycle commuting.

Free parking (E_p) is associated with a decrease in the frequency with which a respondent commutes by bicycle. Given the spatial distribution of free parking in Minneapolis, however, it may be that this variable is simply serving as a proxy for working in Downtown Minneapolis or at the University of Minnesota since the respondents' actual work locations are not available in a geocoded format. Bicycle infrastructure connectivity to Downtown and the University of Minnesota is very strong, with several major north-south bike lanes and the Hiawatha LRT Bike Trail connecting the study area to Downtown. Income is negatively associated with frequency of bicycle commuting.

4.3. Residential preferences

The positive coefficients on the importance of bike routes beyond the neighborhood (I_b) suggests a possible self-selection effect. People who prefer to live near bicycle facilities are more likely to commute by bicycle, and they do so more often.

4.4. Travel attitudes

Both the Pro-bicycling factor (F_b) and the Pro-travel factor (F_u) are positively and significantly associated with participation and frequency. This is unsurprising, but still noteworthy because they represent such a large contribution in this model's explanatory power. The pro-travel factor contained sentiments about enjoying the journey as much as reaching the destination and valuing time spent in travel (versus believing it to be wasted). Utility theory for

Table 4
Results of Heckman selection ordered probit regression.

		Demographics & workplace		Residential preference		Travel attitudes		Built environment	
		Coef	SE	Coef	SE	Coef	SE	Coef	SE
<i>Ordered probit equation for frequency of bicycle commuting</i>									
<i>K</i>	Income (\$1000)	−0.081***	0.026	−0.100***	0.027	−0.089***	0.028	−0.084***	0.028
<i>E_p</i>	Free parking at work	−0.364**	0.168	−0.403**	0.168	−0.440**	0.176	−0.438**	0.179
<i>I_b</i>	Pref bike route			0.438***	0.117	0.227*	0.116	0.214*	0.117
<i>F_b</i>	Pro-bike factor					1.159***	0.176	1.121***	0.183
<i>F_u</i>	Pro-travel factor					0.210***	0.070	0.203***	0.071
<i>A</i>	Accessibility							0.269*	0.143
<i>Probit selection equation for participation in bicycle commuting</i>									
<i>G</i>	Age	−0.022***	0.005	−0.024***	0.005	−0.018***	0.006	−0.016***	0.006
<i>E_d</i>	Commute distance (km)	−0.045***	0.008	−0.051***	0.008	−0.058***	0.010	−0.057***	0.010
<i>E_t</i>	Part-time employee	0.131	0.158	0.200	0.167	0.346*	0.208	0.381*	0.212
<i>I_b</i>	Pref bike route			0.493**	0.066	0.155*	0.084	0.160*	0.084
<i>F_b</i>	Pro-bike factor					1.023***	0.101	1.026***	0.101
<i>F_u</i>	Pro-travel factor					0.170***	0.057	0.167***	0.057
<i>F_d</i>	Pro-drive factor					−0.128*	0.067	−0.120*	0.068
<i>N_t</i>	km of bike lane							0.072**	0.030
	Constant	0.828***	0.233	−0.603*	0.313	−0.420	0.383	−0.769*	0.410
	Cut 1	−1.094***	0.418	0.186	0.627	0.642	0.564	0.705	0.562
	Cut 2	−0.586	0.381	0.722	0.600	1.345**	0.544	1.422***	0.548
	Cut 3	−0.244	0.358	1.078*	0.582	1.837***	0.531	1.927***	0.541
	Cut 4	0.449	0.319	1.798**	0.554	2.773***	0.521	2.887***	0.538
ρ	Rho	0.411	0.245	0.383	0.231	0.314	0.284	0.250	0.273
	McFadden pseudo R^2	0.065		0.125		0.329		0.337	

$N = 614$; Censored Observations: 453; Uncensored Observations: 161.

* Significant at the $p < 0.1$ level.

** Significant at the $p < 0.05$ level.

*** Significant at the $p < 0.01$ level.

predicting mode share assumes that people will choose the mode that minimizes their cost and time investments, but these findings suggest that bicyclists derive value specifically from their commute. This is consistent with literature on positive utility of commuting in which people value mental separation from work and self-report non-zero ideal commute distances (Redmond and Mokhtarian, 2001). Further supporting this finding is a recent study by Paige Willis et al. (2013), in which bicyclists in particular derive satisfaction from many aspects of their commute.

Additionally, the pro-drive factor (F_d) is significant and negative in the participation equation, but not the frequency equation.

