Transportation Research Part D 32 (2014) 15–23

Contents lists available at ScienceDirect

Transportation Research Part D

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

Using the built environment to oversample walk, transit, and bicycle travel

Orion Theodore Stewart [a,b,*], Anne Vernez Moudon [a,c]

[a] Urban Form Lab, University of Washington, 1107 NE 45th Street Suite 535, Seattle, WA 98105, USA
[b] School of Public Health Department of Epidemiology, University of Washington, Box 357236, Seattle, WA 98195, USA
[c] College of Built Environments Department of Urban Design and Planning, University of Washington, Box 355740, Seattle, WA 98195, USA

ARTICLE INFO

Keywords:
Walk
Transit
Bicycle
Land use
Screening tests
Household travel survey
Sampling
Non-motorized travel

ABSTRACT

Characteristics of the built environment (BE) have been associated with walk, transit, and bicycle travel. These BE characteristics can be used by transportation researchers to oversample households from areas where walk, transit, or bicycle travel is more likely, resulting in more observations of these uncommon travel behaviors. Little guidance, however, is available on the effectiveness of such built environment oversampling strategies. This article presents measures that can be used to assess the effectiveness of BE oversampling strategies and inform future efforts to oversample households with uncommon travel behaviors. The measures are sensitivity and specificity, positive likelihood ratio (LR+), and positive predictive value (PPV). To illustrate these measures, they were calculated for 10 BE-defined oversampling strata applied post-hoc to a Seattle area household travel survey. Strata with an average block size of <10 acres within a ¼ mile of household residences held the single greatest potential for oversampling households that walk, use transit, and/or bicycle.

Introduction

Walk, transit, and bicycle travel modes are recognized for their promise to help the transportation sector reduce environmental impacts, increase physical activity, and relieve congestion without sacrificing accessibility (Rastogi, 2011). Recognition of these benefits has resulted in growing support for walk, transit, and bicycle travel modes, especially in metropolitan areas (Moreland-Russell et al., 2013; Pucher et al., 1999). Support often comes in the form of calls for community design and infrastructure improvements to create more "walkable" (Fenton, 2005), "transit-oriented" (Dorsey and Mulder, 2013), or "cycling-friendly" (Pucher et al., 1999) environments. Such strategies recognize the important role that the built environment (BE) plays in influencing travel. Neighborhood BE factors that affect travel fall under the domains of density, diversity, and design (Cervero and Kockelman, 1997). Density captures the intensity of activity through such measures as residential units or jobs per unit area. Diversity captures the variety of land use in an area, such as the ratio of residential units to jobs or the distance from a residential unit to a retail destination. Design refers to street network patterns within an area, such as the percentage of four-way intersections. After socio-demographics and other confounders are controlled, design variables tend to have a greater association with walking and transit use than diversity or density

* Corresponding author at: Urban Form Lab, University of Washington, 1107 NE 45th Street Suite 535, Seattle, WA 98105, USA. Tel.: +1 206 616 8265; fax: +1 206 685 0767.

E-mail address: orions@uw.edu (O.T. Stewart).

http://dx.doi.org/10.1016/j.trd.2014.06.012
1361-9209/

measures (Ewing and Cervero, 2010). The neighborhood BE may also have different effects on sub-groups of the population defined by socio-economic status, health status, and other individual characteristics (Forsyth et al., 2009). In addition to area-based neighborhood BE characteristics, detailed route characteristics of the environment are also important for pedestrian travel (Lee and Moudon, 2006). Pedestrians accessing transit tend to choose paths with more intersections, wider sidewalks, park crossings, more retail, and flatter terrain (Guo, 2009). Bicycle travel is also sensitive to detailed route characteristics such as distance, turn frequency, slope, intersection controls, traffic volumes, and presence of certain bicycle facilities (Broach et al., 2012). For cycling, such route characteristics may be more important than area-based neighborhood BE characteristics (Moudon et al., 2005b; Winters et al., 2010).

