

**EXAMINING THE IMPACTS OF RESIDENTIAL SELF-SELECTION ON TRAVEL
BEHAVIOR: METHODOLOGIES AND EMPIRICAL FINDINGS**

Xinyu (Jason) Cao

Humphrey Institute of Public Affairs, University of Minnesota
301-19th Avenue S, Minneapolis, MN, 55455 USA
Voice: 612-625-5671
cao@umn.edu

Patricia L. Mokhtarian

Department of Civil and Environmental Engineering, University of California, Davis
One Shields Avenue, Davis, CA 95616 USA
Voice: 530-752-7062
plmokhtarian@ucdavis.edu

and

Susan L. Handy

Department of Environmental Science and Policy, University of California, Davis
One Shields Avenue, Davis, CA 95616 USA
Voice: 530-752-5878
slhandy@ucdavis.edu

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 nine categories: direct questioning, statistical control, instrumental variables models, sample selection models, propensity score, joint discrete choice models, structural equations models, mutually-dependent discrete choice models, and longitudinal designs. This report reviews and evaluates these alternative approaches. Virtually all of the 38 empirical studies reviewed found a statistically significant influence of the built environment remaining after self-selection was accounted for. However, the practical importance of that influence was seldom assessed. 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.

Keywords: built environment, causality, land use, new urbanism, residential location, smart growth

This report has been summarized in the following two papers:

Cao, X., Mokhtarian, P. L., and Handy, S. L. (2009). Examining the impacts of residential self-selection on travel behavior: A focus on empirical findings. *Transport Reviews*.

Mokhtarian, P. L. and Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research B*, 42 (3), 204-228.

November 2006
Revised June 2008

List of Acronyms

AO: auto ownership
AT: attitudes
ATE: average treatment effect
BE: built environment
IV: instrumental variables
LATE: local average treatment effect
LR: long-run
MTE: marginal treatment effect
OLS: ordinary least squares
RC: residential choice
SEM: structural equations modeling
SR: short-run
TB: travel behavior
TT: treatment effect on the treated
TUT: treatment effect on the untreated
VKT: vehicle-kilometers traveled
WL: walking level

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 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 (as found by Handy and Clifton, 2001), and thus walk more. 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 overestimate the influence of built environment elements on travel behavior when we use land use policies to try to reduce travel, fuel consumption, and emissions. If, for example, someone with an automobile-oriented lifestyle ends up living in a dense, mixed-use neighborhood (perhaps because of financial incentives or because not enough other housing is available to fulfill his preferences), his travel behavior will probably not match that of those who actively want and choose to live in such neighborhoods.

In the past few years, this complex issue has been addressed in a variety of ways. This report describes and critiques the various methodological approaches adopted to date to assess the causal impact of the built environment on travel behavior, and summarizes the empirical findings of those studies. The organization of this report is as follows: Section 2 reviews the prerequisites of causality inference 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. CAUSALITY REQUISITES

According to the Merriam-Webster online dictionary, causality is defined as “the relation between a cause and its effect or regularly correlated events or phenomena”. Causality must be inferred, because we can

only observe an association between events. The association can be categorized into one (or more) of three principles of connection of events: resemblance, contiguity in time or place, and cause or effect (Hume, 1748). Therefore, association itself is insufficient to establish causality. To *robustly* infer causality, scientific research generally requires at least four kinds of evidence: association, non-spuriousness, time precedence (direction of influence), and causal mechanism (Schutt, 2004; Singleton and Straits, 2005).

Association: The presence of a “statistically significant” relationship between two variables (established, for example, through a t-test, chi-squared test, analysis of variance, or correlation) is often taken as evidence of association. While useful as a general principle, statistical significance does not guarantee even a meaningful association, let alone true causality. The apparent relationship may be spurious (see below), or may simply constitute a Type I statistical error, in which the null hypothesis of no relationship is erroneously rejected due to random variation making the relationship appear to be stronger than it really is. The latter situation may well arise in a given study in which numerous statistical tests are conducted, but is less likely to explain results that persist across a number of independent studies, as is the case for the observed association between the built environment and travel behavior.

On the other hand, while a statistically significant association is often taken to be at least a necessary condition of causality (Singleton and Straits, 2005) if not a sufficient one, this is also not guaranteed to be the case. That is, a weak association does not rule out causality. The causal relationship may be strong for one subgroup of the sample but be diluted when tested across the entire sample; controlling for a third variable may unmask a strong association between the first two (Utts, 1999). Further, an insignificant relationship may be the net outcome of causal forces acting in opposite directions and mostly canceling each other, which is quite a different case than that of no significant forces in either direction.

Nonspuriousness: A nonspurious relationship between variables refers to an association that cannot be explained by a third-party (extraneous or antecedent) variable. If a third-party variable happens to cause both a “dependent” variable and an “explanatory” variable, a statistically significant association may exist even if the explanatory variable inherently has nothing to do with the dependent variable. Therefore, to infer causality, we should eliminate rival hypotheses that can explain the observed association between variables (Singleton and Straits, 2005). The land use–transportation literature offers evidence of possible spurious relationships between the built environment and travel behavior. As an example, in the 1995 Nationwide Personal Transportation Survey, it was found that low income households were disproportionately likely to reside in high-density urban areas, and that they were much more likely to walk than their higher-income counterparts (Murakami and Young, 1997). In this case, household income can be a cause of both residential choice and travel behavior, and hence this rival hypothesis weakens the inference of causality between the latter two variables. To establish non-spuriousness in a nonexperimental study, an appropriate method is to show that the relationship still holds when all third-party variables are controlled for (statistical control). In reality, of course, we are seldom able to control for all variables, but we should account for as many variables as possible (Singleton and Straits, 2005).

Time precedence (direction of influence): To infer causality, a cause must precede its effect in time, or at least the direction of influence must be from a cause to an effect (Singleton and Straits, 2005). A causal relationship is “a relationship in which a change in one event forces, produces, or brings about a change in another” (Singleton and Straits, 2005, p. 20). Therefore, a panel study showing that changes in built environment characteristics at one point in time are associated with changes in travel behavior at a later time will offer more direct evidence of a causal link from the built environment to travel behavior than cross-sectional analysis can. Sometimes, however, an anticipatory travel choice may precede residential choice; for example, those moving to suburban neighborhoods may acquire one more car right before they relocate. Thus, although this travel choice is still a result of the residential change, the

temporal order of observed choices is reversed, complicating efforts to identify the direction of causality even when dynamic data are available.

For cross-sectional data, it can be even *more* difficult to tell whether the choice of the built environment precedes travel choice or travel choice precedes residential choice. For example, it is evident that highly-walkable neighborhoods are significantly associated with a large amount of pedestrian travel (e.g., Cervero and Duncan, 2003). A common inference from this association is that the influence is from the built environment to travel behavior through an intervening variable – travel costs. This is a strong causal mechanism from the perspective of transportation economics, as discussed later in this section. Alternatively, however, as mentioned in the Introduction, this association may mean that individuals who walk a lot intentionally choose a highly-walkable neighborhood in which to live. In this case, travel attitudes (walking preferences) are likely to confound this direction of influence.

As shown in Figure 1, travel attitudes may act as either antecedent or intervening factors in the associations between the built environment and travel behavior. Figure 1a illustrates a potentially spurious relationship between walkable neighborhoods and walking behavior, which can be addressed by controlling for walking preference. In Figure 1b, a large amount of walking (which may or may not have very much to do with the built environment) may stimulate or reinforce an individual's preference for pedestrian travel, which may in turn encourage her choice of highly-walkable neighborhoods. In other words, walking behavior (in that model) is likely to be a proxy for walking preference. If we explicitly account for the influence of walking preference, the influence of the walking behavior on the choice of walkable neighborhood is likely to diminish. Further, an individual's current travel behavior is not a logical indicator of her previous walking preference and residential choice (it may well be correlated with prior attitudes that *are* true antecedents of residential choice, but since the degree of that correlation is unknown, using current behavior as a proxy for past attitudes is in effect assuming what one needs to prove).

[Insert Figure 1 here]

Therefore, when only cross-sectional data on the built environment and travel behavior are available, but not attitudes (as is the case in many studies), the influence from the (previously-chosen) built environment to (presently-chosen) travel behavior is generally inferred more strongly than that from travel behavior to the built environment. In that situation, two roles of walking preference can be distinguished. Travel attitudes may again serve as an intervening variable but in the other direction, as shown in Figure 1c. In particular, if travel attitudes are measured at the current time, these attitudes may be more a function of prior residential choice than the reverse (Chatman, 2005). If we fail to take that into account, we may overstate the influence of travel attitudes. Alternatively, as shown in Figure 1d, the built environment may have a primary and direct influence on travel behavior while travel attitudes may be secondary or irrelevant to this link, as most previous studies have implicitly or explicitly assumed. For example, one may walk to many nearby activities, even if reluctantly (counter to preferences), if the built environment makes it too difficult or expensive to drive; conversely, one may drive to many nearby activities, even if reluctantly, if the built environment is not conducive to walking – heavy, fast, noisy, smelly traffic; no or broken sidewalks; no aesthetic appeal; etc.

Causal mechanism: The identification of a causal mechanism between the built environment and travel behavior can provide strong support for a causality inference (Singleton and Straits, 2005) – and conversely, without such a mechanism, even a strongly significant association is unsatisfying (and more likely to be spurious). In our context, Boarnet and Crane (2001) offer an explicit economic explanation of such a mechanism: the built environment influences the price of travel (an intervening variable), through its impact on travel time and other qualities of travel, which in turn influences the consumption of travel. A similar idea is implicit in discrete choice models of travel behavior: the utility of a particular travel

choice – what mode to take or which destination to choose – is influenced by travel time and other characteristics (intervening variables) of the possible choices, characteristics which are influenced by the built environment.

Experimental design is the key to establishing these evidential components. In the context of the built environment and travel behavior, however, the classic before-after random-assignment control group experimental design is impractical because of its prohibitive costs, ethical deficiency, and/or political impossibility. As an alternative, some studies compared changes in travel behavior between individuals who moved to an environment substantially different from their previous neighborhoods (treatment group) and those who did not move (e.g. Handy et al., 2006; Krizek, 2003a). However, residential relocation is not a treatment randomly assigned by experimenters, but is a “self-selected” result of individuals’ changes in employment location, lifecycle, and, importantly, potentially attitudes toward travel. By contrast, another type of temporal change, a deliberate policy intervention (such as creating and promoting safe routes to school, as discussed in Section 3.7), is to some extent an experimental manipulation. However, intervention programs are implemented at specific locations, which themselves are generally not random but rather (often) chosen on the basis of being more deficient on the dimension that the intervention is expected to improve. Further, participants are automatically classified into the treatment or control group based on their residential locations, not randomly assigned.

On the other hand, numerous studies employed cross-sectional data in lieu of longitudinal data capturing such changes in circumstance, but an overwhelming majority were built upon observational design, in which nonrandom assignment/selection bias is a major concern (Mokhtarian and Cao, 2008; Oakes 2004). That is, those who are observed receiving a treatment often differ significantly from those not receiving the treatment. In practice, if selection bias cannot be eliminated through study design, etiological analysis should be applied to address the selection issue (Oakes, 2004). In other words, if an observational study attempts 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 the four types of evidence.

3. METHODOLOGIES AND FINDINGS

At least in most of the developed world, residential location is a matter of choice to some extent, subject to constraints such as income or even still racism. As a matter of course, then, most households select residential locations at least partly “based on their travel abilities, needs and preferences” (Litman, 2005, p. 6). Residential self-selection generally results from two sources: attitudes and sociodemographic traits. An example of self-selection tied to socio-demographic traits occurs when 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; van Acker et al., 2007), we focus this review on the issue of attitude-induced self-selection. That is, what matters here is the degree to which preferences related to the built environment, and more specifically the transportation opportunities the built environment offers, factor into the residential location choice. If households choose residential locations in significant part because of transportation options, then we say that they have “self-selected” with respect to travel preferences. In this case, their travel preferences, rather than the built environment, may be the primary causal factor for their travel behavior. Unless explicitly indicated, residential self-selection in the remainder of this report refers only to that resulting from attitudinal factors.

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

$$TB = f_1(BE, X) + \varepsilon, \quad (1)$$

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, whether the dependent variable is continuous and observed (as in linear regression models) or representing a discrete choice (as in logit or probit models), requires that observed explanatory variables (BE, X) be uncorrelated with unobserved explanatory variables (ε). Failure to meet this important condition is broadly referred to as *endogeneity bias*, and produces coefficients for BE and X that are biased and inconsistent estimators of the true values. Furthermore, the conventionally-estimated standard errors of the estimated coefficients will also be biased, which renders invalid the usual hypothesis-testing on the significance of variables (Ramanathan, 2002). In other words, this problem will occur if attitudes are unmeasured and if they influence residential location, in effect influencing what built environment characteristics are experienced by the individual.

Endogeneity bias can occur in two conceptually distinct ways, either of which could arise in our context. *Simultaneity bias* is produced when an “explanatory” variable is simultaneously a function of the “dependent” variable it is supposed to explain – that is, when one variable is both a cause and an effect of another. In the present context, this would mean:

$$\begin{aligned} TB &= f_1(BE, X, Y) + \varepsilon_1 \\ BE &= f_2(TB, X, Z) + \varepsilon_2, \end{aligned} \quad (2)$$

where X denotes observed explanatory variables common to both TB and BE, and Y and Z denote observed variables distinctive to TB and BE, respectively. In this formulation, travel behavior is assumed to exert a direct influence on residential choice and thus the built environment (as well as the more conventionally assumed converse direction of causality, from BE to TB), separate from the influence of attitudes. This could occur if travel behavior were largely determined by constraints such as income (X) – e.g. making it impractical to own a car – and then residential location were influenced by the resulting travel behavior, e.g. a reliance on public transportation (as well as separately by income also). In models such as these, it is easy to see that BE is likely to be correlated with ε_1 , because of its correlation (through f_2) with

$$TB = f_1 + \varepsilon_1.$$

The second type of endogeneity bias is *omitted variables bias*. This occurs whenever observed and unobserved explanatory variables are directly correlated, either because one causes the other or because both are functions of the same antecedent variables. The most frequently-discussed form of the residential self-selection problem is of this type, and can be expressed as:

$$TB = f_1(BE(AT), X) + \varepsilon(AT), \quad (3)$$

in which the attitude (AT) portion of ε partly explains or causes BE. However, as illustrated by Figure 1c, the opposite direction of causality between BE and AT is also plausible:

$$TB = f_1(BE, X) + \varepsilon(AT(BE)), \quad (4)$$

in which travel attitudes are influenced by the built environment.

A number of methodological approaches have been applied to test and control for this endogeneity bias in previous studies; we discuss nine such approaches in this section. Generic forms of the statistical control, instrumental variables, sample selection, propensity score, and longitudinal approaches are discussed in the excellent review article by Winship and Morgan (1999), a highly-recommended gateway into the more complex econometric literature on the estimation of causal effects in the presence of selection bias. An archetype of the latter is Heckman and Vytlačil (2005). Bhat and Guo (2007) provide a useful discussion of the statistical control, instrumental variables, and longitudinal approaches in the specific context of residential self-selection.

The general format for each of the nine approaches is that we briefly describe the method, then discuss one or more studies exemplifying the method, and then make some analytical and/or critical observations about the basic approach. Thirty-eight relevant studies and their corresponding methodologies are summarized in Table 1. We identified the studies to include based on our knowledge and our connections with worldwide scholars in the field, but do not claim that they are collectively exhaustive. Most studies included are published in the travel behavior literature; a few studies from the field of physical activity are included because they convey important concepts.

[Insert Table 1 here]

3.1 Direct questioning

To assess whether people's travel and land use predispositions influenced their choice of residential neighborhood, why not just ask them? Although this approach may appear primitive next to more complex quantitative approaches, it requires considerable ingenuity to execute well, and done well it can provide very useful insights.

Using 1,368 respondents to a 1995 survey conducted in six neighborhoods in Austin, TX, Handy and Clifton (2001) investigated the potential of providing local shopping as a strategy to reduce auto dependence. Through group discussions with some of the respondents, they found some evidence for residential self-selection and concluded that "having the option to walk to the store [i.e., living in a neighborhood that facilitates walking] is to some extent an effect of the desire to walk to the store" (p. 344).

Hammond (2005) studied the decision process relating residential choice and travel choice. In a self-administered survey, he first asked respondents living in Century Wharf, Cardiff (an isolated, compact, and mid-size provincial city in the UK), to answer questions regarding residential choice and commute mode choice. He concluded that living in the city center is associated with lower levels of auto use. In fact, living in the city center and workplace proximity are the two most important reasons among others for lower car use. Respondents were also asked to describe their decision sequence with respect to residential choice and commute mode choice. He found that 18% of the 90 respondents selected commute mode before making their decisions on residential location, and that 39% chose residence and commute mode simultaneously. This result indicates that for more than half of the sample, residential choice is either conditional on or interacts with commute mode choice. Through an eight-person focus group, he found that participants incorporated commute mode choice and access to work into their residential choice, and that all participants were commuting by the mode (including car, bus, and train) that they had expected to use when looking for a residence (although one participant planned to change mode). Therefore, people selectively locate in a residential neighborhood to realize their travel preferences. However, almost all of these results are not based on statistical tests but on descriptive analysis.

Although the direct questioning approach may appear primitive next to more complex quantitative methods, it requires considerable ingenuity to execute well. Done well, it may offer valuable information

regarding the process of residential and travel choices, sometimes beyond what multivariate analyses can do. It can also be used effectively in conjunction with quantitative approaches, for example in the development of survey instruments, the identification of appropriate model specifications and/or market segments having different decision-making processes, and the validation of multivariate analyses (Clifton and Handy, 2003; Pendyala, 1998). Nevertheless, used on its own it has several limitations. To begin with, the sample size is generally small and may not be representative of the population of interest. Moreover, direct questioning is likely to suffer from a number of biases, including:

- ◆ *memory*: For even very recent moves and certainly for longer-ago ones, attempts to recall one's attitudes and preferences prior to the move will be unreliable, as those beliefs are likely to be altered by the realities of the new residential environment and other intervening events and changes;
- ◆ *consistency*: If we ask about participants' behavior first, they may later (consciously or subconsciously) express attitudes to be consistent with that behavior;
- ◆ *saliency (recency)*: Participants will tend to focus on the aspects of the situation to which their attention is drawn, and thus their responses will be very much flavored by the specific content and tone of the interview questions; and
- ◆ *social desirability*: As the conversation goes along, participants may anticipate the objective of the study and hence conform their expressed attitudes and choices either to what they think the researcher wants to hear, or to established social norms.

Of course, these biases are also possible with the design of the self-administered questionnaires from which the data for quantitative analyses are often collected. A researcher with heavily biased ideas about which variables are important, the terms in which questions should be phrased (e.g., pejoratively referring to certain situations as being "auto-dependent" and focusing only on the negative aspects of such cases, as opposed to using more neutral phrasing and/or giving equal time to the personal benefits of automobile use), and the nature and direction of influence among a set of variables, is likely to find what she looks for, in either case. However, some scholars (e.g. Dillman, 1978) suggest that all else equal, the extent of at least the latter three biases could be more severe in the case of direct questioning, where the body language and tone of the interviewer can offer additional cues to the participants, and where (even in the case of a prepared script or set of questions) the interviewer generally has a certain amount of discretion over the spontaneous digressions that the interview might take.

Equally importantly, direct questioning does not allow us to quantify the respective influences of the built environment and residential self-selection, and determine which is more important. In addition, this approach is vulnerable to most of the limitations discussed in the following sections.

3.2 Statistical control

The method of statistical control explicitly accounts for the influences of attitudinal factors in analyzing travel behavior, by measuring them and including them in the TB equation (thereby moving them from unobserved to observed). This approach has been operationalized in two different ways in the literature, one incorporating attitudes directly, and the other incorporating an attitudinal-based measure of dissonance between one's preferred and actual neighborhood types.

3.2.1 Direct incorporation of attitudes

In this case, TB is modeled as a function of AT as well as BE:

$$TB = f_3(BE, AT, X) + \xi, \tag{5}$$

which removes AT from the ε of equations (3) and (4), and thereby presumably eliminates any correlation between BE and ξ . If the inclusion of AT drives the influence of BE into insignificance, the natural

conclusion is that the influence of BE was entirely due to predispositional attitudes. If BE is still significant, the conclusion is that the BE exerts some influence of its own, separate from the predisposition that led an individual to locate there in the first place.

