ELSEVIER

Examining the impacts of residential self-selection on travel behavior: A focus on methodologies

Patricia L. Mokhtarian [a,*], Xinyu Cao [b,1]

[a] *Department of Civil and Environmental Engineering, University of California, Davis, One Shields Avenue, Davis, CA 95616, USA*
[b] *Humphrey Institute of Public Affairs, University of Minnesota, 301 19th Ave. S. Minneapolis, MN 55455, USA*

Abstract

Numerous studies have found that suburban residents drive more and walk less than residents in traditional neighborhoods. What is less well understood is the extent to which the observed patterns of travel behavior can be attributed to the residential built environment itself, as opposed to the prior self-selection of residents into a built environment that is consistent with their predispositions toward certain travel modes and land use configurations. To date, most studies addressing this attitudinal self-selection issue fall into seven categories: direct questioning, statistical control, instrumental variables models, sample selection models, joint discrete choice models, structural equations models, and longitudinal designs. This paper reviews and evaluates these alternative approaches with respect to this particular application (a companion paper focuses on the empirical findings of 28 studies using these approaches). We identify some advantages and disadvantages of each approach, and note the difficulties in actually quantifying the absolute and/or relative extent of the true influence of the built environment on travel behavior. Although time and resource limitations are recognized, we recommend usage of longitudinal structural equations modeling with control groups, a design which is strong with respect to all causality requisites.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Built environment; Causality; Land use; Residential location; Smart growth

1. Introduction

Numerous studies have observed that residents of higher-density, mixed-use ("traditional", "neo-traditional", or "new urbanist") neighborhoods tend to walk more and drive less than do inhabitants of lower-density, single-use residential ("suburban") areas (e.g., Cervero and Duncan, 2003; Crane and Crepeau, 1998; Frank et al., 2006). What is less well understood is the extent to which the observed patterns of travel behavior can be attributed to the residential built environment itself, as opposed to the prior self-selection of residents

---

* Corresponding author. Tel.: +1 530 752 7062; fax: +1 530 752 7872.
E-mail addresses: plmokhtarian@ucdavis.edu (P.L. Mokhtarian), cao@umn.edu (X. Cao).
[1] Tel.: +1 612 625 5671.

into a built environment that is consistent with their predispositions toward certain travel modes and land use configurations. For example, residents who prefer walking may consciously choose to live in neighborhoods conducive to walking, and thus walk more (as found by Handy and Clifton, 2001). Therefore, the observed differences in pedestrian behavior in those two types of neighborhoods may be more a matter of residential choice than travel choice. In other words, residential self-selection may be at work. If so, we are likely to over-estimate the influence of built environment elements on travel behavior when we use land use policies to try to reduce travel, fuel consumption, and emissions (e.g., Kitamura et al., 1997).

In the past few years, this complex issue has been addressed in a variety of ways. This paper describes and critiques the various methodological approaches adopted to date to assess the causal impact of the built environment on travel behavior. A companion paper (Cao et al., submitted for publication) focuses more heavily on the empirical studies employing each approach, describing and evaluating their key findings. A companion report (Cao et al., 2006) contains the essential content of both papers, plus considerable additional detail that could not be incorporated into journal-length articles. In particular, it includes a table summarizing the 28 empirical studies reviewed for this work.

The organization of this paper is as follows: Section 2 describes the self-selection problem, reviewing the prerequisites of causality inference and placing the issue in the context of the built environment and travel behavior. Section 3 analyzes the various methodologies that have been used to address this issue, while Section 4 discusses numerous ways of posing the research question(s) of interest, and highlights the difficulties in actually quantifying the absolute and/or relative extent of the true influence of the built environment on travel behavior. The last section summarizes the review and makes some recommendations for future research.

2. The self-selection problem

As indicated above, previous studies have consistently found a significant association between the built environment and travel behavior. However, association itself is insufficient to establish causality. To *robustly* infer causality, scientific research generally requires at least four kinds of evidence (Schutt, 2004; Singleton and Straits, 2005; for a more extensive discussion of these in this context, see Cao et al., 2006): *association* (a statistically significant relationship), *nonspuriousness* (a relationship that cannot be attributed to another variable), *time precedence* (cause precedes effect), and *causal mechanism* (a logical explanation for why the alleged cause should produce the observed effect).

In attempting to ascertain the extent to which the built environment (BE) *causes* travel behavior (TB), therefore, the goal is to use a methodology that is as robust as circumstances will permit with respect to these types of evidence. It is particularly important to ensure that an observed association between BE and TB is not the spurious result of the fact that unmeasured variables (such as attitudes) are causing both. As shown in Fig. 1, there are in fact a number of plausible relationships among attitudes (AT), BE, and TB, and the chosen methodology will ideally be capable of distinguishing among the various possibilities.

Self-selection in this context refers to "the tendency of people to choose locations based on their travel abilities, needs and preferences" (Litman, 2005 p. 6). Residential self-selection generally results from two sources: attitudes and sociodemographic traits. With respect to the latter, an example of self-selection is that low-income and zero-vehicle households may choose to live in neighborhoods with ample transit service and hence use transit more. In this case, it is not good transit facilities but households' economic constraints that have a true and direct influence on their choice of transit mode. However, since most previous studies have employed multivariate analysis and accounted for the sorting effect of sociodemographic characteristics (e.g., Abreu e Silva et al., 2006; Kitamura et al., 2001), we focus this review on the issue of attitude-induced self-selection. Unless explicitly indicated, residential self-selection in the remainder of the paper refers only to that resulting from attitudinal factors.

In simple mathematical terms, the often-observed relationship between the built environment and travel behavior is generally modeled as taking the form:

$$\text{TB} = f_1(\text{BE}, X) + \varepsilon, \tag{1}$$

where X denotes other observed variables such as sociodemographics, and ε represents the collective influence on TB of all unobserved variables. The problem is that the standard estimation of such functional forms,