4.5. Built environment

Of all the built environment variables hypothesized to have a relationship with bicycle commuting, only bike lanes (N_t) and job accessibility (A) are significant. Interestingly, bike lanes appear in the participation portion of the equation, while job accessibility appears in the frequency portion of the equation. Bicycle lanes are more strongly associated with *whether* a person bike commutes at all than *how much* they do so.

The coefficient for job accessibility (A) suggests that close proximity to jobs is an important predictor in how frequently one can make that commute trip by bicycle. It also probably serves as a proxy for other built environment variables that were pushed out of the model due to multicollinearity, such as density.

5. Discussion and conclusions

This study tested three hypotheses, with mixed results:

1. Bicycle commuting participation and frequency are associated with proximity to dedicated bicycle infrastructure ⇒ **Participation is associated with bike lanes.**
2. Bicycle commuting participation and frequency are associated with different factors ⇒ **Different variables were significant in the participation and frequency equations, but we do have three common attribudinal variables among the four ones.**
3. Jointly estimating bicycle commuting participation and frequency is an improvement over modeling either separately ⇒ **ρ is insignificant, suggesting no improvement from jointly estimating participation and frequency.**

5.1. Implications for practice

The model results show that bicycle commuting participation and frequency are associated with different built environment measures. However, interpreting these results requires some measure of caution. As discussed extensively in other literature, establishing causality between travel behavior and the built environment is challenging (Cao et al., 2009a; Mokhtarian and Cao, 2008). The residential preference for bicycle routes (I_b) helps control for residential self-selection effects that confound an observed relationship between built environment characteristics and commute mode choice by predicting participation (self-selection) separately from frequency.

Given the observed residential self-selection in the model, built environment variables can have two possible interpretations. They may be “magnets”, attracting bicycle-inclined individuals to an area that meets their preferences, or they may be “catalysts”, sparking a change in propensity or frequency of bicycling. For example, the significant association with residential preference for bicycle lanes suggests that people who are inclined to bicycle

deliberately move into neighborhoods that support their preferred mode of travel. However, the significance of bike lanes themselves *after* controlling for residential preference implies a potential catalyst effect.

Bicycle infrastructure, measured as kilometers of bike lane within one kilometer of the respondent's home, is significant in the participation portion of the model. However, subsequent bivariate correlation tests (not shown) between survey respondents' pro-bike travel attitude factor (F_b) and their length of tenure in a neighborhood (stratified by age) failed to provide evidence of temporal precedence. Travel attitudes are stronger among residents who moved into their current home more recently, suggesting that travel attitudes *precede* location choice. This suggests that bicycle infrastructure functions more like a magnet than a catalyst. While evidence to infer causality between location choice and infrastructure is lacking, it is evident that people who are more likely to use a bicycle for commuting do in fact live near these facilities. This is an important finding because it implies that placing new bicycle infrastructure around other built environment characteristics that do appear to influence bicycling will magnify their effects by attracting residents with a propensity to commute by bicycle.

In the frequency half of the equation, job accessibility and free parking both have significant coefficients. The relationships to job accessibility is intuitive: respondents with greater accessibility to jobs by bicycling have a higher probability of working within a reasonable bike distance from home. Job accessibility is also correlated with other built environment features that support nonmotorized travel, such as density.

While it is possible that free parking directly influences how often a respondent commutes by bicycle by decreasing the cost of driving, this is probably not its only mechanism of action. Pay-for-parking is located primarily in Downtown Minneapolis and around the University of Minnesota campus, whereas free parking is the norm in most other parts of the city. There are some exceptions, but the general trend suggests a probable association between a lack of free parking at the respondent's work and that workplace being located either Downtown or at the University. Connectivity to Downtown and the University via bicycle is excellent due to several closely-spaced major north-south bicycle routes, whereas routes throughout the rest of the city have fewer and inconsistent dedicated facilities.

The accessibility and parking variables have significant implications for practice, given the finding about bicycle lanes serving as magnets. After controlling for self-selection into close proximity of bike lanes, these two variables still have a significant relationship with frequency of commuting by bicycle.

If a city aspires to increase overall levels of bicycling, these results suggest that new bicycle infrastructure should be deployed in neighborhoods with high accessibility to employment, and the routes should be designed to provide connections to major job centers. Table 5 explains the proposed mechanisms of action for

catalysts and magnets. Magnets alone attract bicycling-inclined residents but lack the relative advantages that enable cyclists to bike more than they would elsewhere. Catalysts alone may encourage cycling among existing residents, but are limited by the latent propensity of existing residents to bicycle. As new residents self-select into the neighborhood because of the bike lane (magnet), other catalyst factors will enable them to bike more frequently than they would in neighborhoods lacking these features. In short, build magnets in neighborhoods that already have catalysts to amplify the effect of each.