Using data collected from 1,114 adults in the San Francisco Bay Area and the San Diego metropolitan area in 2003, Chatman (2005) studied the confounding influence of modal (auto, transit, walk/bike) preferences in the relationship between the built environment and nonwork travel. Through negative binomial regressions, he found that respondents who sought transit and walk/bike access (to shops/services and for all travel purposes) were more likely to conduct nonwork travel by transit and walk/bike, respectively, but auto travel was not significantly influenced by auto access preference. After controlling for these attitudinal factors, he also found that living within half a mile of a heavy rail station and bus frequency had an influence on nonwork travel by transit, and bus frequency and number of four-way intersections influenced walk/bike travel. By further incorporating interaction terms of built environment characteristics and modal preference indicators in the models, Chatman found that the effects of built environment characteristics showed little difference between those with strong and weak preferences. Chatman also modeled non-work auto mileage as a function of built environment traits and modal preferences, but he did not find any meaningful influence of the preferences. Chatman concluded that the residential self-selection problem is not a big concern, at least for his dataset.

Kitamura et al. (1997) incorporated attitudinal measures into the specification of linear regression models of travel behavior. This study explored the effects of both the built environment and attitudinal characteristics on disaggregate travel behavior for about 800 residents in five neighborhoods in the San Francisco Bay Area in 1993. They first regressed sociodemographic and neighborhood characteristics against frequency and share of trips by mode. Measurements of residential density, public transit accessibility, mixed land use, and the presence of sidewalks were found to be significantly related to mode choice and trip generation by mode, controlling for sociodemographic characteristics. After attitudinal measures were incorporated as explanatory variables in the model, they found that attitudes explain travel behavior better than neighborhood characteristics, which lends some support to the self-selection speculation. However, several built environment characteristics (parking spaces available, distance to nearest bus stop, and distance to nearest park) remained significant in the model for fraction of trips by auto, even after including attitudinal variables.

Cao et al. (2006a) investigated the determinants of trip frequencies for two types of pedestrian travel: strolling and walking to the store, using the same data as Handy and Clifton (2001). Two separate negative binomial models showed that although residential self-selection (measured as a preference for stores within walking distance when households were looking for a place to live) impacts both types of trips, it is the most important factor explaining walking to a destination (i.e. for shopping) among the variables tested. However, after accounting for the influence of self-selection, neighborhood characteristics, especially perceptions of various characteristics, impact strolling frequency, while characteristics of local commercial areas are important in facilitating shopping trips. Similar to the previous one, this study indicates that residential self-selection at least partially contributes to differences in pedestrian behavior, but that the built environment does exert a separate influence beyond that. However, the single attitude measurement included may not have completely captured the influence of self-selection (e.g., a preference for recreational strolling was not measured). To the extent that unmeasured influences were at work, their models may overstate the influence of the built environment.

To overcome this limitation, Cao, Handy, and Mokhtarian measured more than 12 dimensions of residential preferences and travel attitudes in a new research design. Using data collected from 1,682 respondents in Northern California in 2003, Handy et al. (2005; 2006), and Cao et al. (2005; 2006b; 2007a) explored the influence of the built environment and residential self-selection on driving behavior, walking behavior, nonwork travel behavior by various modes, vehicle type choice, and auto ownership

decisions, respectively. All studies conducted cross-sectional analyses, and some of these studies also adopted quasi-longitudinal designs (discussed below in Section 3.7). In the cross-sectional analyses, the respective modelling techniques employed were linear regression, negative binomial regression, seemingly unrelated regression, nested logit, and ordered probit. These studies measured the residential built environment both subjectively, through factor analysis of respondents' perceptions of their residential neighborhood, and objectively, through GIS analysis. After controlling for attitudinal and sociodemographic factors, Handy et al. (2006) found that both perceived neighborhood characteristics and objective accessibility variables influence walking to the store frequency, and perceived aesthetic quality and social context of residential neighborhoods affect strolling frequency. Cao et al. (2005) concluded that the built environment has an influence on frequencies of nonwork travel by auto and transit while attitudinal factors have an incremental contribution to explaining the variations in these behaviors, and both the built environment and residential self-selection affect walking/biking nonwork trip frequency. Cao et al. (2006b) found that vehicle type choice is greatly impacted by attitudinal factors, but commute distance and parking availability have a separate influence on the choice of SUVs and pickup trucks, respectively. By contrast, Handy et al. (2005) and Cao et al. (2007a) found that neighborhood characteristics were displaced by preferences for the same aspects when modeling vehicle miles driven and auto ownership, suggesting that the observed associations with neighborhood traits are a consequence of residential self-selection.

3.2.2 *Comparison of consonant and dissonant residents*

The second form of the statistical control approach is to compare the travel behavior of residentially consonant and dissonant individuals. Here, in addition to incorporating travel-related attitudes into the equation for travel behavior, attitudes toward residential location type are used to classify survey respondents as consonant (well-matched) or dissonant (poorly-matched) with respect to their current residential location. The travel behavior of dissonant residents is then compared to that of consonant residents in the type of neighborhood in which they would rather live, and in their current neighborhood. If the travel behavior of dissonant residents is more similar to that of the consonant residents in their *desired* type of neighborhood, it suggests that their predispositions dominate their travel behavior. If their travel behavior is more similar to that of the consonant residents in their *current* neighborhood, it suggests that the built environment exerts a separate influence that outweighs a contrary predisposition. Alternatively, a continuous measure of the degree of dissonance, as well as measures of the built environment, can be incorporated into the travel behavior equation, and tests performed to see whether the built environment remains significant after dissonance is accounted for.

In three studies of a 1998 sample of 1,358 residents of the San Francisco Bay Area, Schwanen and Mokhtarian compared the trip frequency (2003), commute mode choice (2005a), and mode-specific distances traveled (2005b) of dissonant suburban and urban residents (those who preferred a more or less, respectively, dense/diverse neighborhood than the one they currently lived in) to their consonant counterparts in both kinds of neighborhoods. In general, they found that while *suburban* residents' travel behavior was similar whether they were consonant or dissonant, dissonant *urban* residents' behavior fell between that of consonant urban and consonant suburban residents – more auto-oriented than the former but less so than the latter. These findings suggest that the built environment does in fact play a role, at least in constraining and possibly in shaping, one's underlying preferences. Unfortunately for the goal of reducing auto dependence, the role does not appear to be symmetric, and the asymmetry is not in the societally-desirable direction: urban-oriented suburban residents are less able to achieve their preference for non-auto travel than suburban-oriented urban dwellers are able to realize their preference for auto travel. However, in these studies too, residential preferences were captured with a single variable, attitude toward residential density/diversity. Although that attitude was a factor score representing a composite of several different elements (e.g., housing type, having shops and services within walking distance, and yard size), it still leaves room for improved measurement of residential preferences.

Frank et al. (2007) adopted both methods. They first incorporated residential preferences and a walkability index in linear regression models of travel behavior. Then they classified the respondents into a 2 x 2 matrix (based on two binary variables of residential preference and walkability) indicating matched and mismatched residents, and compared their travel behavior. In this study, they applied these techniques to two sub-samples drawn from the 2001-2002 Strategies for Metropolitan Atlanta's Regional Transportation and Air Quality (SMARTRAQ) study. Overall, they found that both residential preference and built environment characteristics affect walking and driving behavior, with particularly the linear regression models for vehicle-miles traveled (VMT) suggesting that the influence of the built environment is stronger. With respect to matched and mismatched residents, somewhat in contrast to Schwanen and Mokhtarian, they found that residents of high walkable areas drove similar amounts on average, regardless of their preferences, while residents of low walkable areas drove less if they preferred pedestrian-/transit-oriented neighborhoods than if they preferred auto-oriented ones. Again, in each sub-sample, residential preference was represented by a single indicator, which is a composite measure of several dimensions of the preferences for the built environment.

Although the statistical control approach can offer insightful evidence of residential self-selection, it is vulnerable to several intrinsic limitations. First, attitudes are not straightforward to measure and analyze, and are often not measured, e.g. not available in standard travel/activity diary data sets, and hence pose significant difficulty in the context of regionwide travel demand forecasting. Even when they *are* measured, they are measured with error, and may not comprehensively capture all the relevant attitudes. Second, when data are cross-sectional, there can be a temporal mismatch: the attitudes measured in the present may differ from those leading to the prior choice of the built environment. Third, these studies modeled only a single causal direction, from the built environment to travel behavior. As illustrated in Figure 1, this is too simplistic a representation of the potential interactions among these variables.

3.3 Instrumental variables models

Another approach to address residential self-selection is to use instrumental variables (IVs) to purge BE of its correlation with ε . A time-honored econometric technique, it involves (as applied in this context) first modeling BE as a function of relevant instrumental variables (or “instruments”), z , that are not correlated with ε , and then replacing the observed BE in equation (1) with its predicted value \hat{BE} from that model:

$$\begin{aligned}
 BE &= b(z) + \eta(AT) \\
 TB &= f_4(\hat{BE}, X) + \varepsilon(AT), \tag{6}
 \end{aligned}$$

where $\hat{BE} = \hat{b}(z)$. The predicted \hat{BE} will then, by construction, be uncorrelated with ε . The implication is that the entire influence of AT on TB will lie in ε ; if \hat{BE} is significant in the equation for TB, it represents an influence of the BE that is purged of the self-selection attitudinal component. Thus, the statistical control and IV methods represent opposite strategies in dealing with the endogeneity problem (Winship and Morgan, 1999): whereas the object of the former method is to identify variables that are *maximally* correlated with ε to use as controls, the object of the latter method is to find variables that are *minimally* correlated with ε to use as instruments.

Boarnet and Sarmiento (1998) employed ordered probit models to estimate nonwork auto trip frequency, using 1993 data from 769 Southern California residents. Population density, retail employment density, service employment density, and street grid patterns at the block group/census tract level and at the zip code level were chosen to measure the built environment. They initially found that none of these built

environment variables were significant in the models. Then they chose four non-transportation neighborhood traits as built environment instruments: percentage of population that was African American, percentage of population that was Hispanic, and percentages of housing built before 1940 and before 1960. After performing instrumental variable regressions, they found that the predicted built environment variables became statistically significant in one specification of the model. In particular, predicted service employment density became significant to nonwork auto trip frequency when both employment densities at the zip code level were instrumented. However, the predicted built environment variables in other specifications were still insignificant.

By contrast, Greenwald and Boarnet (2001) found a different pattern when modeling nonwork walking trip frequency. Using 1,091 individuals from the 1994 Household Activity and Travel Behavior Survey in Portland, Oregon, they employed ordered probit models to test walking frequency against built environment variables and sociodemographic characteristics. The built environment variables were measured at three geographical levels: census block group, census tract, and zip code. They initially found that population density, retail employment density, street grid patterns, and pedestrian environment factor (PEF) score were significantly associated with nonwork walking frequency. Thereafter, they selected six variables as instruments: per capita income in the area (census block group only), percentage of population living in the geographical area with at least a college education, percentage of population that was African American, percentage of population that was Hispanic, percentage of housing units in the area classified as rural but not farms, and percentage of housing units in the area classified as urban dwelling units. After performing instrumental variable regressions, they showed that most predicted built environment variables at the census block group and census tract levels remained significant while those at the zip code level became insignificant. Therefore, they concluded that the built environment influences nonwork walking trip generation at the neighborhood level.

Using 4,328 individuals (whose households have at least one car) in the 1996-2003 German Mobility Panel, Vance and Hedel (2007) investigated the effect of built environment elements on car use (a binary variable) and distance travelled.¹ Following Boarnet and Sarmiento (1998), they chose four non-transportation variables as instruments: the respective percentages of building built before 1945 and between 1945 and 1985, the percentage of senior residents, and the percentage of foreign residents. The results of instrumental variable models showed that commercial density, street density, and walking time to public transportation had true effects on car use and distance traveled.

Instead of using instruments to model a continuous BE variable with linear regression, Khattak and Rodriguez (2005) used them to model a binary residential choice (RC) variable with logit. Using 453 households from a neo-traditional neighborhood in Chapel Hill and a suburban neighborhood in Carrboro, North Carolina, they first developed a binary logit model for neighborhood type choice (pseudo R^2 was 0.27), with residential attitudes as instruments. Then they incorporated the predicted probabilities of neo-traditional neighborhood choice (a new explanatory variable) into three negative binomial regression models for auto trip frequency, external trip frequency, and walking trip frequency, and two linear regression models for trip distance and trip duration. They concluded that households with high predicted probabilities of living in the suburban neighborhood conducted more auto trips and external trips, walked

¹ They incorporated the IV approach into a technique known as a “two-part model (2PM)”, a variant on the Heckman sample selection model discussed in the next section. In both techniques, a first-stage model predicts a binary selector variable (car use, in their case), while a second-stage model predicts the behavior variable of interest, conditional on the selector variable being equal to one. The 2PM differs from the selection model in that the second-stage model doesn’t have any added variables related to the first-stage model (specifically, no inverse Mills ratio), but the first-stage model affects the marginal effects of explanatory variables in the second-stage model on the behavior variable. Vance and Hedel estimated one model system using this approach directly, and a second system in which IVs were used to predict the four urban form variables that were explanatory in both stages.

less, and traveled longer distances than those with high predicted probabilities of living in the neo-traditional neighborhood. However, some if not all instruments that they selected for the RC equation may not be appropriate. Generally, instrumental variables should satisfy two criteria: they must be highly correlated with the endogenous explanatory variable they are predicting (“relevance”), but not be significantly correlated with the error term of the original equation (“exogeneity”; Cameron and Trivedi, 2005; Hall, et al., 1996); in this case, the endogenous explanatory variable is neighborhood choice and the error term reflects unmeasured attitudes. Although Khattak and Rodriguez explicitly stated that they excluded attitudes that are expected to be associated with travel behavior and hence correlated with the error term in an equation for travel behavior, they did not provide any empirical evidence of independence from travel behavior for the attitudes they *did* include. To the contrary, other studies suggest that some of their instruments may be correlated with travel behavior. For example, Cao et al. (2006a) and Handy et al. (2006) found that residential preference for stores within walking distance, a dimension similar to “having shops and services close by is important to me” in Khattak and Rodriguez, is significantly associated with walking frequency.

This discussion illustrates the intrinsic limitations of the IV technique (as stated by Winship and Morgan, 1999, p. 683, “the perfect instrument [is] an apparent contradiction”). The problem is that BE or RC (in this context) must be *substantially* correlated with the error term for the TB equation in order for endogeneity bias to be a problem; small correlations between observed and unobserved variables are tolerated all the time, without remedial measures being required or taken. But in that case, first of all, finding suitably uncorrelated variables with which to model BE (i.e., meeting the exogeneity criterion) can be difficult. Second, modeling BE as a function of variables *uncorrelated* with the error term for the TB equation will therefore necessarily leave a sizable portion of the variance in BE unexplained (thereby falling short on the relevance criterion).

The problem of low relevance or “weak instruments”² occurs quite often with this technique, and has a number of (related) potential deleterious consequences:

The standard error of the coefficient of \hat{BE} in equation (6) is likely to be high (Bound et al., 1995; Shea, 1997), reflecting the imprecision with which the true effect of BE on TB is being captured. In that case, finding \hat{BE} to be insignificant may not reflect a true lack of influence after controlling for self-selection, but rather the inability of the poor \hat{BE} to capture that influence.

² There does not appear to be a consensus in the literature on the definition of “weak”. However, in a linear regression context, the relevant indicators are the R^2 of the first-stage equation estimating BE, and the F-test for the significance of the block of instruments in the same equation. In the special case of one instrument (z) and one regressor (BE), Nelson and Startz (1990) indicate that $\{\text{sample size}\} \times R^2 \ll 2$ is problematic, and Shea (1997) provides a partial- R^2 -based test for the case of multiple instruments and multiple regressors. Bound et al. (1995, p. 446) indicate that an F-statistic near 1 or lower is “cause for concern”, and Staiger and Stock (1997) point to an F-statistic less than 5 as problematic, even with a sample size of several hundred thousand.

Hall et al. (1996) generalize these tests of instrument relevance to one based on the statistical significance of the smallest canonical correlation between instruments and endogenous explanatory variables, with the R^2 - and F-based tests appearing as special cases. In the one-instrument/one-regressor case, their relevance test statistic, $-\{\text{sample size}\} \times \ln(1 - [\text{corr}(z, BE)]^2)$, is asymptotically χ^2 -distributed with 1 d.f. However, they caution against using such measures as an a priori screening device to identify a suitable set of instruments, finding that higher relevance is likely to be achieved at the expense of lower exogeneity. Nevertheless, several authors (e.g. Bound et al., 1995; Staiger and Stock, 1997) urge that R^2 s and/or F-statistics from the first-stage regression of instruments against endogenous explanatory variables be routinely reported, as a basis for judging instrument relevance. Those indicators were not reported for at least three empirical studies reviewed in this section.

Having a poor \hat{BE} can be viewed as an instance of measurement error in the original variable (the true BE), which is known to result in coefficient estimates for that variable that are inconsistent and biased toward zero, and coefficient estimates for the other variables in the equation that are also biased (Greene, 1997).

- (1) The asymptotic properties of IV estimators no longer hold, even for very large samples. The resulting coefficient estimators can be extremely biased and statistically inconsistent, i.e. differing considerably from the true value with non-negligible probability (Bound et al., 1995; Hall et al., 1996; Staiger and Stock, 1997).
- (2) In small samples, the coefficients in equation (6) are biased in the same direction as the ordinary least squares (OLS) coefficients, which, as mentioned in the introduction to this section, are themselves biased and inconsistent estimators of the true values. The weaker the instruments, the more closely the bias of IV estimation approaches that of OLS (Bound et al., 1995).
- (3) When the correlation between BE and ε is high, it is possible for \hat{BE} to appear to be strongly significant in equation (6), even when the true effect of BE is zero, and “[t]hus it is in the cases where least squares is a poor estimator that instrumental variables with a poor instrument will be even worse” (Nelson and Startz, 1990, p. S125; also see Hall et al., 1996).

Finally, independently of the weak instruments problem, special account needs to be taken of the sampling variance in \hat{BE} , or else incorrect statistical inferences on the significance of its coefficient in the TB model may result. The corrections needed are especially tedious when the TB variable is discrete (Bhat and Guo, 2007).

3.4 Sample selection models

Individuals who are observed to be living in a traditional neighborhood are often different from those in a suburban neighborhood. Selection bias is a major concern of observational studies. The basic idea behind sample selection models is to explicitly model the prior selection into (or participation in) different discrete states (residential location types here), and model the outcome of interest (TB) as conditional on that prior selection. Using the notation of our context, probably the most common form of the sample selection model is

$$RC^* = f_R(BE, X, Z) + \varepsilon_R$$

$$RC = 1 \text{ if and only if } RC^* \geq 0; \text{ otherwise } RC = 0$$

$$TB = f_T(BE, X, Y) + \varepsilon_T,$$

where all variables are as defined previously, and TB is observed if and only if residential choice (RC) = 1 (i.e. $RC^* \geq 0$). (More specifically, we could define the outcome equation in terms of a latent variable TB^* , which is equal to the observed variable TB if $RC^* \geq 0$). ε_R and ε_T are allowed to be correlated, and their correlation is generally simultaneously estimated together with all the other parameters of the joint system. This particular form of the model applies to a situation such as the demand for working hours, where the chosen number of hours (corresponding to our TB) is observed only if the individual is actually working (corresponding to our $RC = 1$; Greene, 1997). If that selection bias in the sample of workers is ignored and only the second equation is estimated using ordinary least squares regression, the resulting coefficients will be inconsistent and inefficient. Greene (1997) shows that doing so leads to a form of omitted variables bias, where the omitted variable is one that corrects for the sample selectivity.