Despite growing support for walkable, transit-oriented, and cycle-friendly environments based on a growing body of evidence, mode shares remain low. Recent figures from the 2009 National Household Travel Survey show that 84% of trips in the U.S. are taken by private vehicle (Litman, 2012). Aside from indicating that continued efforts are necessary to make these modes viable to a larger portion of the population, this suggests that travel researchers will continue to obtain a small sample of these "rare populations" of walkers, transit users, and bicyclists in random population-based surveys. Small samples yield a limited set of observations which in turn inhibit analyses and weaken policy-relevant inferences. Some researchers have increased the sample of rare populations through intercept sampling techniques, in which travelers using specific modes are surveyed at places where they are likely to show up, such as at transit stations (Cambridge Systematics, 1996, 2007) or on bicycle or pedestrian facilities (Thakuriah et al., 2012). Other researchers have utilized convenience samples obtained through targeted outreach (Dill, 2009). These non-probability sampling methods fail, however, to provide results that can be generalized to the target population. They cannot be used to accurately describe regional trends in these travel modes and understand the factors that may influence their use. Efficient probability sampling techniques to reach travelers who use uncommon modes has become an area of focus among travel researchers, particularly those studying walk, transit, and bicycle modes (Forsyth et al., 2012) as well as Metropolitan Planning Organizations conducting Household Travel Surveys (HTS) where they must obtain a statistically stable sample and enough walk, transit, and bicycle households to develop accurate travel demand models (NCHRP, 2008).

The association between travel and the BE suggests that BE characteristics could be used to focus efforts to sample rare travelers. For example, reckoning that more walkers could be found in more walkable neighborhoods, the Walkable–Bikable Communities study in Seattle restricted its sample frame to areas defined by high population densities and proximity to retail and service destinations (Lee et al., 2006). Its results however, could not be generalized beyond people living in such places. In contrast, choice-based probability sampling provides an efficient way to focus on rare travelers while achieving a sample that is representative overall. Choice-based probability sampling divides a sample frame into strata, with certain strata being sampled at a higher rate than others (Stopher, 2012). These strata must be mutually exclusive, exhaustive, and based on data predicting that a sampling unit (usually an individual or a household) will be more or less likely to exhibit the travel behavior of interest. Sample weights are then used to adjust for the varying probability of selection in each stratum so that valid inferences can be made to the target population. Because BE and demographic variables are good predictors of travel mode choice, they become ideal stratification variables. In any survey with a spatial sample frame and sampling units that exist in space, such as a regional HTS survey, spatial data that capture area BE and demographic characteristics can easily be joined to sampling units in a Geographic Information System (GIS) and used for stratification. Accordingly, census tract geographies have been used to define strata based on income, race and ethnicity, age, car ownership, commute mode, and residential and employment density, all of which are factors that are known to be associated with higher rates of non-motorized and transit travel (Bricka and Korepella, 2002; NIRPC, undated). These census data are available nationally. Strata are also sometimes based on data developed by planning organizations, such as transportation analysis zone measures of jobs accessible via transit or walking (NuStats, undated), distance to transit service (Cambridge Systematics, 2007; NIRPC, undated; NuStats, undated), and even detailed neighborhood pedestrian audit scores (Cambridge Systematics, 1996).

Despite the common practice of using socio-demographic, land use and BE characteristics to oversample for uncommon travel behaviors, little has been reported on the effectiveness of such strategies in identifying these rare populations. Long ago, in 1994, Portland Metro HTS reported that strata with a good Pedestrian Environment Factor performed well in identifying populations of people who walked, biked, or used transit. Results published in a preliminary report (Cambridge Systematics, 1996) showed that walking, bicycling, and transit mode shares in these strata were as high as 28.5%, 2.5%, and 8.8%, respectively, compared to 10%, 1.6%, and 2.9% across the region.

The premise of the present work is that travel researchers would benefit from more precise guidance in selecting geographic strata for oversampling rare populations of walkers, bicyclists, and transit users. We present formal measures for evaluating the effectiveness of oversampling strata derived from simple characteristics of the built environment. The measures are sensitivity and specificity, positive likelihood ratio (LR+), and positive predictive value (PPV). These measures are commonly used in epidemiology to assess the performance of so-called screening tests (Koepsell and Weiss, 2003). Screening is the "examination of asymptomatic people in order to classify them as likely, or unlikely, to have the disease that is the object of screening" (Morrison, 1982). Thus a mammogram screening takes an otherwise healthy individual and classifies them as likely or unlikely to have breast cancer. Applying the concept of screening for disease to the stratification process for choice-based probability sampling, we can classify the sampling units as likely, or unlikely, to exhibit the behaviors of interest (i.e. walking, using transit, or bicycling) based on their stratum. Once true travel behaviors are known, measures of sensitivity and specificity, LR+, and PPV provide useful information on how effectively a stratification scheme correctly classified individuals likely and unlikely to exhibit the behaviors of interest. Epidemiologists and demographers have long