5.2. Limitations and areas for further study

The issue of free parking at a respondent's work location raises several questions about how the route choices available to a person might influence their mode choice. This study measured built environment characteristics around participants' homes, but not their work locations nor along possible routes connecting the two. Previous studies have found that characteristics along the route are stronger predictors of nonmotorized travel (Srinivasan, 2002; Winters et al., 2010), and the free parking variable in this study demonstrates the need of exploring this issue in greater detail.

Future research should identify respondents' work locations and construct separate measures for bicycle infrastructure along the routes in between home and work. A measure of job accessibility that uses a modified network with stronger weights on jobs accessible via a route comprised mostly of dedicated infrastructure might serve as a sufficient proxy, if work locations are not available. While this still would not resolve whether infrastructure serves as a catalyst or magnet, it would clarify the contexts that make infrastructure relevant to travel decisions.

We modeled the participation component of bicycle commuting as a choice between being a bicycle commuter, or not being a bicycle commuter who uses any and all other modes. Some studies consider bicycling versus a specific choice, such as driving (Handy and Xing, 2011). Focusing specifically on bicycling versus driving removes possible dampening effects on the results from consolidating walking, transit, and driving into a single "non-bicycle" mode category. Built environment characteristics that support walking and transit have more in common with bicycle-friendly spaces than car-supportive environments. However, walking and transit are also small mode shares, so any dampening effect may not have much of an impact. Additionally, despite research and policy that collapses all non-auto modes into a broader category of "alternative transportation", there are distinct differences in how each mode functions and what needs it serves. Comparing bicycling to all other modes in aggregate, as was done here, has a lower potential to overestimate results.

The survey that produced data for this study asked respondents to self-report their estimated bicycle commuting frequency "in a typical month with good weather". This limits the analysis in two ways. First, self-reporting may be subject to social desirability bias, where people over-report what they perceive to be altruistic behavior. Second, focusing exclusively on good weather limits the range of seasons and geographies where the results may be applicable.

Finally, this study was not able to establish causality in the relationship between bicycling and the built environment. Cross-sectional data, as used in this paper, is notably weak in this regard. As previously mentioned, a bivariate correlation test between the strength of travel attitudes length of tenure in a neighborhood confirms that temporal precedence is a likely issue in this study. Nonetheless, the rest of the evidence presented in this study considerably advances the conversation about self-selection and the built environment.

Table 5
Catalyst and magnet mechanisms of action.

		Catalyst	
		No	Yes
Magnet	No	N/A	Affects whoever lives there already regardless of bicycle propensity
	Yes	Attracts bike-inclined people but lacks other BE support for cycling	Attracts bike-inclined people and provides BE support for cycling

Acknowledgements

The study was funded by the Transitway Impact Research Program (TIRP) in the Minneapolis–St. Paul Metro Area. We also thank Hennepin County and the City of Minneapolis planning staff for their help and support.