That typical form of the selectivity model is not quite appropriate in our context, however: although RC is often treated as binary in the literature (e.g., representing a stereotypical “urban” (U) or “suburban” (S) neighborhood), we observe TB in either case, not only if $RC = 1$. A more general form of the sample

selection model, referred to as a switching regression model with endogenous switching, is needed (Lee, 1983; Maddala, 1983; Heckman, 1990)³:

$$RC^* = f_R(BE, X, Z) + \varepsilon_R,$$

$RC = 1$ (urban neighborhood chosen) if and only if $RC^* \geq 0$; otherwise $RC = 0$ (suburban neighborhood chosen),

$$TB_U = f_U(BE, X, Y) + \varepsilon_U, \text{ and}$$

$$TB_S = f_S(BE, X, Y) + \varepsilon_S, \tag{7}$$

where TB_U is observed if $RC^* \geq 0$, and TB_S is observed if $RC^* < 0$.⁴ (More specifically, we could define the outcome equations in terms of latent variables TB^*_U and TB^*_S , with an observed variable TB equal to TB^*_U if $RC^* \geq 0$ and TB^*_S if $RC^* < 0$, respectively).

Where is AT in this model? If observed, it is represented by X, Y, and/or Z, depending on its expected relationships with RC and TB (generally expected to influence both, i.e. to be represented by X). Most commonly, however, AT is unobserved, and the implicit assumption is that the influence of AT on BE is controlled for by the presence of BE in the RC^* equation. In reality, since the measurement of BE will not be perfect and its relationship to RC^* will not be perfectly captured, some (perhaps much) influence of AT will remain in ε_R and possibly be correlated with counterparts in ε_U and ε_S . The model formulation allows for this eventuality.

Greenwald (2003) extended the participation equation into a multinomial choice model following Lee (1983; also see Bourguignon et al., 2007 for additional information on estimating this type of selection model). In particular, he classified 4,235 respondents from the 1994 Household Activity and Travel Behavior Survey in Portland, Oregon, into six types of residential conditions based on residential tenure (own, rent) and three levels of the PEF score. A multinomial logit model was developed to predict individuals' residential choice (pseudo- R^2 was 0.33), with sociodemographics and some variables derived from census data being the explanatory variables. Then, he inserted the predicted probability of the *observed* residential choice into eight separate models of "substitution rates", intended to represent travel time choices for a "typical" trip in each category. Specifically, the dependent variables were ratios of median walking trip times to driving trip times and transit trip times to driving trip times, for consumption, communication, socialization, and all purposes (where the median was taken over all trips in the given category, made by that individual, and estimated for each mode using network-based travel times for the reported trips). He found that the predicted probability significantly (negatively) influenced the ratio of median transit time to median driving time, for consumption and socialization purposes, as well as in the model for all trips. After accounting for the influence of residential self-selection in this way, he also found that some built environment variables were significant in all models.

However, Greenwald's specification departs from the classic formulation of a selectivity model. First, in a two-stage selectivity model, the new explanatory variable in the outcome equation is not the predicted participation probability (which in essence would make it an IV model similar to that of Khattak and Rodriguez, 2005) but the inverse Mills ratio (IMR: for simplicity, if the participation equation is a binary probit model, the $IMR = \phi(\beta' X) / \Phi(\beta' X)$, where ϕ and Φ are the PDF and CDF of a standard normal distribution, respectively) derived from the participation equation (Cameron and Trivedi, 2005; Lee, 1983). Second, in a multinomial logit-OLS model, the number of outcome equations is not one but

³ See Section 3.3 for the discussion of Vance and Hedel (2007), who apply a related technique, the two-part model (2PM), to control for the endogeneity of urban form measures in a model of car distance driven. See Dow and Norton (2003) for a useful discussion of the differences between the 2PM and the classic Heckman selectivity model (Heckit).

⁴ Train (1986, Ch. 5) has a useful treatment of this model when the RC variable is multinomial rather than binary.

depends on the number of alternatives in the multinomial logit model (Lee, 1983). It is problematic to interpret the impact of the estimated probability of the single *chosen* residential alternative (which could be any of the six types, whether more or less pedestrian-oriented) on the travel time ratio: interpreted directly, the model indicates that the better residential choice is predicted (i.e. the higher the predicted probability of the chosen residential location type) – no matter what that choice may be – the higher median driving trip time tends to be, compared to median transit trip time. Therefore, the ability of this model to correct for selectivity bias is unclear.

Zhou and Kockelman (2008) pioneered the application of the sample selection model to determine the relative contributions of the built environment and self-selection to travel behavior. Specifically, they classified 1,903 households in the 1998-1999 Austin Travel Survey into two groups: CBD and urban residents, and rural and suburban residents. They chose rural and suburban residents as a treatment group and the others as a control group. Using a sample selection model, they first modeled the prior residential choice (pseudo- R^2 was 0.07) and then inserted a derived lambda (which is the IMR for the treatment group and $\phi(\beta' X)/[1 - \Phi(\beta' X)]$ for the control group) into the two equations for VMT of the treatment and control groups. They calculated and compared the average treatment effect (ATE: the average increase in VMT of moving a *randomly-selected person* from an urban neighborhood to a suburban one, or the *true* influence of the built environment) and the effect of treatment on the treated (TT: the average increase in VMT of having moved a *randomly-selected suburban resident* from an urban neighborhood to a suburban one, or the *total* influence of the built environment) (Heckman et al., 2001). They found that self-selection (the difference between TT and ATE) accounted for 10% to 42% of the total influence of the built environment on travel behavior, depending on model specifications. In their outcome equations, however, they included population density and job density besides demographics and lambda. The appropriateness of including these built environment variables is debatable. The type of the residential neighborhood is a result of residential self-selection. So are built environment elements associated with the neighborhood. For example, those who are observed to be living in high density (or accessibility, land use mix) areas may differ from those in low density (or accessibility, land use mix) areas. Therefore, these variables could still be correlated with unobserved characteristics influencing travel behavior, in violation of the model's assumption. Further, although there appears to be little discussion in the literature on the efficacy of the participation equation (i.e. its "relevance", to borrow the IV terminology), the low pseudo- R^2 of that equation in this application is cause for concern (Ed Vytlačil, personal communication with Cao, May 30, 2008), analogous to the problem of weak instruments in the IV model.

Neighborhood type is a coarse measurement of the built environment because two examples of the same type of neighborhood can differ on many dimensions such as density and land use mix. It is specific built environment elements rather than neighborhood type that can guide planners on how to improve the environment. Therefore, although Heckman's sample selection model is valuable for separating various types of treatment effects (Heckman et al., 2001), its model specification precludes the inclusion of other built environment variables in travel behavior equations because of their endogenous nature as discussed above. This challenge is not insurmountable. For example, as Greenwald (2003) did, we can classify residents into groups along a few dimensions of the built environment and then apply a multinomial logit model for residential choice (Lee, 1983). In this case, the model involves multiple treatments, and generalized point estimates for those treatment effects are needed.

The fundamental unity between sample selection models and those using the IV method has been formally explored, e.g. by Heckman and Vytlačil (2005). At a superficial level, both involve multiple equations: one or more "outcome" equations (for TB, in our context) and one or more equations modeling the "troublesome" (endogenous explanatory) variable (RC or BE, in our case), referred to in the sample selection context as the "participation" or "selection" equation(s). The latter equation is typically a discrete choice model in the sample selection context, but as seen for the Khattak and Rodriguez (2005)

study above, the same can be true for an IV model (although they are more often linear regressions on continuous dependent variables). However, whereas in an IV model the instrumental predictors of BE should be variables that are not expected to have a direct impact on TB (e.g. Bound et al., 1995), in sample selection models it is not only permissible but customary (though not essential) for the participation and outcome equations to share some explanatory variables (X and BE in equation (7)), and permissible for the observed explanatory variables in the participation equation to be correlated with the unobserved variables in the outcome equation. It is generally assumed, though, that a sample selection model contains at least one explanatory variable (Z above) that influences participation but has no direct effect on the outcome; such variables are instrumental variables for participation. When all the explanatory variables in the participation equation fit that description, the sample selection model is essentially an IV model.⁵

In our context the two approaches differ in the forms of the structure they place on the relationships among the key endogenous variables. The IV approach explicitly incorporates BE into the TB equation, whereas the sample selection approach need not – the effect of the BE could be completely captured by the RC and RC^* variables (although that would represent the special case in which the BE has no separate influence on TB , beyond its influence on RC). On the other hand, the concept of discrete observed participation (RC) controlling the entire outcome (TB) equation is integral to the sample selection approach, whereas in the IV approach, BE (or RC) is just one of many potential explanatory variables that enters the TB equation in the usual linear compensatory fashion.

In practical terms, the IV technique is usually applied in two stages, as outlined in Section 3.3. By contrast, modern applications of the sample selection approach often estimate the parameters of all equations simultaneously, yielding an increase in efficiency and enabling explicit estimation of the correlations of unobserved variables across equations.

3.5 Propensity score

Propensity score is highly recommended in social epidemiology (Oakes and Johnson, 2006). In a non-randomized observational study, treatment is a result of self-selection. A direct comparison of outcomes (travel behavior) between those in treatment and control groups (traditional vs. suburban neighborhoods) tends to produce biased treatment effects because individuals in these two groups may have systematic differences on some characteristics. Alternatively, we can mimic a randomized experiment using a propensity score. The propensity score is the conditional probability that an individual receives a treatment given a set of observed covariates (Rosenbaum and Rubin, 1983). Simply put, the propensity score is the predicted probability of a binary choice model in which the dependent variable is the decision to receive a treatment and personal characteristics are independent variables. In practice, the propensity score has been used to address the nonrandom assignment of treatment through such applications as stratification, matching, and regression (covariance) adjustment (D’Agostino, 1998; Rosenbaum and Rubin, 1983).

In *regression adjustment*, the propensity score (the predicted probability) enters travel behavior equations as an additional independent variable. Therefore, Khattak and Rodriguez (2005) is more or less an application of propensity score regression adjustment. In *propensity score matching*, once we know the propensity score of a random individual in the treatment group, we can match this individual with a person who has the same propensity score (or within a predefined range) in the control group. Because both individuals have the same propensity score, this matching reduces the bias and hence produces balances in those personal characteristics. This approach mimics an experimental design in which two exchangeable individuals are randomly assigned to the treatment group and the control group. Once the

⁵ At the other extreme, for the first formulation of the selection model above, if the selection and participation equations are identically specified, the Tobit model results (e.g., Sigelman and Zeng, 1999).

matching is complete, the ATE is the difference in the mean outcomes between those in the treatment and control groups. *Propensity score stratification* classifies respondents into several groups based on the predicted probability, and then compares the mean outcomes in each stratum. The ATE for the whole sample is an average of the differences in the mean outcomes. This approach mimics an experimental design for each stratum (Rosenbaum and Rubin, 1983; 1984).

The propensity score method is different from the statistical control method (Section 3.2). Conceptually, the propensity score method controls for the observed characteristics that affect *whether an individual is assigned to a treatment group or a control group*. The attention is directed to the imbalance in the values of covariates between treatment and control groups. The statistical control method *identifies the determinants of travel behavior* through incorporating them directly into the behavior equation, so that we can account for all differences between treatment and control groups that affect the behavior. The attention is directed to the behavioral outcome (Oakes and Johnson, 2006; Winship and Morgan, 1999). In reality, however, when the propensity score is used as a regressor in the outcome equation, it is acting as one type of statistical control, namely a composite of the variables differentiating the treatment and control conditions. Empirically, the model used to estimate a propensity score is a prediction model so it is not necessary to evaluate multicollinearity and statistical significance of explanatory variables; interaction and polynomial terms are always encouraged for propensity score estimation (Oakes and Johnson, 2006). However, multicollinearity and statistical significance are important for a model aiming to explain travel behavior in the statistical control approach.

The sample selection model for a binary endogenous variable is essentially a generalized propensity score approach, although the application of the former is earlier than that of the latter (Winship and Morgan, 1999). The difference between the two approaches is that the sample selection model requires a strong normality assumption and inserts a lambda into the behavior equation, but the application of a propensity score as a regressor inserts the estimated propensity score (probability) into the behavior equation (Winship and Morgan, 1999).

The propensity score approach has recently been applied in the field of travel behavior. Boer et al. (2007) explored the influence of built environment elements on walking choice (binary), using the 1995 National Personal Transportation Survey (NPTS). Their logistic regression propensity model included individual and household traits as regressors. The goodness of fit of the model was not provided. Without the propensity score matching, they found that land use mix, density, and parking pressure (defined as number of residents per foot of parkable street length) were significantly associated with walking choice. With the matching, many previously significant relationships became insignificant although a few remained significant. Therefore, self-selection played an important role in walking choice.

Using 1,553 residents in Northern California, Cao (2008) applied propensity score stratification to estimate the true effect of neighborhood type on travel behavior. He developed a binary probit model for residential choice (traditional and suburban neighborhoods), with demographics, residential preferences, and travel attitudes as explanatory variables. Then he classified these residents into quintiles based on the propensity score and calculated the ATEs of neighborhood type. The results showed that, on average, the true effect of neighborhood type on driving distance is 18.0 miles per week, which accounts for 12% of individuals' overall vehicle miles driven. The influences of neighborhood type are likely to be overstated by 29% for driving distance, by 64% for utilitarian walking and 16% for recreational walking if residential self-selection is not controlled for.

The propensity score method has many limitations (Oakes and Johnson, 2006). Because the computation of propensity score relies on observed characteristics, the method does not consider any bias due to unobserved characteristics. Thus, if unmeasured attitudes are a source of self-selection, this approach cannot compensate for that. Also, the approach can be employed only if there is a large amount of

overlap in the scores for those in the treatment and control groups. Otherwise, for matching, one may not be able to find many matched pairs and most information in the data will be lost. Excluding the unmatched cases in the treatment group is likely to discard important information about the types of people selecting treatment, and the outcomes for those types, biasing the sample (and the resulting coefficient estimates) in a different way. For stratification, few overlaps in the propensity score may yield only few observations in either treatment or control group and hence make comparison impossible. Further, the matching may be sensitive to the predefined caliper-width (i.e. range for determining whether or not two observations are matched). If the width is too small, one may not be able to find a match for many individuals in the treatment group; if the width is too large, the ability to reduce bias is limited. Finally, the propensity score approach is commonly used for a binary (yes-no) treatment variable. There is some extension to multiple treatment conditions, but the usefulness of the extension is unclear (Oakes and Johnson, 2006).

3.6 Other joint models

The quantitative approaches discussed thus far have progressed from single-equation models of TB that explicitly control for attitudes, to two-step instrumental variables models where the BE variable is purged of its correlation with attitudes before incorporating it into the TB equation, to multi-equation models where selection into discrete conditions (residential neighborhood types) is modeled jointly with an outcome variable (TB) having a specification that will differ by condition. In terms of application to the problem at hand, three other types of models that simultaneously account for multiple endogenous choices appear in the literature: joint discrete choice models involving nominal and/or ordinal endogenous variables, structural equations models involving continuous endogenous variables, and mutually-dependent discrete choice models. These models have recently become very popular, as shown by the number of studies published since 2006 alone. Although in principle a combination of these types of simultaneous models is possible – that is, a system of structural equations with discrete (binary or ordinal) endogenous variables (see, e.g., Lee, 1981; Lewbel, 2004; Louviere et al., 2005; Muthén, 1983) – applications involving such systems are still relatively rare and we are not aware of any in the present context.

3.6.1 Joint discrete choice models

In *joint discrete choice models*, the observed endogenous variables measuring residential choice (RC) and travel behavior (TB) are both discrete, whether nominal or ordinal, and the joint probability of an (RC, TB) bundle being chosen is modeled. Such models have been recognized as an approach to dealing with residential self-selection for several decades. For example, Horowitz (1986) reported on a 1976 multinomial logit estimation of joint probabilities of residential location (census tract) and commute mode choices (auto or bus) in Washington, DC using data collected in 1968. In justifying the approach, he commented (pp. 207-208) that preferences for travel by certain modes could affect one’s choice of residential location as well as the converse, and “[t]hus, there is not a clear direction of causality from one choice to the other. Causality may run in both directions simultaneously, thereby making the choices interdependent.” He explained that a model of commute mode choice *conditional on residential location* could account for short-term impacts of (for example) policy changes on mode choice, whereas modeling the two choices *jointly* could account for long-term impacts of the policy on residential location, as well as on mode choice.⁶

⁶ To see the difference, consider the following simplified and hypothetical example. Suppose there are only two kinds of residential neighborhoods, RC = urban (“urb”) or suburban (“sub”), and that travel behavior consists of choosing TB = car or bus. Suppose that for an auto-oriented person living in an urban neighborhood the probability of choosing car is 0.6, while for such a person living in a suburban neighborhood the probability is 0.9. Suppose that at time 0, the probabilities that such a person will choose to live in an urban or suburban area are 0.2 and 0.8, respectively. Then we have that the total probability such a person will choose car is $\Pr_0[\text{TB} = \text{car}] = \Pr_0[\text{TB} = \text{car} \mid \text{RC} = \text{urb}] \Pr_0[\text{RC} = \text{urb}] + \Pr_0[\text{TB} = \text{car} \mid \text{RC} = \text{sub}] \Pr_0[\text{RC} = \text{sub}]$

This category can be further subdivided into two: “sequential” and simultaneous models. The “*sequential*” approach is represented by the multidimensional nested logit model (Ben-Akiva and Lerman, 1985), where both choices are treated as nominal, and in which one choice (most naturally, TB) is conditioned on the other (RC) so that the joint probability of an (RC, TB) bundle being chosen is modeled as $\Pr[\text{RC}] \Pr[\text{TB}|\text{RC}]$. The latent utility for a generic (r, t) combination is formulated as an additive function of observed variables common to both the residential choice and travel behavior alternatives (X_{rt}) and unique to each type of choice (Y_t and Z_r , respectively), unobserved variables common to both (ε_{rt}), and unobserved variables unique to the “upper” choice (ε_r):

$$U(r, t) = f_5(X_{rt}, Y_t, Z_r) + \varepsilon_r + \varepsilon_{rt}. \quad (8)$$

The conditional probability of the lower choice, $\Pr[\text{TB} = t \mid \text{RC} = r]$, will be a function of X_{rt} and Y_t . The marginal probability of the upper choice, $\Pr[\text{RC} = r]$, will be a function of the expected maximum utility (referred to as the “inclusive value”) across all alternatives available for the lower choice (conditioned on the upper choice), as well as Z_r . It can be shown that the correlation θ of the error terms for the utility functions of choices within the same nest (i.e. the utilities of all (r, TB) choices for a specific residential choice r) is:

$$\theta = \frac{\text{Var}(\varepsilon_r)}{\text{Var}(\varepsilon_r) + \text{Var}(\varepsilon_{rt})}.$$

$$= (0.6 \times 0.2) + (0.9 \times 0.8) = 0.84.$$

Now suppose that a number of changes are made to the urban area to increase the cost of driving and make transit more attractive, so that this decreases the probability that even the auto-oriented urban dweller will use the car, say to 0.5 (from 0.6). The short run (SR) computation of the total probability the person will choose car assumes the residential location remains fixed, and reflects a forecast that would be made if self-selection were not taken into account:

$$\begin{aligned} \Pr_{\text{SR}}[\text{TB} = \text{car}] &= \Pr_1[\text{TB} = \text{car} \mid \text{RC} = \text{urb}] \Pr_0[\text{RC} = \text{urb}] + \Pr_1[\text{TB} = \text{car} \mid \text{RC} = \text{sub}] \Pr_0[\text{RC} = \text{sub}] \\ &= (0.5 \times 0.2) + (0.9 \times 0.8) = 0.82, \text{ lower than before, as expected.} \end{aligned}$$

But in the long run (LR), the increased densification and higher cost of driving will increase the probability that the auto-oriented person will choose a suburban area (say to 0.9, from 0.8) over an urban one, so that the long run total probability of choosing car is:

$$\begin{aligned} \Pr_{\text{LR}}[\text{TB} = \text{car}] &= \Pr_1[\text{TB} = \text{car} \mid \text{RC} = \text{urb}] \Pr_1[\text{RC} = \text{urb}] + \Pr_1[\text{TB} = \text{car} \mid \text{RC} = \text{sub}] \Pr_1[\text{RC} = \text{sub}] \\ &= (0.5 \times 0.1) + (0.9 \times 0.9) = 0.86, \end{aligned}$$

actually *higher* than before the policy changes intended to make driving less attractive. Thus might policy changes executed in only one part of a region have unintended consequences, if there is still freedom to “vote with one’s feet” (or car, as the case may be) and move elsewhere.

The alert reader will of course suggest making the policy changes to the *suburban* area rather than the *urban* area, since if the auto-lover doesn’t like it, then where will she go? Well in real life (aside from the political difficulty, in a democracy, of adopting such solutions in areas where voters are self-selected to prefer otherwise), there is likely to be an even more suburban alternative that will be available, at least for a long time into the future. The main point here, however, is not so much the specifics of this particular example – the example focused on what would happen to a particular type of person (auto-oriented), but hopefully there will be enough transit- and walking-oriented people who become *more* likely to move into the urban area, to more than compensate for the auto-oriented people who move out. Rather, the point is to illustrate how a model that *jointly* accounts for changes in residential location as well as travel behavior provides a forecast of the impacts of transportation and land use policies that is (a) different from, and (b) more realistic than, a model that is implicitly conditioned on residential location and therefore fails to take the self-selection factor into account.

This quantity measures the proportion of total variation in the unobserved portion of utility that is due to unmeasured variables (such as attitudes) common to all (r, TB) alternatives (ε_r) as opposed to variables unique to each individual (r, t) combination (ε_{rt}). As such, it could be viewed as an indicator of the extent to which selection into a particular residential neighborhood has not been controlled for by observed variables: the more completely variables that are related to residential choice in a systematic way are accounted for (observed), the more the variation in *unobserved* utility will be due to idiosyncratic features of specific (RC, TB) combinations, which vary randomly from one combination to the next.

It is important to realize that although such a model *can* represent a temporal sequence of choices, it need not do so (Sobel, 1980). Mathematically, the nested logit model simply represents a particular structure for the correlations of unobserved variables across sets of alternatives; the choices themselves theoretically could take place in any order or simultaneously. However, although nested logit models do not *impose* a sequential structure on multiple decisions, they can certainly *reflect* one when it exists (in other words, sequentially dependent decisions are a sufficient, though not necessary, condition for nested logit to be a potentially appropriate model structure). The present context is one such natural application, since residential choice has long and widely been held (e.g. Salomon and Ben-Akiva, 1983) to be a longer-term choice which is antecedent to short-term choices related to individual trips. But it must be emphasized that finding such a structure to fit the data well cannot be taken as *confirmation* of a sequential decision process, only as being *consistent* with it. Further, analysts should not let a presumed temporal sequence of decisions blind them to alternative possibilities such as those shown in Figure 1.

Although a number of nested logit models incorporating residential location and travel behavior have been developed (e.g. Abraham and Hunt, 1997), at least two have employed the technique explicitly to account for residential self-selection. Cervero (2007) developed a two-level nested logit model, with the upper level indicating the binary choice of residential location (whether or not to live within half a mile of a rail station) and the lower level representing the binary choice of commute mode (rail or auto). Using 11,369 workers in the 2000 San Francisco Bay Area Travel Survey, he calibrated the model, and then compared the average odds of choosing rail over auto for those living near rail stations ($0.1547 / 0.8453 = 0.1830$) to the average odds for those living farther away ($0.1144 / 0.8856 = 0.1292$). Cervero concluded that (p. 2082) “the odds of rail commuting are 41.6% [the percentage by which 0.1830 exceeds 0.1292] greater if one lives near versus away from transit, all else being equal... This suggests that around 40% of the higher rail commuting shares among Bay Area workers living near transit is accounted for by self-selection.” However, we have two main concerns about this result. The first is that the odds ratio is not necessarily a percentage of a whole (i.e. mathematically, it could exceed 100%), and we do not see a logical translation from the factor by which the odds change, to a percent of commuting shares accounted for by residential self-selection. The second concern is that it is not clear how the *conditional* probabilities (e.g. $\Pr[\text{TB} = \text{rail} \mid \text{RC} = \text{close to transit}]$) account for prior self-selection, as opposed to the direct influence of the built environment (i.e. the fact that rail by definition is more convenient for those living closer to stations) once a residential selection has been made.

Using travel diary data from the Regional Travel – Household Interview Survey, Salon (2006) estimated a three-tiered nested logit model of residential choice (census tract: the chosen tract plus 10 randomly-selected alternatives), auto ownership (AO: 0, 1, or 2+ cars), and walking level (WL: zero = no trips that were walk-only, low = 1-49% of trips walk-only, and high = 50% or more trips walk-only) for 4,382 residents of New York City. Given the available variables, she used population density as an indicator of neighborhood walkability. Using the full joint model, she then computed various elasticities of WL with respect to population density (BE, for the sake of argument). She suggested that the effect of locational self-selection can be quantified by taking the difference between unconditional elasticities of WL and those computed to be conditional on RC. That is, the self-selection effect is the difference between (1) the elasticity of walking level with respect to population density calculated from the marginal $\Pr[\text{WL}]$

(obtained from the unconditional probabilities $\Pr[\text{WL}, \text{RC}, \text{AO}]$ by $\Pr[\text{WL}] = \sum_a \sum_r \Pr[\text{WL}, \text{RC}, \text{AO}]$), and (2) that calculated from the conditional $\Pr[\text{WL} | \text{RC}]$ (obtained from $\Pr[\text{WL}, \text{AO} | \text{RC}]$ by $\Pr[\text{WL} | \text{RC}] = \sum_a \Pr[\text{WL}, \text{AO} | \text{RC}]$). With the application of this approach, Salon concluded that self-selection accounted for one-third to one-half of the effect of a change in population density (BE) on walking level (TB) in most areas of New York City.

In the *simultaneous* joint discrete choice model, latent utilities for each choice, RC^* and TB^* (where, in the two applications to date, RC is nominal and TB is respectively ordinal or nominal), are formulated in separate equations, with the probability of a particular (RC, TB) bundle being estimated jointly. The separate utility equations may have overlapping sets of explanatory variables, but (together with the other joint models discussed so far) do not include one endogenous variable directly in the equation for the other. Bhat and Guo (2007) pioneered the theoretical development and empirical application of such a joint structure modeling discrete residential choice and ordinal car ownership, parameterizing the error terms as follows:

$$\begin{aligned} \text{RC}^* &= b(\text{BE}, Z, X) + u \text{BE} \pm w \text{BE} + \zeta \\ \text{TB}^* &= t(\text{BE}, Y, X) + v \text{BE} + w \text{BE} + \delta, \end{aligned} \tag{9}$$

where u and v are unobserved (individual-specific) factors (such as attitudes) impacting households' sensitivity to built environment traits in residential choice alone and travel choice alone, respectively; w stands for unobserved individual factors impacting both residential and travel choices; and ζ and δ are idiosyncratic terms. By including the common error term $w \text{BE}$, Bhat and Guo's model simultaneously corrects for the endogeneity of the built environment. Using data from 2,954 Alameda County households in the 2000 San Francisco Bay Area Travel Survey, they calibrated this joint mixed multinomial logit-ordered response model. In their operationalization, RC is measured as a discrete indicator of one of 233 transport analysis zones, BE variables include measures for zonal density, zonal land-use structure, regional accessibility, local transportation network, and commute-related variables, and TB is the ordinal measure of number of vehicles owned by the household. Their results showed that the built environment has a true influence on auto ownership, and the lack of a significant common error term (more precisely, estimates of the variance of w that did not significantly differ from zero) failed to support the speculation that attitude-based residential self-selection influences auto ownership choice in this application.

Pinjari et al. (2007) extended Bhat and Guo's approach to incorporate a multinomial mode choice representation of TB. Using 1,878 commuters from the same survey, they developed a joint model for residential location choice and commute mode choice. They found that the effect of self-selection results from both observed variables and unobserved factors, and the built environment has an independent influence on mode choice beyond the influence of self-selection. Potential limitations of this extension are similar to those of Bhat and Guo (2007): if the correlations of error terms across equations are due to unobserved variables such as attitudes rather than to BE and other observed variables, inclusion of terms analogous to $w \text{BE}$ may not solve the self-selection problem. Further, as elaborated below, these joint discrete choice models do not represent direct causal relationships between the endogenous variables of the system, only correlated error terms.

3.6.2 Structural equations models

The second category of joint models is *structural equations models*. By contrast to the joint discrete choice models of the previous subsection, here the endogenous variables are typically continuous, and they are usually modeled as directly influencing other endogenous variables. Recognizing that AT influences both BE and TB, and therefore including it in a single-equation model for TB, as in equation (5), constitutes a useful improvement in the realism of a model of TB. In fact, however, the influence

between attitudes and behavior is probably not entirely unidirectional, as Figure 1 illustrates. It is quite possible that over time, both the BE and TB may affect AT as well, and AT and TB could affect BE (bringing about a residential relocation). There is a sizable literature in transportation (and other fields) on the mutual causality between attitudes and behavior, with ample evidence for impacts in both directions (e.g., Tardiff, 1977; Golob, 2001). Thus, improving the realism of the model even further suggests the need for multiple interrelated equations, reflecting the multiple likely directions of causality. Specifically, one could postulate the following Structural Equations Model (SEM):

$$\begin{aligned}
 TB &= t(AT, BE, W, X, Y, Z) + \omega_1 \\
 BE &= b(AT, TB, W, X, U, V) + \omega_2 \\
 AT &= a(TB, BE, W, Y, U, S) + \omega_3,
 \end{aligned}
 \tag{10}$$

where W = observed variables common to all three equations, X = observed variables influencing both TB and BE but not AT; similarly for Y and U; Z, V, and S are observed variables whose influences are unique to TB, BE, and AT respectively; and the ω s represent the net impacts of the unobserved variables relevant to each left-hand side.

Using 1993 data on 515 individuals in the San Francisco Bay Area, Bagley and Mokhtarian (2002) employed SEM to investigate the relationships among the built environment, travel behavior, and attitudes. This study is also the first application of covariance structural analysis in exploration of the relationships between the built environment and travel behavior. In this study, nine endogenous variables were incorporated into the structural model specification: two continuous residential type measures, three measures of travel demand, three measures of attitudes, and one measure of job location. The exogenous variables consisted of sociodemographic characteristics, lifestyle factor scores, and other measurements of attitudes. They found that with respect to direct and total effects, attitudinal and lifestyle variables had the greatest impact on travel demand among all explanatory variables, while residential location type had little separate influence on travel behavior. These results lend strong support to the speculation that the observed relationships between the built environment and travel behavior are not direct causal links, but are primarily attributed to interactions of these measures with other variables.

Using a sample of 1,217 workers collected from Northern California in 2003, Circella et al. (2007) explored the connections among six groups of variables: sociodemographics, travel attitudes, land use preferences, neighborhood characteristics, auto ownership, and vehicle miles driven. They treated the former two groups as exogenous variables and the remaining as endogenous variables. They found that travel attitudes and land use preferences were associated with neighborhood characteristics and driving behavior, and neighborhood characteristics had an impact on driving behavior. Therefore, they concluded the concurrent influences of the built environment and residential self-selection on travel behavior.

Scheiner and Holz-Rau (2007) applied SEM to a sample of 2,691 residents in the region of Cologne, German, in 2002 and 2003. In this study, they investigated the interactions among life situation (sociodemographics), lifestyle factors, residential location attitudes, urban form, and travel behavior, with the latter four groups being endogenous variables. Travel behavior variables include mode use (the share of trips by a specific mode) and vehicle kilometers travelled. Quality of public transportation, density of supply (retail, service, and leisure opportunities), and density and mixed use were chosen as the measurements of urban form. They developed eight SEMs with different model specifications. They found that individuals who prefer high quality of transit, good access to retail and service, and urban life were more likely to live in urban areas with high density and mixed use, and those living in such neighborhoods tended to drive less and use alternative modes more. So both the built environment and self-selection have influences on travel behavior.

3.6.3 Mutually-dependent discrete choice models

The joint discrete choice models of Section 3.6.1 do not allow direct causality between their endogenous variables. By contrast, as indicated above, the SEMs of Section 3.6.2 are built around the concept of direct (potentially mutual) causality among endogenous variables, but in most travel behavior applications to date, those variables are continuous. One recent application, however, conceptually blends both these approaches, jointly modeling discrete endogenous variables as mutually dependent.

Also applying the concept of latent utility, Chen et al. (2008) constructed two simultaneous equations, in which car ownership level and the propensity to use a car influence each other. In this specification, car use for commute trips (a binary variable) is observed but the underlying propensity to use a car is unobserved. Thus, this latent propensity is presumed by the authors to include unobserved attitudes toward car use. If residential self-selection is at work, it is expected that the propensity will have a significant impact on car ownership level. Chen and colleagues applied a two-stage estimation method to a sample of 2,089 commuters (having cars in the household) in the New York metropolitan region. In particular, they first estimated a probit model for car use, as a function of car ownership, demographics, built environment variables, and tour complexity (pseudo- R^2 was 0.60). They then inserted the predicted probability for the propensity to use a car into an ordered probit model for car ownership. They found that car ownership level was not significant in the model for car use, but the predicted probability of car use had a significant influence on car ownership level. After controlling for the propensity to use a car, population density and transit-based job accessibility at home had an association with car ownership.

The fact that the causality appeared to go only in one direction (car use propensity influencing car ownership level) in this application suggests that, empirically, this model could be viewed as an example of several other methods discussed earlier, and accordingly subject to their limitations. Specifically, the probit model of car use could be viewed as a propensity score model, where the predicted propensity (probability) is then included as a regressor in the outcome (car ownership) equation. As such, as mentioned in Section 3.5 it is not clear how a propensity that is estimated as a function of observed characteristics can resolve an endogeneity bias caused by the correlation of observed with unobserved characteristics. The predicted car use probability is also related to, and serves a role similar to, that of the lambda term in the outcome equation of a sample selection model (although the lambda term is derived from theoretical considerations). As such, the concern we expressed with respect to the Zhou and Kockelman (2008) study in Section 3.4 similarly applies here: inclusion of built environment terms in the outcome equation may also be perpetuating an endogeneity bias.

The temporal mismatch issue is also a concern with this type of model; e.g., present car use (or even the propensity to use a car, based on current variables) would not logically be expected to influence previously-determined car ownership levels.

3.6.4 Discussion

As indicated above, in joint discrete choice models, the RC and TB choices are not directly modeled as affecting each other. A dependent relationship or *correlation* of unobserved factors influencing those choices can be ascertained through statistical tests of certain parameters in the formulation, and in the Bhat and Guo formulation, the correlation is even modeled as a function of specific observed explanatory variables common to both equations (in this respect, it goes one step beyond an approach such as seemingly-unrelated regressions for continuous dependent variables, or a sample selection model where the outcome variable as well as the participation variable is discrete, where error terms are allowed to be correlated across equations but are not parameterized). Although their initial application did not have observed attitudinal variables, the inclusion of such variables in future applications could provide additional insight into the sources of the relationship between the two choices: it is possible that the wBE terms in their system were insignificant because the correlation of the error terms for the two choices was

due to unmeasured variables such as attitudes toward walking and/or driving rather than the BE variables that *were* measured.

In general, then, one limitation of this approach is that unobserved portions of the RC and TB equations are assumed to be correlated only through their relationship to observed variables (though not necessarily BE variables alone, as was assumed in their initial application for simplicity only); the remaining error terms in equation (9), ζ and δ , are assumed to be uncorrelated. Thus, if the available observed variables fail to largely capture the effects of attitudes common to both choices, a key assumption of the model might not be met. It would be interesting to test the extent to which this assumption holds, in a future study which obtains attitudinal measures as well as typical “objective” measures of the BE and sociodemographic traits, and then examines the extent to which parameterizing the error terms with only the objective measures serves to capture the effects of attitudes common to both equations. Of course, it is fair to say that *all* techniques are limited by the variables for which observations are available, and can be improved by the measurement and inclusion of additional relevant variables. Our point is precisely that although this method uses to the utmost the information embedded in commonly-available measures, we can only learn so much about behavior without moving more explanatory variables from unobserved to observed.

In any case, it may also be considered a limitation in some respects that the direction(s) of *causality* (if any) between RC and TB cannot be statistically tested. Rather, the two choices are modeled as if they occur simultaneously, potentially jointly influenced by common antecedent variables, but not by each other directly. (This is formally true even of the nested logit approach: while a sequential interpretation may be placed on the model as mentioned above, the essence of the model is the estimation of the joint probability $\text{Pr}[\text{RC}, \text{TB}]$, and the assumption of direct causality between those two choice dimensions, in either direction, is neither required nor implied. However, see, e.g., Tringides et al., 2004 for a transportation application of the recursive bivariate probit model, involving two binary endogenous variables with a unidirectional relationship between them; and Schmidt and Strauss, 1975 and Ye et al., 2007 for empirical estimations of simultaneous equation systems with two mutually-dependent discrete endogenous variables. The latter paper also has a useful discussion of the nature of causal inferences that can be made from such systems). This is a reasonable approach when two choices are made close together in time, but such a model may not reflect a situation in which the choices are in fact temporally decoupled, as RC and TB often are.

In structural equations models, by contrast, the fact that endogenous variables are modeled as directly influencing other endogenous variables provides the ability to conduct tests to ascertain which directions of causality (if any) are statistically supported by the data. This ability can enrich our insight into the behavioral processes of interest: beyond learning that RC and TB are correlated through having the BE in common, it can be valuable to determine whether that correlation is due to TB influencing the choice of the bundle of BE attributes that constitutes RC, or to the direct influence of the BE on TB, or both. The difference could be important to properly predicting the reaction to a change in the BE: potentially little reaction in the first instance (especially in the short run), and considerable reaction in the second.

However, although allowing multiple directions of causality arguably constitutes a conceptual improvement over the single-equation and joint (simultaneous) model methodologies, the use of cross-sectional data is still a practical drawback to this approach. The same temporal mismatch described in connection with model (5) of Section 3.2 may occur here. Further, properly estimating the parameters of a dynamic process using a static snapshot of data requires that the process be stable (not trending over time), and have achieved an equilibrium – conditions that not only may be unrealistic, but for which there is not a good statistical test (Kline, 2005).

Structural equations models have other limitations as well. For example, identifiability requirements may limit the specifications that can be empirically tested, and it is possible for several different specifications – representing substantively distinct behavioral processes, with different policy implications – to fit the data roughly equally well (MacCallum, 1995; although the same can be true for nested logit models as well – see, e.g., Forinash and Koppelman, 1993). Further, structural equations models are not well-suited to situations where one or more endogenous variables are multinomial, a case of considerable interest in travel behavior research (particularly, in the present context, mode choice, but destination and route choices are potentially also measures of travel behavior that could be relevant in a residential self-selection study).

3.7 Longitudinal designs

A longitudinal design can be used to control for attitudes that do not vary over time: if AT does not change across time, then $\Delta AT = 0$, and in the model

$$\Delta TB = f_6(\Delta BE, \Delta X) + \eta, \quad (11)$$

ΔBE and η ($=\Delta \varepsilon$) will be uncorrelated (if BE and ε were only correlated through AT). This formulation also controls for any other important variables that are either observed (ΔX) or remain constant (0 change) over the same time period. For these reasons, conventional wisdom holds that modeling the change in a given dependent variable is easier (produces better-fitting models, all else equal) than modeling its absolute level. The situations to which this model has been applied include residential moves, as well as changes “in place” to the built environment, e.g., the “Safe Route to Schools” (SR2S) program described below.

In contrast to the disaggregate studies that underlie the rest of this report, some studies have used a before-after design to investigate the influence of a specific change to the built environment on aggregate travel behavior. For example, Painter (1996) found that street light improvements in three urban streets and on a pedestrian footpath (previously prone to crime) in London greatly increased pedestrian street use after dark. McBeth (1999) concluded that installation of bike lanes in downtown Toronto increased bike volume. An advantage of these studies is that they concentrated on the observed changes in travel behavior of people exposed to the study areas, rather than reported changes. However, they did not employ control locations or control for other variables. The lack of controls may confound the intervention effects with other potential effects. Further, with only aggregate measures of travel demand, it cannot be determined how the changes are distributed: are the same people using the facility more, are more people using it, or both? This question can be answered with true disaggregate panel data.

In an evaluation of California SR2S projects, Boarnet et al. (2005) examined the relationship between improvements in the walking and biking infrastructure and children’s walking and bicycle travel to school, based on retrospective responses of 1,244 parents. Changes in this infrastructure (sidewalks, crossings, and traffic control) serve as a “treatment” for the children who passed the SR2S projects on their way to school (experimental group). The control group consists of those who did not pass the SR2S projects. Through paired-sample t-tests, they found that 15.4% of the 486 children who passed the SR2S projects increased their walking or bicycle travel to school, while only 4.3% of the 376 children who did not pass the projects increased their non-motorized travel. However, memory biases and social desirability biases (given that the “desirable” answer was probably especially apparent to the treatment respondents) may be concerns of this study.