References

- Akar, G., Clifton, K.J., 2009. Influence of individual perceptions and bicycle infrastructure on decision to bike. *Transp. Res. Rec.* 2140, 165–172.
- Buehler, R., 2012. Determinants of bicycle commuting in the Washington, DC region: the role of bicycle parking, cyclist showers, and free car parking at work. *Transp. Res. Part D: Transp. Environ.* 17 (7), 525–531.
- Cao, X., 2015. Examining the impacts of neighborhood design and residential self-selection on active travel: a methodological assessment. *Urban Geogr.* 36, 236–255.
- Cao, X., Mokhtarian, P.L., Handy, S.L., 2008. Differentiating the influence of accessibility, attitudes, and demographics on stop participation and frequency during the evening commute. *Environ. Plann. B: Plann. Des.* 35 (3), 431–442.
- Cao, X., Mokhtarian, P.L., Handy, S.L., 2009a. Examining the impacts of residential self selection on travel behaviour: a focus on empirical findings. *Transp. Res. Part D: Transp. Environ.* 13 (3), 359–395.
- Cao, X., Mokhtarian, P.L., Handy, S.L., 2009b. The relationship between the built environment and nonwork travel: a case study of Northern California. *Transp. Res. Part A: Policy Pract.* 43 (5), 548–559.
- Cao, X., Schoner, J.E., 2014. The influence of light rail transit on transit uses: an exploration of station area residents along the Hiawatha line in Minneapolis. *Transp. Res. Part A* 59 (3), 134–143.
- Geller, R., 2006. Four Types of Cyclists. Technical Report. Portland Office of Transportation, Portland, OR.
- Handy, S.L., Cao, X., Mokhtarian, P.L., 2006. Self-selection in the relationship between the built environment and walking: empirical evidence from Northern California. *J. Am. Plann. Assoc.* 72 (1), 55–74.
- Handy, S.L., Xing, Y., 2011. Factors correlated with bicycle commuting: a study in six small US cities. *Int. J. Sustain. Transp.* 5 (2), 91–110.
- Heinen, E., Maat, K., van Wee, B., 2011a. Day-to-day choice to commute or not by bicycle. *Transp. Res. Rec.* 2230, 9–18.
- Heinen, E., Maat, K., van Wee, B., 2011b. The role of attitudes toward characteristics of bicycle commuting on the choice to cycle to work over various distances. *Transp. Res. Part D: Transp. Environ.* 16 (2), 102–109.
- Heinen, E., van Wee, B., Maat, K., 2010. Commuting by bicycle: an overview of the literature. *Transp. Res. Part D: Transp. Environ.* 14 (1), 59–96.
- Hunt, J.D., 2009. 'An Examination of Bicycle Use Sensitivities Over Time Using Stated Preference'. IAPR Technical Paper Series TP-06015.
- Hunt, J.D., Abraham, J.E., 2007. Influences on bicycle use. *Transportation* 34 (4), 453–470.
- Krizek, K.J., Johnson, P.J., 2006. Proximity to trails and retail: effects on urban cycling and walking. *J. Am. Plann. Assoc.* 72 (1), 33–42.
- McKenzie, B., Rapino, M., 2011. Commuting in the United States: 2009. Technical Report September. US Census Bureau, Washington, D.C.
- Mokhtarian, P.L., Cao, X., 2008. Examining the impacts of residential self-selection on travel behavior: a focus on methodologies. *Transp. Res. Part B: Methodol.* 42 (3), 204–228.
- Moudon, A.V., Lee, C., Cheadle, A.D., Collier, C.W., Johnson, D., Schmid, T.L., Weather, R.D., 2005. Cycling and the built environment, a US perspective. *Transp. Res. Part D: Transp. Environ.* 10 (3), 245–261.
- OpenTripPlanner Development Team, 2013. OpenTripPlanner. OpenPlans. GNU Lesser Public General License. Version 0.9.1.
- Paige Willis, D., Manaugh, K., El-Geneidy, A., 2013. Uniquely satisfied: exploring cyclist satisfaction. *Transp. Res. Part F: Traffic Psychol. Behav.* 18, 136–147.
- Parkin, J., Wardman, M., Page, M., 2007. Estimation of the determinants of bicycle mode share for the journey to work using census data. *Transportation* 35 (1), 93–109.
- Porter, C., Suhrbier, J., Schwartz, W.L., 1999. Forecasting bicycle and pedestrian travel: state of the practice and research needs. *Transp. Res. Rec.* 1674, 94–101.
- Redmond, L., Mokhtarian, P., 2001. The positive utility of the commute: modeling ideal commute time and relative desired commute amount. *Transportation* 28, 179–205.
- Rodríguez, D.A., 2004. The relationship between non-motorized mode choice and the local physical environment. *Transp. Res. Part D: Transp. Environ.* 9 (2), 151–173.
- Schoner, J.E., Cao, X., 2014. Walking for purpose and pleasure: influences of light rail transit, the built environment, and residential self-selection on pedestrian travel. *Transp. Res. Rec.: J. Transp. Res. Board*, 67–76.
- Sener, I.N., Eluru, N., Bhat, C.R., 2009. Who are bicyclists? Why and how much are they bicycling? *Transp. Res. Rec.* 2134, 63–72.
- Sommer, B.B., Sommer, R., 1997. *A Practical Guide to Behavioral Research: Tools and Techniques*, fourth ed. Oxford University Press, Oxford.
- Srinivasan, S., 2002. Quantifying spatial characteristics of cities. *Urban Stud.* 39 (11), 2005–2028.
- Thakuriah, P.V., Metaxatos, P., Lin, J., Jensen, E., 2012. An examination of factors affecting propensities to use bicycle and pedestrian facilities in suburban locations. *Transp. Res. Part D: Transp. Environ.* 17 (4), 341–348.
- Wardman, M., Tight, M., Page, M., 2007. Factors influencing the propensity to cycle to work. *Transp. Res. Part A: Policy Pract.* 41 (4), 339–350.
- Winters, M., Brauer, M., Setton, E., Teschke, K., 2010. Built environment influences on healthy transportation choices: bicycling versus driving. *J. Urban Health: Bull. NY Acad. Med.* 87 (6), 969–993.
- Xing, Y., Handy, S.L., Mokhtarian, P.L., 2010. Factors associated with proportions and miles of bicycling for transportation and recreation in six small US cities. *Transp. Res. Part D: Transp. Environ.* 15 (2), 73–81.