Krizek (2000) examined the changes in households’ travel behavior before and after their residential relocation, using the Puget Sound Transportation Panel data. Households’ residential relocation may expose them to different built environments, serving as a “treatment”. Households’ travel behavior was

measured by a variety of variables, including trip distance, trip minutes, tour distance, tour minutes, trips per tour, and percentage of total trips taken by alternative modes. Paired-sample t-tests were conducted to examine the changes in households' travel behavior against the changes in the "Less Auto-Dependent Urban Form (LADUF)" ranking (high, medium, and low), a measurement of built environment based on an assessment of density, street pattern, and land use mix. The results showed relatively weak correlations between changes in the built environment and changes in travel behavior. He also found that more than half of his sample moved to a neighborhood whose environmental characteristics were similar to their previous neighborhood. This result suggests that households may decide to live in a neighborhood at least partly to match their travel preferences, lending additional support to residential self-selection. However, this study is vulnerable to a lack of control for other determinants of TB change.

Using the same dataset, Krizek (2003a) applied linear regression models to test whether changes in travel behavior can be attributed to changes in neighborhood accessibility, controlling for changes in sociodemographic characteristics, workplace accessibility, and regional accessibility. Travel behavior variables used in this study are VMT, person miles traveled, number of tours, and number of trips per tour. The measurements of neighborhood accessibility are dependent on a combination of density, street pattern, and land use mix. Regional accessibility is computed using a simple exponential function of travel impedance with employment as the attractiveness measure. In addition to the changes in sociodemographic characteristics and accessibility, their base values were included in the model specification to capture the effects of starting levels of these variables. The results showed that the base values of neighborhood accessibility and most sociodemographic characteristics are significant in all four models, supporting the premise that starting levels of these variables affect the changes in travel behavior. Also, the changes in neighborhood accessibility are statistically significant in all models, which suggests that when households' neighborhood accessibility changes, their travel behavior also changes, all else being equal. The author pointed out, however, that the results should be interpreted with caution, as the changes in both neighborhood accessibility and travel behavior may be the result of changes in attitudinal predispositions toward the residential environment and travel choices.

Meurs and Haaijer (2001) investigated the extent to which changes in residential environment characteristics led to changes in travel patterns, using Dutch Time Use Study data from 1990 and 1999. For the dynamic analysis, the respondents were divided into two segments: movers and non-movers. Regression analyses were conducted on both segments, in which changes in the number of trips by various modes were regressed against changes in residential environment and personal characteristics. For the people who moved, changes in residential environment characteristics influence travel behavior, and changes in employment and auto ownership as well as other sociodemographic factors greatly influence changes in auto trip frequency. For the people who did not move, the observed effects of spatial changes (which were relatively minor and incremental, such as an extra garage, the installation of traffic calming measures, and the provision of a bike path) are limited, as they expected. However, although the authors controlled for major sociodemographic determinants of mobility such as "family composition, work or activities, education, car ownership, etc." (p. 434), the nine-year span between waves means that the age distribution of respondents in the study will be biased upward by the end. Further, it appears that non-movers were primarily analyzed separately from movers, rather than explicitly used as controls.

Similarly, Cao, Handy, and Mokhtarian classified their Northern California respondents into movers and nonmovers, based on whether they moved within the last year. Unlike Krizek (2000, 2003a) and Meurs and Haaijer (2001) using objective measures, they measured changes in the built environment by taking the differences between movers' perceptions of current and previous neighborhoods, and assumed the residential environment of nonmovers to remain constant over the measurement period. Changes in travel behavior were measured using a series of general indicators of the use of different modes compared to previously, on a five-point scale ranging from "a lot less now" to "a lot more now." Residential preferences and travel attitudes were measured at only one point in time: currently. For these reasons,

they refer to this design as “quasi-longitudinal”. Handy et al. (2005, 2006) developed three ordered probit models to investigate whether changes in the built environment influence changes in driving, walking, and biking. After accounting for the influence of current attitudes and changes in sociodemographics, they found that for all respondents, changes in neighborhood characteristics consistently affect changes in these behaviors, and changes in neighborhood characteristics are the most important in explaining changes in driving and walking. Using linear regression, Cao et al. (2007b) found that for movers, a change in the perceived outdoor spaciousness of their neighborhood impacts change in auto ownership, and its influence is equivalent to that of sociodemographics.

This finding raises a concern that changes in auto ownership may be endogenous to the associations between changes in the built environment and travel behavior. Accordingly, Cao et al. (2007b) employed a structural equations modeling approach to explore the complex relationships among endogenous variables, namely changes in the built environment, auto ownership, and travel behavior. They assumed that current attitudes and changes in sociodemographics affect changes in the built environment, which in turn influence changes in auto ownership and travel behavior, and the latter two changes impact each other. They found that changes in neighborhood characteristics have a true influence on changes in driving and walking. However, the latter two studies lack a control group of non-movers due to a survey design limitation. Although the dynamic structural equations modeling approach adopted by Cao et al. (2007b) is an improvement in terms of methodology, there are still limitations in their application of it. Because it is not feasible to retrospectively measure attitudes, they have data on current attitudes only, and thus their models only control for current attitudes rather than changes in attitudes. So, they cannot rule out the competing hypothesis that an attitude change preceded and (partly) prompted the residential location change. To the extent that is true, the attitude change is confounded with the change in built environment and may account for some of the apparent effect of the built environment seen in this study.

Recently, Wells and Yang (2008) analyzed cross-sectional data (70 low-income women in Florida, Alabama, and Georgia) and longitudinal data (32 women) collected from 2003 to 2006. They employed linear regression modeling. The results of the cross-sectional analysis showed that there is no significant association between neighborhood type and post-move walking. By contrast, the longitudinal analysis demonstrated that changes in the number of cul-de-sacs and changes in the number of service jobs per resident are negatively associated with post-move walking. Further, these two variables explained 16% of the variation in walking, in addition to 44% of variation being explained by pre-move walking and demographics. For the longitudinal analysis, they took advantage of a natural experiment, a critical approach in the field of neighborhood design and physical activity (Oakes, 2004). These low-income women moved to either a neo-traditional community or a suburban neighborhood, with the help from a housing program. The women did not have an alternative choice because only one type of neighborhood was available in each region. However, some women may opt out of this program because they cannot afford a car, which may be necessary to live in a suburban neighborhood. Therefore, although this study represents the commendable use of an unusual and valuable opportunity, the extent to which it is a true natural experiment is debatable. Also, of course, a sample size larger than 32 would be desirable.

Overall, longitudinal designs can offer substantial improvement over cross-sectional designs, providing a more robust causal inference on the relationship between the built environment and travel behavior. In classical experimental design (e.g., Shadish et al., 2002), “before” measurements are taken, then participants are randomly assigned to either the experimental or the control group, the experiment is performed on the first group, and then “after” measurements are taken and changes are compared between the two groups. Such an approach is very strong on the nonspuriousness and time precedence causality requisites discussed in Section 2 (Singleton and Straits, 2005).

Longitudinal designs still have a number of limitations, however, both inherently and in the way they are likely to be applied in the present context. For example, even here, to be able to use the temporal

sequence of observations to sort out multiple potential directions of causality, it is necessary to assume that the process is stationary (meaning that the relationships of interest are stable over time), which is not always realistic. Another intrinsic issue is that it can be difficult to determine the optimal time(s) at which to take each measurement, especially since the optimal spacing between measurements may differ by individual: too short, and the changes of interest will not have had time to occur; too long, and measurement of variables and relationships will be unreliable due to memory lapses and other noise in the system (Kline, 2005).

With respect to limitations specific to the likely application of longitudinal methods in the current context, an important issue is that neither the treatment nor the assignment to experimental or control group is completely random, which was discussed in Section 2. A further practical difficulty of true longitudinal studies is that they can be more expensive and are certainly more time-consuming than cross-sectional ones. Finally, although not intrinsic limitations of the approach, applications to date have been hampered by not measuring attitudes across time (when in point of fact, it may be precisely a change in attitudes that prompted a residential relocation in the first place), and by not including feedback loops from the built environment to attitudes.

4. EXACTLY WHAT QUESTION(S) ARE WE TRYING TO ANSWER, AGAIN?

Given the preceding immersion in technical detail, it is worthwhile at this juncture to step back and ask ourselves specifically what it is we are interested in knowing! The purpose of this discussion is not to definitively enunciate the question(s) of interest, but rather to explore a variety of relevant questions and to more clearly delineate the issues involved in choosing which one(s) to address. Several pertinent questions are discussed in the subsections below.

4.1 Is there a statistically significant effect of BE on TB after self-selection has been accounted for?

This is the first and simplest way to pose the question of interest. We have indicated in the discussion of each methodological approach to this issue how that question could be answered, and indeed, if the answer is “no”, then the remaining questions of this section become moot. In point of fact, however, based on the empirical evidence to date, the answer would have to be a straightforward and resounding “yes”. Virtually every quantitative study reviewed for this work, after controlling for self-selection through one of the various ways discussed above, found a statistically significant influence of one or more built environment measures on the travel behavior variable of interest.

4.2 What is the size of the true impact of BE on TB?

On the other hand, although many academic studies tend to focus purely on the question of statistical significance, the magnitude and practical relevance of an effect is arguably at least as important (Ziliak and McCloskey, 2004). In other words, is the true influence of the built environment even worth bothering over, after we go to all the trouble to assess it properly? For example, to ascertain whether changes to the BE are a cost-effective way to change TB, given the opportunity costs of spending resources another way, it is necessary to determine the magnitude of the effect, not just whether one occurred or not. The ways to answer this question differ depending on whether the built environment is measured as continuous (BE) or discrete (RC). We treat each case in turn.

4.2.1 True marginal effects on TB for continuous-valued measures of BE

The second column of Table 2 summarizes how the true marginal effects of continuous-valued measures of BE on the expected value of TB can be ascertained for the methodologies discussed here. For the statistical control, instrumental variables, and longitudinal models, the magnitude of the true marginal impact (purging BE of the influence of AT) can easily be obtained from the appropriate coefficient of BE, \hat{BE} , or ΔBE , respectively, in the equation for TB. Note that in these cases, expressing TB as a linear

function of BE (or its variations) implies the assumption of a constant marginal effect of BE on TB. For the remaining methodologies, however, ascertaining the marginal impact is not so straightforward.

[Insert Table 2 here]

For the selection model approach, a number of definitions of the marginal impact are of potential interest (similar to the discussion in Section 4.2.2 below, for assessing the effects of the discrete-valued RC variable). The application literature is not always clear about which definition a given study uses, nor which is most appropriate to the problem at hand. It is first of all important to distinguish between (see equation system (7)) (i) effects on the expected values of the *potential* TB^*_S and TB^*_U , if those quantities could be observed for the entire population instead of only for those for whom $RC = 0$ and $RC = 1$, respectively; versus (ii) effects on the expected values of the *actual* TB_S and TB_U , where TB_S is observed if $RC = 0$ and TB_U is observed if $RC = 1$. With respect to effects on the actual TB_S and TB_U , it is further vital to distinguish between *conditional* effects – on $E[TB_S | RC = 0]$ and $E[TB_U | RC = 1]$, and *unconditional* effects – on $E[TB_S]$, $E[TB_U]$ and $E[TB]$, where $E[\bullet]$ is the expectation operator. It is also necessary, in the most general case, to account for the appearance of BE in both the participation and the outcome equations. Respectively, the effects of interest have the following interpretations in this context (see Huang et al., 1991 and Vance and Geoghegan, 2004, for applications in different contexts):

- ◆ $\partial E[TB^*_S] / \partial BE$ is the effect of increasing BE by one unit on the expected potential travel behavior of a randomly-selected person living anywhere, if that random person were to be governed by the TB_S equation. This is simply given by the coefficient of BE in the equation for TB_S (Maddala, 1983; Huang et al., 1991).
- ◆ $\partial E[TB^*_U] / \partial BE$ is the effect of increasing BE by one unit on the expected potential travel behavior of a randomly-selected person living anywhere, if that random person were to be governed by the TB_U equation. Similarly, this is given by the coefficient of BE in the equation for TB_U . The remaining effects, however, are more complex and will differ by individual.
- ◆ $\partial E[TB_S | RC = 0] / \partial BE$ is the effect of increasing BE by one unit for those living in suburban neighborhoods, on the expected actual travel behavior of those living in suburban neighborhoods. This conditional effect is not the same as would be obtained from the regression of TB against BE (and other variables) for the subpopulation of individuals living in suburban neighborhoods, i.e. not just the coefficient of BE in a stand-alone equation for TB_S . The proper formula for the conditional effect (found in Huang et al., 1991) corrects for the bias inherent in self-selection into a given type of neighborhood, by incorporating the effect of BE on RC in the participation equation.
- ◆ Similarly, $\partial E[TB_U | RC = 1] / \partial BE$ is the effect of increasing BE by one unit for those living in urban neighborhoods, on the expected actual travel behavior of those living in urban neighborhoods.
- ◆ $\partial E[TB_S] / \partial BE$ is the effect on expected actual TB of increasing BE by one unit for a randomly-selected person living anywhere, if that random person were to be governed by the TB_S equation. As Huang et al. (1991) show using the product rule of differentiation, this effect can be decomposed into (1) the change in TB_S (because of changing BE) weighted by the probability that the selected person lives in a suburb, plus (2) the change in the probability that the selected person lives in a suburb (i.e. the effect on RC of the change in BE) weighted by the expected value of TB_S given that the person lives in a suburb. In other words, a change in the built environment will alter travel behavior in two ways: by directly affecting it, and by affecting the probability of living in an urban versus suburban neighborhood, which itself affects travel behavior (by controlling which TB equation is in effect).

- ◆ $\partial E[\text{TB}_U] / \partial \text{BE}$ is the effect on expected actual TB of increasing BE by one unit for a randomly-selected person living anywhere, if that random person were to be governed by the TB_U equation; it can be decomposed in a similar way.
- ◆ $\partial E[\text{TB}] / \partial \text{BE}$ is the effect on the expected TB (regardless of whether it is observed for S or U) of increasing BE by one unit for a randomly-selected person living anywhere. $E[\text{TB}]$ is the weighted average of the conditional expected values, where the weights are the probabilities of living in the respective types of neighborhoods:

$$E[\text{TB}] = E[\text{TB}_S | \text{RC} = 0] \Pr[\text{RC} = 0] + E[\text{TB}_U | \text{RC} = 1] \Pr[\text{RC} = 1],$$

and $\partial E[\text{TB}] / \partial \text{BE}$ is just the sum of the unconditional marginal effects $\partial E[\text{TB}_S] / \partial \text{BE}$ and $\partial E[\text{TB}_U] / \partial \text{BE}$. It decomposes into the weighted average of the coefficients of BE in the two equations for TB_S and TB_U (where the weights are the respective selection probabilities), plus a term representing the correction for self-selection into neighborhood type (incorporating the influence of BE on RC).

It is interesting to realize that the complexity of marginal effects for the selection model is not purely due to potentially correlated error terms of the participation and outcome equations (although that is one factor). Rather, it is mainly due to the presence of the same *observed* variables in both types of equations. Thus, if AT is unobserved, as is often a motivation to use selection models, the error terms in the participation and outcome equations are likely to be correlated – but *even if AT is observed*, if it is present in both types of equations (as would be expected) the complexities described above will arise.

Which of these marginal effects is most appropriate for the problem at hand? In some cases it might be the final one; in other cases it might be one or both of the conditional marginal effects. In some cases the choice is clear, in others it may require some debate till consensus is reached or until it is agreed that multiple measures have value. The key point here is that it is imperative to identify (and justify) which effect is being analyzed in a given context. It is all too easy to misstate the effects of changing BE by applying a conditional effect unconditionally, or conversely. Specifically, for example, it would not be appropriate to project the conditional marginal effect of a change in BE on TB_U to the population as a whole.

As indicated above, to obtain the *total* marginal effect of BE on TB, it is generally necessary to account both for its effect on RC and its direct effect on TB, given RC. By contrast, the “*true*” effect of BE referred to in Table 2 is the portion of the total effect that is *not* due to residential self-selection. Again, however, that effect is in general not simply the coefficient of BE in the equation for TB, and its computation differs across the different measures listed above. For the final total unconditional marginal effect $\partial E[\text{TB}] / \partial \text{BE}$, for example, the true effect of BE would be the weighted average of the coefficients of BE in the equations for TB_S and TB_U , where the weights are the respective selection probabilities.

Turning to joint discrete choice models, the idea of the decomposition of the effect of BE on TB into two components – that of the direct effect on TB and the indirect effect through its effect on RC – also applies here as well, as illustrated by the simple example of footnote 3. For such models, the elasticity approach applied to nested logit by Salon (2006) and described in Section 3.5.1 above would be appropriate for the Bhat and Guo methodology as well.

For structural equations models, an important distinction is whether the model is *recursive* (i.e. having neither feedback loops, nor correlated error terms between directly linked endogenous variables) or *nonrecursive* (the in-between types of block and block-recursive are also possible, but for simplicity we focus on the two extremes). To illustrate a recursive model, consider the simple structure of Figure 2,

where BE is represented by the residential location construct. In this situation, it is useful to start with the *total association* (zero-order correlation) between BE and TB. This total association can be decomposed into (1) “spurious” components due to the common dependence of BE and TB on the antecedent variables AT and sociodemographic traits, (2) terms related to the unanalyzed correlations between the predetermined variables AT and sociodemographics themselves, and the (3) direct and (4) indirect effects of BE on TB (for an example, see the block-recursive model analyzed by Wolfle, 1980, and let his X_1 represent our TB, his X_3 be BE, and X_4 be AT; also see Alwin and Hauser, 1975). In Figure 2 there are no indirect effects of BE on TB (i.e. those occurring through the impact of BE on an intervening variable that then affects TB), so the total effect, which is the sum of the direct and all indirect effects, is simply the direct effect. This is captured by the coefficient a of BE in the equation for TB, and represents the true effect of BE on TB (however, see Cao et al., 2007a for a discussion of the case in which c is insignificant).

[Insert Figure 2 here]

For nonrecursive models, e.g. if in Figure 2, TB were modeled as affecting AT (consistent with Figure 1b) as well as the converse, computation of the total effect of BE on TB becomes more complex, and also includes the effect of AT on BE (see, e.g., Mueller, 1996). In such cases, the familiar regression-model interpretation of a coefficient as representing the marginal effect on the target variable of changing another variable by one unit, holding all other variables constant, is essentially meaningless (Hayduk, 1987). Changing BE would change TB, which would change AT – so AT could not be held constant. For the most robust SEMs, then (i.e. those allowing multiple directions of causality), it is quite difficult, if not impossible, to isolate the true effect of BE on TB (that is, the separate BE effect remaining after the influence of AT is accounted for). The stability index of a nonrecursive system (Bentler and Freeman, 1983) gives a mathematical indication of whether the infinite loops of impacts converge or diverge, but does not help separate out individual effects, nor confirm whether the system is truly in equilibrium or not (as required for the model to be valid).

4.2.2 True effects of discrete-valued measures of residential choice on TB

For the case of a discrete-valued RC measure, the program/policy evaluation context provides a useful framework within which to view the current topic. In fact, in view of the linkage between the built environment and physical activity, some scholars addressing the residential self-selection question come from a medical/health perspective (see, e.g., Vol. 18, No. 1 of the *American Journal of Health Promotion*, a special issue on health-promoting community design; and TRB-IOM, 2005), in which it is routine to evaluate the impacts of discretely-measured new medical treatments or public health-promotion programs or policies. The recent theoretical and applied econometric literature on policy evaluation (e.g., Winship and Morgan, 1999; Heckman and Vytlacil, 2005) offers a deepening understanding of this framework. The basic scenario in this literature is that there is a discrete *treatment* (e.g. a new policy), which is chosen by or applied to some of the population to whom it is available (the *treated*) but not others (the *untreated*; there can be multiple categories of (non)treatment, but for simplicity we will restrict the discussion to two).

The generic question is, what is the true effect of the treatment on an *outcome* variable of interest, particularly in the presence of non-random selection into the treated versus untreated groups (specifically, when selection into groups is correlated with variables affecting the outcome)? In cross-sectional analysis, each individual is observed in either a treated or untreated state, but not both. To evaluate the treatment effect, what we typically *want* to do is compare the outcomes of a randomly selected person who is moved from untreated to treated (but see below for further discussion of this point). What we *can* do, instead, is compare the observed outcomes of the treated to those of the untreated. That comparison will be a biased estimator of the true effect if (1) treated individuals initially differ from the untreated on variables relevant to predicting the outcome (which means the two groups would have different outcomes even in the absence of the treatment); and/or (2) the treated differ from the untreated in their potential

reaction to the treatment (different functions, or different parameters of the functions, relating explanatory variables to outcomes).

In our context, we have generically referred to the outcome of interest as travel behavior (TB) and the treatment as residential choice (RC). We have alluded to both forms of bias: one in which the treated (say, residents of traditional neighborhoods) differ from the untreated (residents of suburban neighborhoods) on the levels of their initial attitudes (AT) toward BE and TB (among other variables), and the other in which the impact of a *given* level of, e.g., BE or AT on TB could differ by residential neighborhood type. We are interested in the effect of “switching from 0 (suburban) to 1 (urban)”, but what exactly does that mean? A fundamental contribution of the recent literature is the articulation of numerous potentially relevant effects of interest, and their unification under the common framework of marginal treatment effects (see, e.g., Heckman and Vytlacil, 2005). We mention several possibilities here, described in terms of the current topic. For concreteness, we take the treatment to be the choice to live in an “urban” (meaning traditional or neotraditional, transit- and/or pedestrian-oriented, etc.) neighborhood, and the outcome of interest to be vehicle-kilometers of automobile travel (auto VKT) for residents of such neighborhoods:

- ◆ ***(conditional) marginal treatment effect (MTE(x, ε_R))***: What would be the average effect on auto VKT of moving from a suburban neighborhood to an urban one, given observed variables affecting the outcome (the X , Y , and BE of equation system (7), here referred to collectively as x), and unobserved variables (say, including AT) affecting participation (ε_R)? In essence, this question asks, what would be the effect of moving a specific randomly-selected person from suburban to urban, averaged over *all people in the population having identical observed characteristics affecting TB and identical unobserved characteristics affecting RC*?
- ◆ ***the average treatment effect (ATE(x))***: What would be the average effect on auto VKT of moving a randomly-selected person having observed characteristics x from suburban to urban? This effect would be obtained by averaging the MTE over the distribution of ε_R in the entire population.
- ◆ ***the average effect of treatment on the treated (TT(x))***: What would be the average effect on auto VKT of having moved a randomly-selected *urban resident* with observed characteristics x from a suburban neighborhood to an urban one? This effect is obtained by averaging the MTE over the distribution of ε_R in the population of urban residents.
- ◆ ***the average effect of treatment on the untreated (TUT(x))***: What would be the average effect on auto VKT of moving a randomly-selected *suburban resident* with observed characteristics x to an urban neighborhood? Naturally enough, this effect is obtained by averaging the MTE over the distribution of ε_R in the population of suburban residents.
- ◆ ***local average treatment effect (LATE(x, z_0, z_1))***: As mentioned in Section 3.4, it is generally assumed that a sample selection model contains at least one explanatory variable Z that influences participation but has no direct effect on the outcome; such variables are instrumental variables for participation. A typical example of such a variable is one indicating the presence or degree of an *incentive* to participate. In a study of the impact of a college education on earnings, for example, a variable indicating the availability or amount of financial aid might be considered a useful instrumental variable in the participation model, without affecting the outcome given participation. In our context, such a variable might be the presence of financial incentives for households to move to urban neighborhoods. Then it is of interest to ask what would be the average effect on auto VKT for a person with characteristics x , if Z changes from z_0 to z_1 . That is, loosely speaking, *what is the effect for those who needed an incentive to move from a suburban neighborhood to an urban one*? This is the LATE (Imbens and Angrist, 1994; Imbens, 2001); it focuses on the effectiveness of the incentive

by eliminating those who would have moved anyway (but see Winship and Morgan, 1999, p. 685 for some problems of the LATE approach).

In general, we expect $TT \geq ATE \geq TUT$. If the effect on auto VKT is independent of ϵ_R given x , then MTE, ATE, TT, and LATE are equal; estimating the true effect is challenging precisely because the effect of treatment is assumed to differ with different values of AT (which is a component of ϵ_R). On the other hand, if AT is observed (part of the x), then independence of the treatment effect from unobserved influences on participation may be a plausible assumption.

The answer to “which effect do we want?” is similar to that for the marginal effects of a continuous-valued BE in the previous subsection: “It depends”. Any of the measures above could be of interest in a given context. As with BE in Section 4.2.1, the key message here is that they do differ, in general, and accordingly should not be confused.

4.3 What is the effect of BE on TB in a relative sense?

Identifying the size of the true effect of the built environment on travel behavior (the subject of Section 4.2) is important, but in isolation it is often not very informative. It is desirable to have a way to put that effect in context, to scale it or judge it in a relative sense. This suggests two other questions of interest, differing in whether it is the “totality” of BE or of TB which is taken as the benchmark. In the first question, the total variation of TB constitutes the relevant denominator; in the second question it is the total impact of BE (on TB).

4.3.1 What proportion of the total (or even explained) variation in TB is truly due to BE?

Answering this question is important to properly evaluating the ability of changes in the BE to stimulate meaningful changes in TB: is a given-magnitude marginal change large in relative terms, or a tiny drop in the bucket relative to an individual’s overall travel?

It is telling that while virtually all of the studies reviewed for this work emphasize the statistical significance of the BE after self-selection has been controlled for, and several comment that the BE is only one (type) of a number of variables influencing travel behavior, only one study (Salon, 2006) directly quantified the true contribution of the BE to the explained variation in travel behavior. (At least two SEM studies – Meurs and Haaijer, 2001 and Bagley and Mokhtarian, 2002 – illustrated the total contribution of the BE to TB, but did not separate out the impact of self-selection of the BE, for reasons explained in Section 4.2.1). This is despite the fact that it is relatively easy to assess the proportionate magnitude of the impact of the BE, controlling for self-selection (at least approximately) for several of the methodologies reviewed here, by assessing the incremental contribution of BE variables to an R^2 or model log-likelihood measure, after AT is included. Alternatively, one could examine the change in TB predicted by a change in the BE (which has the advantage of expressing the BE influence in terms of “real” measures such as trips or distance traveled), but in so doing, it is essential to control for confounding factors. For example, Schwanen and Mokhtarian (2005b) compare the predicted weekly distance traveled of “average” consonant and dissonant urban and suburban residents, but the resulting differences are the combined effect of differences in sociodemographic and other characteristics across the four categories, as well as differences in one’s personal interaction with the built environment. If the average consonant urban resident is (attitudinally and sociodemographically) equivalent to the average dissonant suburban resident (and similarly for consonant suburban and dissonant urban residents), then the differences in their travel behavior could reasonably be attributed to the true effect of the built environment – however, that information was not presented.

We suspect this type of analysis is missing from published studies in part because the answer is expected (or found) to be “very little”, compared to the contributions of sociodemographic and unmeasured variables (as implied by the elasticities reported and computed in the review article of Ewing and Cervero,

2001, and by the earlier work of Hanson, 1982 and Weisbrod et al., 1980). However, such an outcome should be neither terribly surprising nor embarrassing for a complex behavior such as travel, which has numerous influences both systematic and idiosyncratic. For the contribution of the BE to TB to be small would not render pointless any attempt to reshape the BE – as discussed elsewhere (e.g. Handy et al., 2006), there are many reasons for improving the BE beyond influencing travel behavior (such as increasing the diversity of available housing options), and even small contributions can be useful at the margin. But as long as changing travel behavior *is* one of the reasons evinced for changing the BE, it is relevant to know how effectively that particular goal is likely to be met (not to mention, *which* elements of the BE are most effective at influencing TB).

4.3.2 Of the total influence of the built environment on travel behavior, what proportion is due to residential self-selection, and what proportion due to the separate influence of the built environment itself?

This question reflects the desire simply to decompose the total influence of BE on TB (whether determined to be large or small) into the component that is due to AT, versus the component due to the true influence of BE. The final column of Table 2 summarizes how to answer this question using each of the approaches studied here. Several points are noteworthy.

First, using the the IV approach, it is not possible to answer this question: while the BE explanatory variables are purged of their correlation with attitudes (thus allowing the separate influence of the BE itself to be determined), the attitudes themselves remain in the error term, and the extent of their influence cannot be distinguished from that of other unobserved variables.

Second, we have not been able to find a discussion of this question in the literature on selection models. However, since as discussed in Section 4.2.1 above it is possible to decompose the effect of a change in BE on expected TB into components due to (1) the direct change in TB due to the change in BE and (2) the change in the probability of a particular RC outcome due to the change in BE, it seems natural to answer the title question with (loosely speaking) $(2)/[(1) + (2)]$ and $(1)/[(1) + (2)]$, respectively.

Third, with respect to using the SEM approach to answer this question, the situation again differs depending on whether the model in question is recursive or nonrecursive. For recursive models, computing equation-by-equation R^2 s is appropriate and relatively straightforward (Bentler and Raykov, 2000), with interpretation (proportion of total variance in the “left-hand side” variable explained by the model) identical to that of the R^2 for a single-equation regression model (see, e.g., Mueller, 1988 for an application). Thus, it is reasonable to suggest that, for recursive models, the title question of this subsection can be answered in the same way as for the statistical controls method. With respect to Figure 2, specifically, we can compute the ratio of the incremental change in the R^2 of the equation for TB when BE is added to a system containing all other variables (including AT) and relationships, to the incremental change when AT and BE are added together. For nonrecursive models, not surprisingly, the situation is more complex. An important recent paper (Hayduk, 2006) offers a meaningful definition of R^2 for an endogenous variable in either a recursive or nonrecursive SEM⁷, but decomposing such an R^2 into components due to specific explanatory variables has not, to our knowledge, been addressed by the literature.

Empirically, this question is being addressed. As shown in the final column of Table 1, several studies indicated which of the two factors – residential self-selection or the built environment itself – was stronger (and of the ten that did, the answers were somewhat mixed: self-selection for two; the built environment for eight). Using the statistical control approach, Kitamura et al. (1997) evaluated the

⁷ In our context, the measure, called the “blocked-error R^2 ”, is defined as the ratio of the reduced variance in TB accounted for when the ω_1 of equation system (10) is blocked from affecting TB but the rest of the model is in place, to the full variance in TB accounted for when the entire model (including the effect of ω_1 on TB) is operating fully.

relative contributions of built environment variables and attitudes by gradually including different groups of variables in their model specifications; Schwanen and Mokhtarian (2003) also discussed the relative importance by comparing the behavior of consonant and dissonant residents. These studies stopped short of explicitly decomposing the total influence of the built environment into parts due to self-selection and the BE itself, although in principle at least an approximate partition could be made along the lines suggested in the last column of Table 2. Bagley and Mokhtarian (2002) made a qualitative statement that the influence of the BE per se on travel behavior was small relative to that of attitudes, but given their nonrecursive SEM structure, did not quantify the relative proportions of each.

5. CONCLUSIONS AND RECOMMENDATIONS

Over the past few years, disentangling the influences of the built environment and residential self-selection and determining their relative importance has become one of the most important emerging issues in understanding the relationship between the built environment and travel behavior (Krizek, 2003b; TRB-IOM, 2005). This report systematically discussed the requisites of causality inference in the context of the built environment and travel behavior, and identified nine approaches used in previous research to empirically address the issue of residential self-selection. The direct questioning method is designed to qualitatively evaluate the process of residential choice and travel choice, and hence is statistically unable to establish the evidence required for confident causal inference. The ability of the other approaches to meet the causality requisites is summarized in Table 3.

[Insert Table 3 here]

The characterization of each approach is somewhat subjective, but is based on “best practice”, in terms of operationalization to date, with respect to each method (with the exceptions of the sample selection and longitudinal structural equations methods, which have not yet been applied in pure form). Thus, a deficient application of a given approach would not rise to the levels indicated here (for example, a statistical control model that omitted numerous important variables would be weaker than indicated on the nonspuriousness criterion), while an improved application could exceed them (e.g. a joint simultaneous discrete choice model that included attitudes would be stronger than indicated on nonspuriousness).

All the statistical methods reviewed here can rely on the kinds of causal mechanisms described in Section 2, and all can be considered strong in terms of their ability to identify significant associations between the BE and TB. Thus, they differ only in how well they meet the nonspuriousness and time precedence criteria. In our view, approaches that explicitly include attitudes can perform well on the nonspuriousness criterion (by leaving little room for significant results to be due to spurious correlation with unmeasured variables), while those that permit multiple directions of causality and/or involve measurement at multiple points in time can excel on the time precedence criterion. In many cases of interest, the conceptual ideal is the longitudinal structural equations modeling approach, which combines most of the strengths of the other methods: measurement of attitudes, allowance of multiple directions of causality, and measurement at multiple points in time. If, when used to evaluate a “treatment” such as a residential move or BE intervention, control groups as well as experimental groups are involved, this approach comes very close to being “airtight” (though questions about generalizability could still remain, and the limitations discussed in Section 3 should be kept in mind). Although this method has not yet been fully operationalized in the present context (Cao et al., 2007b, comes the closest, to our knowledge, but does not include a control group and is only quasi-longitudinal in that “prior” measures are obtained only retrospectively, and do not include attitudes – though current attitudes are measured), a project is underway in Australia (Giles-Corti et al., 2008) which aims to do exactly that.

With respect to findings, research using the direct questioning method qualitatively found some evidence for residential self-selection. Studies using the statistical control approach consistently found a pervasive confounding influence of self-selection in the association between the built environment and travel

behavior, and most studies also found that the built environment has a separate influence on travel behavior (e.g., Cao et al., 2006a; Kitamura et al., 1997). Instrumental variables regression and sample selection models found evidence (strong in some studies; weak in others) that the built environment had an impact after controlling for self-selection. Nested logit applications (Salon, 2006) report a sizable influence of self-selection on travel behavior, with the built environment sustaining a direct influence beyond that. Two studies using propensity score found a true influence of the built environment. With respect to the joint discrete choice model, Bhat and Guo (2007) found no influence of self-selection due to unmeasured variables such as attitudes, but its extension concluded an influence of both the built environment and self-selection (Pinjari et al., 2007). The studies adopting a structural equations modeling approach (e.g. Bagley and Mokhtarian, 2002; Cao et al., 2007b) found an influence of residential selection, although the influence of the built environment appeared to be stronger than that of self-selection in the latter study. Investigations employing a longitudinal design tended to support the argument that the built environment has a causal influence on travel behavior, although they acknowledged the potential influence of attitudinal factors.

Unfortunately, given the various limitations discussed throughout this paper, we are unable at this point to confidently specify the nature and extent of the causality between the built environment and travel behavior. We identified a number of different ways the question of interest could be posed. In general, ironically, it seems as though the more sophisticated the approach to treating self-selection (and therefore, presumably, the more trustworthy the resulting effects that are identified), the more difficult it becomes to answer questions about the absolute and relative magnitudes of the true impacts of the built environment on travel behavior. In fact, those impacts differ by mode and trip purpose, as some studies have shown. They are also likely to differ for different segments of the population, an issue not addressed by any of the empirical applications reviewed here (though most of the methodological approaches can accommodate it conceptually).

Nevertheless, we can improve our understanding by designing studies to satisfy as many requisites of causality inference as possible. Future studies adopting research designs that more closely resemble a true experimental design will lead to more definitive inferences regarding causality. Two types of studies are important (both of them ideally to include comparison groups of unaffected individuals similar in other relevant ways): (1) true panel studies of residents who move from one type of neighborhood to another, with measurements of attitudes as well as sociodemographic traits and travel behavior before and after, and further exploration of the reasons behind the move; and (2) natural experiments that examine the impact on travel behavior in response to a change in the built environment, such as the implementation of a traffic calming program. Only by causal findings based on such evidence can we determine whether land use policies designed to increase opportunities for driving less and walking more will actually lead to the desired behavioral outcomes.

ACKNOWLEDGEMENTS

This research was partly supported by a University of California Transportation Center Dissertation Grant. Chandra Bhat, David Brownstone, Hani Mahmassani, David Ory, Deborah Salon, Tim Schwanen, and Ed Vytlačil, as well as several anonymous referees, have helped clarify some ideas and improve this work. Any remaining errors of fact or interpretation are, of course, our responsibility.

REFERENCES

- Abraham, J. E. and Hunt, J. D. (1997). Specification and estimation of nested logit model of home, workplaces, and commuter model choices by multiple-worker households. *Transportation Research Record*, 1606, 17-24.
- Abreu e Silva, J., Golob, T., and Goulias, K. (2006). The effects of land use characteristics on residence and employment location and travel behavior of urban adult workers. *Transportation Research Record*, 1977, 121-131.
- Alwin, Duane F. and Hauser, Robert M. (1975). The decomposition of effects in path analysis. *American Sociological Review*, 40 (1), 37-47.
- Bagley, M. N. and Mokhtarian, P. L. (2002). The impact of residential neighborhood type on travel behavior: A structural equations modeling approach. *The Annals of Regional Science*, 36, 279-297.
- Ben-Akiva, Moshe and Lerman, Steven R. (1985). *Discrete Choice Analysis: Theory and Application to Travel Demand*. Cambridge, Mass.: The MIT Press.
- Bentler, P. M. and Freeman, E. H. (1983). Tests for stability in linear structural equation systems. *Psychometrika*, 48, 143-145.
- Bentler, Peter M. and Raykov, Tenko (2000). On measures of explained variance in nonrecursive structural equation models. *Journal of Applied Psychology*, 85, 125-131.
- Bhat, C. R. and Guo, J. (2007). A comprehensive analysis of built environment characteristics on household residential choice and auto ownership levels. *Transportation Research Part B*, 41 (5), 506-526.
- Boarnet, M. G., Anderson, C. L., Day, K., McMillan, T., and Alfonzo, M. (2005). Evaluation of the California Safe Routes to School Legislation: Urban form changes and children's active transportation to school. *American Journal of Preventive Medicine*, 28, 134-140.
- Boarnet, M. G. and Crane, R. (2001). *Travel by Design: The Influence of Urban Form on Travel*. New York: Oxford University Press.
- Boarnet, M. G. and Sarmiento, S. (1998). Can land use policy really affect travel behaviour? A study of the link between non-work travel and land-use characteristics. *Urban Studies*, 35 (7), 1155-1169.
- Boer, R., Zheng, Y., Overton, A., Ridgeway, G. K., and Cohen, D. A. (2007). Neighborhood design and walking trips in ten U.S. metropolitan areas. *American Journal of Preventive Medicine*, 32, 298-304.
- Bound, John, Jaeger, David A., and Baker, Regina M. (1995). Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak. *Journal of the American Statistical Association*, 90 (430), 443-450.
- Bourguignon, Francois, Fournier, Martin, and Gurgand, Marc (2007). Selection bias corrections based on the multinomial logit model: Monte Carlo comparisons. *Journal of Economic Surveys*, 21 (1), 174-205.
- Cameron, A. Colin and Trivedi, Pravin K. (2005). *Microeconomics: Methods and Applications*. New York: Cambridge University Press.

- Cao, X., Mokhtarian, P. L., and Handy, S. L. (2005). The impacts of the built environment and residential self-selection on nonwork travel: A seemingly unrelated regression approach. Paper 06-1595 on the Transportation Research Board 85th Annual Meeting CD-ROM. Washington, DC.
- Cao, X., Handy, S. L., and Mokhtarian, P. L. (2006a). The influences of the built environment and residential self-selection on pedestrian behavior: Evidence from Austin, TX. *Transportation*, 33 (1), 1-20.
- Cao, X., Mokhtarian, P. L., and Handy, S. L. (2006b). Neighborhood design and vehicle type choice: Evidence from Northern California. *Transportation Research Part D*, 13 (2), 133-145.
- Cao, X., Mokhtarian, P. L., and Handy, S. L. (2007a). Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership. *Environment and Planning A*, 39, 830-847.
- Cao, X., Mokhtarian, P. L., and Handy, S. L. (2007b). Do changes in neighborhood characteristics lead to changes in travel behavior? A structural equations modeling approach. *Transportation*, 34 (5).
- Cao, X. (2008). Exploring causal effects of neighborhood design on travel behavior using stratification on the propensity score. Paper 09-0115 on the Transportation Research Board 88th Annual Meeting CD-ROM. Washington, DC.
- Cervero, R. (2007). Transit-oriented development's ridership bonus: A product of self-selection and public policies. *Environment and Planning A*, 39, 2068-2085.
- Cervero, R. and Duncan, M. (2003). Walking, bicycling, and urban landscapes: Evidence from San Francisco Bay Area. *American Journal of Public Health*, 93 (9), 1478-1483.
- Chatman, Daniel G. (2005). *How the Built Environment Influences Non-work Travel: Theoretical and Empirical Essays*. Dissertation submitted in fulfillment of the Ph.D. in Urban Planning, University of California, Los Angeles.
- Chen, C., Gong, H., and Paaswell, R. (2008). Role of the built environment on mode choice decisions: additional evidence on the impact of density. *Transportation*, 35, 285-299.
- Clifton, K. and Handy, S. (2003). Qualitative methods in travel behaviour research. In D. Jones and P. Stopher, eds., *Transport Survey Quality and Innovation*. Oxford, UK: Elsevier Science, Ltd.
- Circella, Giovanni, Mokhtarian, Patricia L., and Handy, Susan L. (2008) Land use, attitudes, and travel behavior relationships: A cross-sectional structural equations model for Northern California. Manuscript under review for publication; available from the authors.
- Crane, R. and Crepeau, R. (1998). Does neighborhood design influence travel? A behavioral analysis of travel diary and GIS data. *Transportation Research Part D*, 3 (4), 225-238.
- D'Agostino, Ralph B. Jr. (1998). Propensity score methods for bias reduction in the comparison of a treatment and a non-randomized control group. *Statistics in Medicine*, 17, 2265-2281.
- Dillman, Don (1978). *Mail and Telephone Surveys: The Total Design Method*. New York: John Wiley.
- Dow, William H. and Norton, Edward C. (2003). Choosing between and interpreting the Heckit and two-part models for corner solutions. *Health Services and Outcomes Research Methodology*, 4, 5-18.

- Ewing, Reid and Cervero, Robert (2001). Travel and the built environment: A synthesis. *Transportation Research Record*, 1780, 87-113.
- Forinash, Christopher V. and Koppelman, Frank S. (1993). Application and interpretation of nested logit models of intercity mode choice. *Transportation Research Record*, 1413, 98-106.
- Frank, Lawrence Douglas, Saelens, Brian E., Powell, Ken E., and Chapman, James E. (2007). Stepping towards causation: Do built environments or neighborhood and travel preferences explain physical activity, driving, and obesity? *Social Science and Medicine*, 65, 1898-1914.
- Frank, Lawrence D., Sallis, James F., Conway, Terry L., Chapman, James E., Saelens, Brian E., and Bachman, William (2006). Many pathways from land use to health: Associations between neighborhood walkability and active transportation, body mass index, and air quality. *Journal of the American Planning Association*, 72 (1), 75-87.
- Giles-Corti, B., Knuiman, M., Timperio, A., van Niel, K., Pikora, T. J., Bull, F. C. L., Shilton, T. and Bulsara, M. (2008). Evaluation of the implementation of a state government community design policy aimed at increasing local walking: Design issues and baseline results from RESIDE, Perth Western Australia. *Preventive Medicine*, 46, 46-54.
- Golob, Thomas F. (2001). Joint models of attitudes and behavior in evaluation of the San Diego I-15 congestion pricing project. *Transportation Research Part A*, 35 (6), 495-514.
- Greene, William H. (1997). *Econometric Analysis, 3rd edition*. Upper Saddle River, NJ: Prentice Hall.
- Greenwald, Michael J. (2003). The road less traveled: New urbanist inducements to travel mode substitution for nonwork trips. *Journal of Planning Education and Research*, 23, 39-57.
- Greenwald, M. J. and Boarnet, M. G. (2001). Built environment as determinant of walking behavior: Analyzing nonwork pedestrian travel in Portland, Oregon. *Transportation Research Record*, 1780, 33-42.
- Hall, Alastair R., Rudebusch, Glenn D., and Wilcox, David W. (1996). Judging instrument relevance in instrumental variables estimation. *International Economic Review*, 37 (2), 283-298.
- Hammond, Daniel (2005). *Residential Location and Commute Mode Choice*. Dissertation submitted in fulfillment of the MSC in Transport and Planning, University of Wales, Cardiff.
- Handy, S. L. and Clifton, K. J. (2001). Local shopping as a strategy for reducing automobile travel. *Transportation*, 28 (4), 317-346.
- Handy, S., Cao, X., and Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D*, 10 (6), 427-444.
- Handy, S., Cao, X., and Mokhtarian, P. L. (2006). Self-selection in the relationship between built environment and walking? Evidence from Northern California. *Journal of the American Planning Association*, 72 (1), 55-74.
- Hanson, Susan (1982). The determinants of daily travel-activity patterns: Relative location and sociodemographic factors. *Urban Geography*, 3 (3), 179-202.

- Hayduk, Leslie A. (1987). *Structural Equation Modeling with LISREL: Essentials and Advances*. Baltimore, MD: Johns Hopkins University Press.
- Hayduk, Leslie A. (2006). Blocked-error-R²: A conceptually improved definition of the proportion of explained variance in models containing loops or correlated residuals. *Quality and Quantity*, 40, 629-649.
- Heckman, James (1990) Varieties of selection bias. *The American Economic Review*, 80 (2), 313-318.
- Heckman, J., Tobias, J. L., and Vytlačil, E. (2001). Four parameters of interest in the evaluation of social programs. *Southern Economic Journal*, 68 (2), 210-223.
- Heckman, James J. and Vytlačil, Edward (2005). Structural equations, treatment effects, and econometric policy evaluation. *Econometrica*, 73 (3), 669-738.
- Horowitz, Joel (1986). Modeling choices of residential location and mode of travel to work. Chapter 9 in *The Geography of Urban Transportation*, Hanson, Susan, ed. New York: The Guilford Press. 207-226.
- Huang, Chung L., Raunikar, Robert, and Misra, Sukant (1991). The application and economic interpretation of selectivity models. *American Journal of Agricultural Economics*, 73 (2), 496-501.
- Hume, David (1748). An enquiry concerning human understanding. Available at http://www.infidels.org/library/historical/david_hume/human_understanding.html, accessed on July 22, 2005.
- Imbens, Guido W. (2001). Some remarks on instrumental variables. In *Econometric Evaluation of Labour Market Policies*, Lechner, Michael and Pfeiffer, Friedhelm, eds. New York: Physica-Verlag. 18-42.
- Imbens, Guido W. and Angrist, Joshua D. (1994). Identification and estimation of local average treatment effects. *Econometrica*, 62 (2), 467-475.
- Khattak, A. J. and Rodriguez, D. (2005). Travel behavior in neo-traditional neighborhood developments: A case study in USA. *Transportation Research Part A*, 39, 481-500.
- Kitamura, R., Mokhtarian, P. L., and Laidet, L. (1997). A micro-analysis of land use and travel in five neighborhoods in the San Francisco Bay Area. *Transportation*, 24, 125-158.
- Kitamura, R., Akiyama, T., Yamamoto, T., and Golob, T. F. (2001). Accessibility in a metropolis: Toward a better understanding of land use and travel. *Transportation Research Record*, 1780, 64-75.
- Kline, Rex B. (2005). *Principles and Practice of Structural Equation Modeling*, 2nd ed. New York: The Guilford Press.
- Krizek, Kevin (2000). Pretest-posttest strategy for researching neighborhood-scale urban form and travel behavior. *Transportation Research Record*, 1722, 48-55.
- Krizek, Kevin (2003a). Residential relocation and changes in urban travel: Does neighborhood-scale urban form matter? *Journal of the American Planning Association*, 69 (3), 265-281.
- Krizek, Kevin (2003b). Planning, household travel, and household lifestyles. In *Transportation Systems Planning: Methods and Applications*, ed. Goulias, K. G. Boca Raton, FL: CRC Press.

- Lee, Lung-Fei (1981). Simultaneous equations models with discrete and censored variables. Chapter 9 in *Structural Analysis of Discrete Data with Econometric Applications*, eds. Manski, Charles F. and McFadden, Daniel. Cambridge, Mass: The MIT Press. 346-364.
- Lee, Lung-Fei (1983). Generalized econometric models with selectivity. *Econometrica*, 51 (2), 507-512.
- Lewbel, Arthur (2004). Simple estimators for hard problems: Endogeneity in discrete choice related models. Working paper 604, Department of Economics, Boston College. Available at <http://fmwww.bc.edu/EC-P/WP604.pdf>, accessed June 12, 2006.
- Litman, Todd A. (2005). *Land Use Impacts on Transport: How Land Use Factors Affect Travel Behavior*. Victoria Transport Institute. Available at <http://www.vtpi.org/landtravel.pdf>, accessed on January 13, 2006.
- Louviere, Jordan, Train, Kenneth, Ben-Akiva, Moshe, Bhat, Chandra, Brownstone, David, Cameron, Trudy Ann, Carson, Richard T., Deshazo, J. R., Fiebig, Denzil, Greene, William, Hensher, David, and Waldman, Donald (2005). Recent progress on endogeneity in choice modeling. *Marketing Letters*, 16 (3/4), 255-265.
- MacCallum, Robert C. (1995). Model specification: Procedures, strategies, and related issues. Chapter 2 in *Structural Equation Modeling: Concepts, Issues, and Applications*, ed. Hoyle, Rick H. Thousand Oaks, California: Sage Publications. 16-36.
- McBeth, Andrew G. (1999). Bicycle lanes in Toronto. *ITE Journal*, 69, 38-46.
- Maddala, G. S. (1983) *Limited-Dependent and Qualitative Variables in Econometrics*. Cambridge: Cambridge University Press.
- Meurs, H. and Haaijer, R. (2001). Spatial structure and mobility. *Transportation Research Part D*, 6 (6), 429-446.
- Mokhtarian, P. L. and Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B*, 43 (3), 204-228.
- Mueller, Ralph O. (1988). The impact of college selectivity on income for men and women. *Research in Higher Education*, 29 (2), 175-191.
- Mueller, Ralph O. (1996). *Basic Principles of Structural Equation Modeling: An Introduction to LISREL and EQS*. New York: Springer.
- Murakami, E. and Young, J. (1997). *Daily Travel by Persons with Low Income*. Available at <http://www.cta.ornl.gov/npts/1995/Doc/LowInc.pdf>, accessed July 1, 2004.
- Muthén, Bengt (1983). Latent variable structural equation modeling with categorical data. *Journal of Econometrics*, 22, 43-65.
- Nelson, Charles R. and Startz, Richard (1990). The distribution of the instrumental variables estimator and its t-ratio when the instrument is a poor one. *The Journal of Business*, 63 (1, Part 2), S125-S140.
- Oakes, Michael J. (2004). The (mis)estimation of neighborhood effects: Causal inference for a practicable social epidemiology. *Social Science & Medicine*, 58, 1929-1952.

- Oakes, M. J. and Johnson, P. J. (2006). Propensity score matching for social epidemiology. Chapter 15 in *Methods in Epidemiology*, eds. Oakes, M. J. and Kaufman, J. S. New York: John Wiley & Sons, Inc.
- Painter, Kate (1996). The influence of street lighting improvements on crime, fear and pedestrian street use, after dark. *Landscape and Urban Planning*, 35, 193-201.
- Pendyala, Ram M. (1998). Causal analysis in travel behaviour research: A cautionary note. Chapter 3 in *Travel Behaviour Research: Updating the State of Play*, eds. Ortuzar, Juan de Dios, Hensher, David, and Jara-Diaz, Sergio. Amsterdam: Pergamon. 35-48.
- Pinjari, Abdul Rawoof, Pendyala, Ram M., Bhat, Chandra R., and Waddell, Paul A. (2007). Modeling residential sorting effects to understand the impact of the built environment on commute mode choice. *Transportation*, 34 (5), 557-573.
- Ramanathan, Ramu (2002). *Introductory Econometrics with Applications*. Fort Worth, TX: Harcourt College Publishers.
- Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observation studies for causal effects. *Biometrika*, 70 (1), 41-55.
- Salomon, I. and Ben-Akiva, M. (1983). The use of the life-style concept in travel demand models. *Environment and Planning A*, 15, 623-638.
- Salon, Deborah (2006). *Cars and the City: An Investigation of Transportation and Residential Location Choices in New York City*. Dissertation submitted in fulfillment of the Ph.D. in Agricultural and Resource Economics, University of California, Davis.
- Scheiner, Joachim and Holz-Rau, Christian (2007). Travel mode choice: Affected by objective or subjective determinants? *Transportation*, 34, 487-511.
- Schmidt, Peter and Strauss, Robert P. (1975). Estimation of models with jointly dependent qualitative variables: A simultaneous logit approach. *Econometrica* 43 (4), 745-755.
- Schutt, Russell K. (2004). *Investigating the Social World: The Process and Practice of Research, 4th edition*. Thousand Oaks, CA: Pine Forge Press.
- Schwanen, T. and Mokhtarian, P. L. (2003). Does dissonance between desired and current neighborhood type affect individual travel behaviour? An empirical assessment from the San Francisco Bay Area. Proceedings of the European Transport Conference (ETC), October 8-10, 2003, Strasbourg, France.
- Schwanen, T. and Mokhtarian, P. L. (2005a). What affects commute mode choice: Neighborhood physical structure or preferences toward neighborhoods? *Journal of Transport Geography*, 13 (1), 83-99.
- Schwanen, T. and Mokhtarian, P. L. (2005b). What if you live in the wrong neighborhood? The impact of residential neighborhood type dissonance on distance traveled. *Transportation Research Part D*, 10 (2), 127-151.
- Shadish, W. R., Cook, T. D., and Campbell, D. T. (2002). *Experimental and Quasi-experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.

- Shea, John (1997). Instrument relevance in multivariate linear models: A simple measure. *The Review of Economics and Statistics*, 79, 348-352.
- Sigelman, Lee and Zeng, Langche (1999). Analyzing censored and sample-selected data with Tobit and Heckit models. *Political Analysis*, 8 (2), 167-182.
- Singleton, R. A. Jr. and Straits, B. C. (2005). *Approaches to Social Research*, 4th edition. New York: Oxford University Press.
- Sobel, Kenneth L. (1980). Travel demand forecasting with the nested multinomial logit model. *Transportation Research Record*, 775, 48-55.
- Staiger, Douglas and Stock, James H. (1997). Instrumental variables regression with weak instruments. *Econometrica*, 65 (3), 557-586.
- Tardiff, Timothy J. (1977). Causal inferences involving transportation attitudes and behavior. *Transportation Research*, 11 (6), 397-404.
- Train, Kenneth (1986). *Qualitative Choice Analysis: Theory, Econometrics, and an Application to Automobile Demand*. Cambridge, Mass.: The MIT Press.
- Train, Kenneth (2003). *Discrete Choice Methods with Simulation*. New York: Cambridge University Press. Available at <http://elsa.berkeley.edu/~train/distant.html>.
- Transportation Research Board and Institute of Medicine (TRB-IOM). (2005). *Does the Built Environment Influence Physical Activity? Examining the Evidence – Special Report 282*. Washington, DC. January. Available at <http://onlinepubs.trb.org/onlinepubs/sr/sr282.pdf>.
- Tringides, Constantinos A., Ye, Xin, and Pendyala, Ram M. (2004). Departure-time choice and mode choice for nonwork trips: Alternative formulations of joint model systems. *Transportation Research Record*, 1898, 1-9.
- Utts, Jessica M. (1999). *Seeing through Statistics*, 2nd edition. Pacific Grove, CA: Duxbury Press.
- van Acker, Veronique, Witlox, Frank, and van Wee, Bert (2007). The effects of the land use system on travel behavior: A structural equation modeling approach. *Transportation Planning and Technology*, 30(4), 331-353.
- Vance, Colin and Geoghegan, Jacqueline (2004). Modeling the determinants of semi-subsistent and commercial land uses in an agricultural frontier of southern Mexico: A switching regression approach. *International Regional Science Review*, 27 (3), 326-347.
- Vance, Colin and Hedel, Ralf (2007). The impact of urban form on automobile travel: Disentangling causation from correlation. *Transportation*, 34 (5), 575-588.
- Wells, Nancy M. and Yang, Yizhao (2008). Neighborhood design and walking - A quasi-experimental longitudinal study. *American Journal of Preventive Medicine*, 34(4), 313-319.
- Weisbrod, G. E., Lerman, S. R., and Ben-Akiva, M. (1980). Tradeoffs in residential location decisions: Transportation versus other factors. *Transport Policy and Decision Making*, 1, 13-26.

Winship, Christopher and Morgan, Stephen L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology*, 25, 659-706.

Wolfle, Lee M. (1980). Strategies of path analysis. *American Educational Research Journal*, 17 (2), 183-209.

Ye, Xin, Pendyala, Ram M., and Gottardi, Giovanni (2007). An exploration of the relationship between mode choice and complexity of trip chaining patterns. *Transportation Research Part B* 41, 96-113.

Zhou, B. and Kockelman, K. (2008). Self-selection in home choice: Use of treatment effects in evaluating the relationship between the built environment and travel behavior. *Transportation Research Record*.

Ziliak, Stephen T. and McCloskey, Deirdre N. (2004). Size matters: The standard error of regressions in the American Economic Review. *The Journal of Socio-Economics*, 33 (5) (special issue on Statistical Significance, ed. Altman, M.), 527-546.

Figure 1. Some Potential Relationships among Travel Attitudes, Built Environment, and Travel Behavior

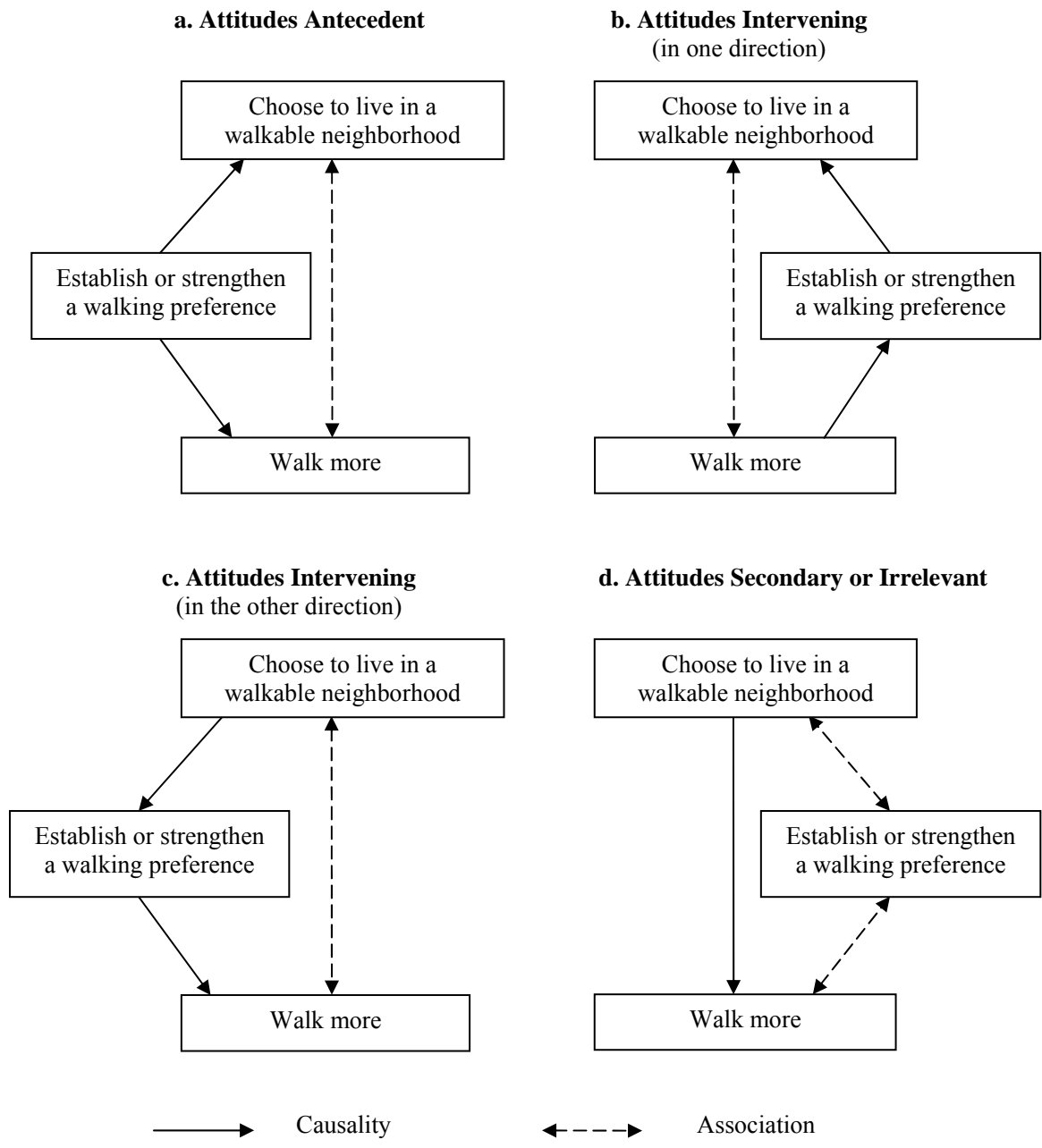


Figure 2: A Simple Recursive Structural Equations Model of Residential Location and Travel Behavior

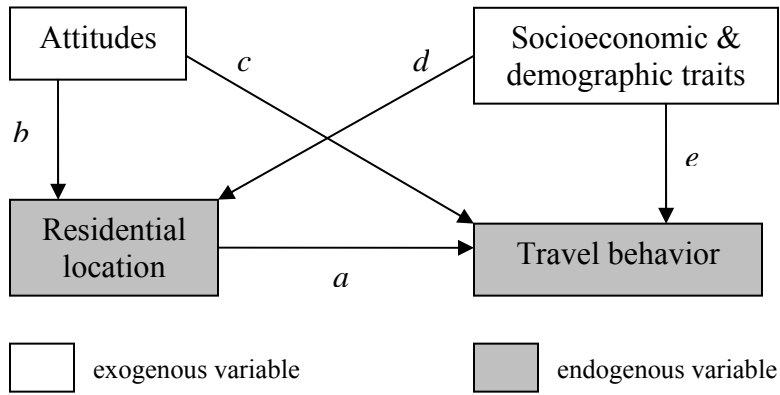


Table 1. Overview of Studies Addressing Residential Self-Selection

Studies	Sample	Methodology	Travel Behavior Measurements	Built Environment Measurements	Attitude Measurements	Conclusions¹
Direct questioning						
Hammond, 2005	90 respondents and 8 interview participants in Century Wharf, Cardiff, UK, 2004	Descriptive and correlational analyses	Changes in car use to work	Moving to the city center	8 measures for residential preferences	BE and SS ² . Residents moving to the city center reduced car use to work; residential choice was either conditional on or interacted with current commute mode choice for most respondents.
Handy and Clifton, 2001	1,368 individuals and 75 interview participants in Austin, TX, 1995	Descriptive analysis and linear regression	Walking to store frequency	Miles to store, perceived store characteristics, and neighborhood dummy	Not available	BE and SS. Local store characteristics influenced walking frequency; but “having the option to walk to the store is to some extent an effect of the desire to walk to the store.”
Statistical control models						
Cao et al., 2006a	1,368 individuals in Austin, TX, 1995	Negative binomial regression	Strolling frequency and walking to store frequency	Objective and perceived neighborhood characteristics, perceived store characteristics	Residential preference for stores within walking distance	BE and SS. Residential preference is the most important single factor explaining walking to store frequency; neighborhood characteristics also had a separate influence on strolling frequency, while characteristics of local commercial areas had a separate influence on shopping trips.
Cao et al., 2006b	1,682 individuals from Northern California, 2003	Nested logit model	Vehicle type choice	Objective and perceived neighborhood characteristics	Various measures for residential preferences and travel attitudes	BE and SS. Attitudinal factors influenced vehicle type choice; BE also had a separate influence.
Cao et al., 2005	1,682 individuals from Northern California, 2003	Seemingly unrelated regression	Frequencies of nonwork trips by auto, transit, and walking/biking	Objective and perceived neighborhood characteristics	Various measures for residential preferences and travel attitudes	BE and SS. Residential self-selection is more likely to influence walking/biking trips than auto and transit trips; BE also had a separate influence on all trips.
*Chatman, 2005	1,114 adults in San Francisco and San Diego metro areas,	Negative binomial regression	Number of nonwork activities accessed by auto, transit, and walk/	Distance from rail, bus frequency, no. of 4-way intersections	Preferences for auto, transit, walk/bike	BE > SS. Mode preferences affected nonwork transit and walk/bike travel, but not auto; BE affected nonwork transit and

	2003		bike; nonwork auto mileage			walk/bike travel after preferences included; impact of BE independent of preferences.
*Frank et al., 2007	two sub-samples (n=2,056 and n=1,466) from the 2001-2002 SMARTRAQ	Linear regression	Percent of respondents taking walking trips, and vehicle miles travelled	Walkability index	Two factors of residential preferences	BE > SS for driving; BE and SS for walking. BE had an influence on walking and driving, controlling for residential preference; BE had a stronger influence on driving than residential preference, while the latter had a stronger influence on walking.
*Kitamura et al., 1997	963 households in the San Francisco Bay Area, CA, 1993	Linear regression	Numbers of trips by non-motorized modes, transit, and all modes; fractions of auto trips, transit trips, and non-motorized trips	Residential density, land use mix, and rail transit accessibility	8 attitude factors	BE < SS. The residential environment had some influence on travel behavior, but attitudes carried more explanatory power in explaining the variation in travel behavior.
*Schwanen and Mokhtarian, 2003	1,358 workers in the San Francisco Bay Area, CA, 1998	Ordered probit model	Respective trip frequencies for 6 purposes	Traditional and suburban neighborhoods	Various measures for lifestyle, personality, and travel attitudes, neighborhood type mismatch (dissonance) indicators	BE > SS. BE exerted a significant influence on trip frequencies after SS controlled for, apparently more so on suburban dwellers than on urban residents.
*Schwanen and Mokhtarian, 2005a	1,358 workers in the San Francisco Bay Area, CA, 1998	Multinomial logit model	Commute mode choice	Traditional and suburban neighborhoods	Various measures for lifestyle, personality, and travel attitudes; neighborhood type mismatch (dissonance) indicators	BE > SS. Suburban-oriented urban dwellers were substantially more likely to commute by car than other urbanites, whereas urban-oriented suburban residents commuted by car at rates almost as high as other suburbanites. Thus, the separate influence of the BE, after controlling for SS, appears stronger in the suburbs.
*Schwanen and Mokhtarian, 2005b	1,358 workers in the San Francisco Bay Area, CA, 1998	Tobit model	Respective distance traveled by auto, rail, bus, walking/ jogging/	Traditional and suburban neighborhoods	Various measures for lifestyle, personality, and travel attitudes;	BE > SS. Significant influence of the BE on distance traveled remained after SS was controlled for, with the BE

			biking, and all modes		neighborhood type mismatch (dissonance) indicators	influence stronger for suburban dwellers than for urban dwellers.
Instrumental variables models						
Boarnet and Sarmiento, 1998	769 Southern California residents, 1993	Instrumental regression	Nonwork auto trip frequency	Density measures and street grid pattern at the block group/census tract and zip code levels	Not available	BE. BE at the neighborhood level had little influence on nonwork auto travel.
Greenwald and Boarnet, 2001	1,091 individuals in the 1994 Household Activity and Travel Behavior Survey in Portland, OR	Instrumental regression	Nonwork walking trip frequency	Density measures, street grid pattern, and pedestrian environment factor at census block group, census tract, and zip code levels	Not available	BE. The residential environment influenced nonwork walking trip generation at the neighborhood level.
Vance and Hedel 2007	4,328 individuals in the 1996-2003 German Mobility Panel	Instrumental regression	Car use and distance travelled	Commercial density, commercial diversity, street density, and walking minutes to public transit	Not available	BE. All measurements but commercial diversity had a causal effect on travel behavior.
Khattak and Rodriguez, 2005	453 households in Chapel Hill and Carrboro, NC	Binary choice model and negative binomial/linear regression	Frequencies of auto trips, walking trips and external trips; distances for all trips and nonwork trips; trip duration	Neo-traditional and suburban neighborhoods	8 measures for residential preference	BE. BE influenced most measures of travel behavior.
Sample selection models						
Greenwald, 2003	4,235 respondents in the 1994 Household Activity and Travel Behavior Survey in Portland, OR	Multinomial logit model and then linear regression	Eight substitution rates (walking/driving and transit/driving) for consumption, communication, socialization, and all trips	Six groups based on housing tenure and three levels of pedestrian environment factor, and zone-based land use characteristics	Not available	BE. New Urbanist designs increased walking substitution for driving, but had few effects on transit substitution for driving.
*Zhou and	1,903	Sample	Vehicle miles	CBD and urban vs.	Not available	BE > SS.

Kockelman 2008	households in the 1998-1999 Austin Travel Survey	selection model	travelled	rural and suburban		Self-selection accounted for 10% to 42% of the total influence of BE on driving behavior depending on model specifications.
Propensity score models						
Boer et al. 2007	10 metropolitan areas in the 1995 NPTS	Propensity score matching	Choice of walking	Land use mix, density, housing age, block length, parking pressure, and share of four-way intersections	Not available	BE and SS. A few influences from built environment elements remained after matching, but most became insignificant.
*Cao 2008	1,553 residents from Northern California, 2003	Propensity score stratification	Vehicle miles driven, strolling frequency, walking to the store frequency	Traditional vs. suburban neighborhoods	Various measures for residential preferences and travel attitudes	BE > SS. Self-selection accounts for 39% of the observed influence of neighborhood type on walking to store frequency, 14% for strolling frequency, and 22% for vehicle miles driven.
Simultaneous models						
*Bagley and Mokhtarian, 2002	515 individuals in the San Francisco Bay Area, CA, 1993	Structural equations model	Vehicle miles, transit miles, and walk/bike miles	Two factor scores: traditional and suburban, based on various measures such as residential density and land use mix	Various lifestyle and attitude factor scores	BE < SS. Residential location type had little separate impact on travel behavior; attitudes and lifestyles were the most important predictors of travel behavior.
Bhat and Guo, 2007	2,954 Alameda County households in the 2000 San Francisco Bay Area Travel Survey	Joint nominal/ordinal discrete choice model	Number of autos	Indicator of 233 TAZs, regional accessibility, density, land use traits, and transportation network characteristics	Not available	BE; no SS. BE had true effects on auto ownership; no evidence of self-selection was found.
Cervero 2007	11,369 workers in the 2000 San Francisco Bay Area Travel Survey	Nested logit model	Rail commute choice	Residential location within or beyond half a mile of a rail station	Not available	BE and SS. Odds of commuting by rail ~40% higher for those living within 1/2-mile of rail station, compared to those living farther away.
Chen et al., 2008	2,089 commuters in the New York	Simultaneous equation	Car ownership	Density, job accessibility	Not available	BE and SS. The unobserved attitude toward using a car influenced car ownership; the

	metropolitan region, 1997-1998					built environment had a separate influence beyond the attitude and demographics.
Circella et al., 2007	1,217 workers from Northern California, 2003	Structural equations model	Weekly miles driven	perceived neighborhood characteristics	Various measures for residential preferences and travel attitudes	BE and SS. Attitudes influenced both the choice of the built environment and driving behavior; the latter two were also associated.
Pinjari et al., 2007	1,878 Alameda County commuters in the 2000 San Francisco Bay Area Travel	Joint nominal/nominal discrete choice model	Commute mode choice	Indicator of 233 TAZs, regional accessibility, density, land use traits, and transportation network characteristics	Not available	BE and SS. Self-selection resulted from observed and unobserved factors; BE had a separate effect on mode choice, after accounting for the influence of self-selection.
*Salon, 2006	4,382 New York City respondents to the Regional Travel-HH Interview Survey	Nested logit model (resid. choice, auto ownership, walking level)	Walking level (none, some, a lot)	Population density	Not available	BE > SS. Self-selection accounted for 1/3 – 1/2 the total influence of BE.
Scheiner and Holz-Rau 2007	2,691 residents in Cologne, German, 2002-2003	Structural equations model	Mode use, vehicle kilometers travelled	Quality of transit, density of supply, and density and mixed use	Lifestyle factors and attitudes toward residential choice	BE and SS. Attitudes influenced both the choice of BE and TB; the latter two were also associated.
Longitudinal designs						
Boarnet et al., 2005	862 respondents to SR2S program, CA, 2002	T-tests	Walking/biking to school	SR2S projects including sidewalk, crossing, and traffic control improvements	Not available	BE. All improvements increased walking/biking to school for children.
Cao et al., 2007a	1,682 individuals from Northern California, 2003	Linear regression (ordered probit model ³)	Number of autos; changes in number of autos	Objective neighborhood characteristics; perceived neighborhood characteristics and their changes	Various measures for residential preferences and travel attitudes	BE and SS. The cross-sectional analysis showed the influence of attitudes on auto ownership; the longitudinal analysis showed separate effects of BE
Cao et al., 2007b	547 movers from Northern California, 2003	Structural equations model	Respective changes in driving, walking/ biking, and number of	Objective neighborhood characteristics; perceived	Various measures for residential preferences and travel attitudes	BE and SS. Attitudes influenced auto ownership and travel behavior; BE also had a separate effect.

			autos	neighborhood characteristics and their changes		
Handy et al., 2005	1,682 individuals from Northern California, 2003	Ordered probit model (linear regression ³)	Vehicle miles driven per week; changes in driving	Objective neighborhood characteristics; perceived neighborhood characteristics and their changes	Various measures for residential preferences and travel attitudes	BE and SS. The cross-sectional analysis showed the influence of attitudes on driving distance; the longitudinal analysis showed separate effects of BE.
Handy et al., 2006	1,682 individuals from Northern California, 2003	Ordered probit model (negative binomial regression ³)	Strolling frequency, walking to the store frequency; changes in walking and changes in biking	Objective neighborhood characteristics; perceived neighborhood characteristics and their changes	Various measures for residential preferences and travel attitudes	BE and SS. The cross-sectional analyses showed the influence of attitudes on walking behavior; the longitudinal analyses showed separate effects of BE on walking and biking behavior.
Krizek, 2000	549 households moving over the seven waves of the Puget Sound Transportation Panel, WA	Pairwise t-tests	Respective changes in trip distance, trip time, tour distance, tour time, trips per tour, and percentage of total trips by alternative modes	Changes in the Less Auto Development Urban Form (LADUF) index	Not available	BE. Individuals chose residential neighborhoods partially to match their travel preference; moving to a different residential environment had little influence on travel behavior given that only 9 out of 36 t-tests are significant.
Krizek, 2003a	6,144 individuals over the seven waves of the Puget Sound Transportation Panel, WA	Linear regression	Respective changes in vehicle miles traveled, person miles traveled, number of tours, and number of trips/tour	Respective changes in neighborhood accessibility and regional accessibility at the residence and workplace	Not available	BE. Changes in neighborhood accessibility and regional accessibility at the residence influenced most changes in travel behavior; regional accessibility at the workplace affected some changes in travel behavior.
McBeth, 1999	The central area of Toronto, 1993-1998	Descriptive analysis	Bicycle volume	Bicycle lane installations	Not available	BE. The installation of bicycle lane increased bicycle volume.
Meurs and Haaijer, 2001	189 movers and 524 nonmovers participating in the Dutch Time Use Study in	Linear regression	Respective changes in the number of trips by auto, bicycle, walking, transit,	Respective changes in home characteristics, street characteristics, and neighborhood characteristics	Not available	BE. Individuals' travel behavior was changed when moving to a different residential environment; nonmovers' travel behavior was also changed but

	1990 and in 1999		and all modes			not great when the environment was changed.
Painter, 1996	Three streets and a footpath, London	Descriptive analysis	Pedestrian volume after dark	Street light improvements	Not available	BE. Street light improvements increased pedestrian volume after dark.
Wells and Yang 2008	70 low-income women in Florida, Alabama, and Georgia collected from 2003-2006	Linear regression	Weekly walking steps	Neighborhood type, land use mix, street patterns, density	Not available	BE. The cross-sectional analysis did not show an association between neighborhood type and walking; the longitudinal analysis found changes in land use mix and street patterns had influences on walking.

Notes:

1. The limitations discussed in the text should be kept in mind while reviewing these conclusions.
 2. BE means evidence found for the influence of the built environment on travel behavior and SS means evidence found for the influence of residential self-selection on travel behavior.
 3. Both a statistical control approach and a longitudinal design were adopted in these studies. The modeling technique in parentheses was used for the statistical control approach.
- * These studies qualitatively explored which of BE and self-selection is more important. Among them, Cao (2008), Salon (2006), and Zhou and Kockelman (2008) quantitatively presented their relative contribution.

Table 2. Detecting the True Effect of the (Continuous-valued) Built Environment on Travel Behavior under the Assumption that Attitudes Affect Both BE and TB

Method	True effect on TB of increasing BE measure by one unit ¹	Proportion of total effect of BE on TB that is due to the BE alone rather than due to the effect of AT on BE
Statistical control	Coefficient of BE in eq. (5) for TB	Incremental contribution to R ² of BE (given AT and all other variables included), divided by incremental contribution to R ² of BE and AT entered together (given all other variables included)
Instrumental variables	Coefficient of \hat{BE} in eq. (6) for TB	Not possible to calculate
Selection models	Multiple possible effects, conditional and unconditional. Must separate out the effect of BE on RC* (or participation probabilities) from the effect of BE on TB _U , TB _S , and/or TB; the latter component is the “true” effect.	Proportion of total marginal effect of BE that is due to its direct effect on TB _U , TB _S , and/or TB (as opposed to its effect on RC* or participation probabilities)
Nested logit (NL)	¹ Elasticity of marginal probability of a given TB outcome, minus elasticity of the conditional probability of that outcome given fixed RC	Quantity to left, divided by elasticity of marginal prob. of a given TB outcome
Simultaneous jt. discrete choice	¹ Same as for NL	Same as for NL
Structural equations model	<i>Recursive models</i> : Total effect of BE on TB. <i>Nonrecursive models</i> : Difficult or impossible to isolate from the impact of changes in AT	<i>Recursive models</i> : Same as for the statistical control method. <i>Nonrecursive models</i> : No guidance from the literature
Longitudinal model	Coefficient of ΔBE in eq. (11) for ΔTB	Incremental contribution to R ² of ΔBE (given ΔAT and all other variables included), divided by incremental contribution to R ² of ΔBE and ΔAT entered together (given all other variables included)

¹ For the discrete choice models, the percentage effect, on the probability of a discrete TB outcome, of increasing BE by a percentage.

Table 3. Causal Inference Capability of the Different Quantitative Approaches

Approach	Association	Nonspuriousness	Time Precedence	Causal Mechanism
Statistical control	strong	strong	weak ¹	yes
Instrumental variables models	strong	moderate ²	weak ¹	yes
Sample selection models	strong	moderate ³	weak	yes
Joint simultaneous discrete choice models	strong	moderate ⁴	weak	yes
Nested logit models	strong	weak ⁴	weak	yes
Cross-sectional structural equations models	strong	strong	moderate ⁵	yes
Longitudinal models – single equation	strong	moderate ⁶	moderate ⁶	yes
Longitudinal models – structural equations	strong	very strong	very strong	yes

Notes:

1. The statistical control and instrumental variables approaches implicitly assume that attitudes are a cause rather than an effect of residential choice and travel behavior, which is open to debate. Further, with cross-sectional data, there is a temporal mismatch in that current attitudes are assumed to account for prior residential choices.
2. Reflects the limited ability of the IV approach to find instruments for BE that are both uncorrelated with ε and explain BE well enough to be useful.
3. Depending on how well RC is modeled.
4. The joint simultaneous discrete choice approach is an improvement over the nested logit approach in that it parameterizes the error terms of the RC and TB equations as functions of BE (and potentially other observed variables), thus reducing the potential for their correlations to be due to a spurious third-party variable.
5. Even with cross-sectional data, a structural equations model can provide some evidence for the direction of influence if attitudes are explicitly controlled for.
6. Strong if attitudes are measured at each wave; otherwise, a Δ BE that precedes a Δ TB may itself be preceded and caused by a Δ AT.