JTLU

Residential self-selection in quasi-experimental and natural experimental studies: An extended conceptualization of the relationship between the built environment and travel behavior

Eva Heinen
University of Leeds and
University of Cambridge
e.heinen@leeds.ac.uk

Roger Mackett
University College London
roger.mackett@ucl.ac.uk

Bert van Wee
Delft University of Technology
g.p.vanwee@tudelft.nl

David Ogilvie
University of Cambridge
dbo23@medschl.cam.ac.uk

Jenna Panter
University of Cambridge
jrp63@medschl.cam.ac.uk

Abstract: Despite a large body of research suggesting that the built environment influences individual travel behavior, uncertainty remains about the true nature, size, and strength of any causal relationships between the built environment and travel behavior. Residential self-selection, the phenomenon whereby individuals or households select a residential area based on their transport attitudes, is a frequently proposed alternative explanation for the reported associations. To resolve the issue of residential self-selection, longitudinal studies are often recommended. In this paper, we argue that intervention study designs are insufficient to fully resolve the problem and that intervention studies on the built environment and travel behavior may still be biased by residential self-selection. The aim of this paper is to extend existing conceptualizations of the relationships between the built environment, travel behavior, and attitudes and to provide suggestions for how a causal relationship between the built environment and travel behavior may be ascertained with more accurate estimates of effect sizes. We discuss the complexities of determining causal effects in intervention studies with participants who relocate, and the biases that may occur. We illustrate the complexities by presenting extended conceptualizations. Based on these conceptualizations, we provide considerations for future research. We suggest repeating analyses with and without individuals who relocated during the study, and with and without statistical controls for residential relocation. Additional quantitative and qualitative analyses will be necessary to obtain more accurate effect size estimates and a better understanding of the causal relationships.

1   Introduction

Over decades, transport policies have aimed to contribute to economically prosperous, attractive, healthful, and sustainable cities. In particular, the built environment has received much policy attention as a modifiable driver of travel behavior, particularly mode choice. Much scientific effort has been invested in determining this relationship, with most studies suggesting that aspects of the built environment influence individual travel behavior (e.g., Handy, Cao, & Mokhtarian, 2005; Ewing & Cervero, 2010; Handy, Boarnet, Ewing, & Killingsworth, 2002; Saelens, Sallis, & Frank, 2003). The main weakness of most studies is that they rely on cross-sectional data, which hampers the determination of causal relationships (e.g., van de Coevering, Maat, & van Wee, 2015; Mokhtarian & Cao, 2008; Handy et al., 2005).

Causality is an important concept in many fields, but there is no single universally accepted definition, and the notion of causality in human behavior is not always accepted (Naess, 2016; Parascandola & Weed, 2001). It is outside the scope of this paper to discuss all approaches and philosophies in detail. In the existing discussion of causality of the built environment on travel behavior, many studies take the probabilistic approach, and we therefore will follow Suppes' (1970) probabilistic theory of causality in this paper: "… one event is the cause of another if the appearance of the first event is followed with a high probability by the appearance of the second, and there is no third event that we can use to factor out the probability relationship between the first and the second events" (Suppes, 1970, p.10). From a more practical perspective, we follow Bradford Hill, who provided a list of viewpoints to be considered when aiming to determine causal relationships between one condition and another (Hill, 1965):

- *strength*: the larger the association, the more likely it is causal, although a small effect size does not exclude this possibility;
- *consistency*: the relationship is repeatedly observed;
- *specificity*: one factor specifically affects one outcome or group;
- *temporality*: the cause precedes the effect;
- *plausibility*: the relationship between cause and effect must make sense in the light of current theories and results; if not, further testing and hypothesizing are required;
- *(biological) gradient*: a greater dose of exposure should result in a larger effect; coherence: the finding should correspond with existing knowledge;
- *experiment*: (if possible) an experiment provides stronger evidence;
- *analogy*: a similar exposure may have a comparable effect, and therefore it is important to consider alternate explanations.

Some of these viewpoints, such as consistency, plausibility, and coherence, are widely taken into account in existing research on the relationship between the built environment and travel behavior. However, most studies do not use an experimental design, sufficiently describe temporal associations, or sufficiently consider alternative explanations. As a result, uncertainty remains about the true nature, size, and strength of any causal relationship between the built environment and travel behavior.

One particularly important potential alternative explanation for an observed relationship between the built environment and travel behavior is that of residential self-selection. This can be described as the phenomenon whereby individuals or households select their residential area based on their transport attitudes and preferences (hereafter referred to as attitudes) (Mokhtarian & Cao, 2008; Cao, Mokhtarian, & Handy, 2009a; Naess, 2009). For example, someone with positive attitudes towards travelling by public transport may choose to live in a location with good public transport access. In such a situation, attitudes are an antecedent to both travel behavior and the residential built environment to which an individual is exposed, which complicates the attribution of a given travel behavior to a particular built environment. In this paper, we limit our discussion of attitudes to those relating specifically to transport, which are of direct relevance to the debate about residential self-selection. However, our argument may

in principle be extended to other attitudes and preferences, particularly those relating to other characteristics of the built environment (van Wee, 2009).

To determine a causal effect of the built environment on travel behavior, it is important to accurately estimate effect sizes for the independent association of the built environment with travel behavior, following the strength consideration of Hill (1965). Accounting for residential self-selection enables one to improve the estimate of the effect and strengthen the case for causal inference. For this, the potential confounding by residential self-selection needs to be reduced and many authors recommend longitudinal studies (e.g., Boone-Heinonen, Guilkey, Evenson, & Gordon-Larsen, 2011; Cao et al., 2009a). These may explore the impact of changes in the built environment on changes in travel behavior in an observational panel study, or—in the spirit of Hill's (1965) "experimentation" viewpoint—by examining the impact of an intervention in the built environment on changes in travel behavior in a quasi- or natural experimental study[1] (McCormack & Shiell, 2011; van de Coevering et al., 2015). For the remainder of this paper, we will use the term "intervention studies" to refer to quasi-experimental and natural experimental studies as most of our argument applies to both research designs. Although intervention studies have the potential to offer stronger causal evidence, they may still be subject to bias and "steps to rule out competing explanations and biases" are recommended (McCormack & Shiell, 2011, p. 9).

In this paper, we argue that intervention studies on the built environment and travel behavior may still be biased by residential self-selection. This paper focusses on residential self-selection, but it should be noted that similar reasoning could be applied to other forms of self-selection, such as workplace self-selection, and that other forms of bias may also be present, such as attrition bias and selection bias. We focus on residential relocation, which—particularly if it takes place during a natural or quasi-experimental study—may introduce several threats to causal inference as it introduces the possibility of alternative causal structures. Uncertainty about the causal structure complicates the determination of the independent effect of the built environment on travel behavior as different structures may require different analytical approaches, thereby increasing uncertainty about the strength of the relationship, which is one of the aspects of causal inference emphasized by Hill (1965). Recently, more studies have been conducted on the effects of interventions in the built environment on travel behavior. Therefore, an extended conceptualization of residential self-selection applicable to intervention studies is needed.

The aim of this paper is to extend existing conceptualizations of the relationship between the built environment, travel behavior, and attitudes and to provide suggestions for how a causal relationship between the built environment and travel behavior may be ascertained with more accurate estimates of effect sizes. Although this is most important for conceptual and methodological development, it has practical importance for providing planners and decision makers with more robust evidence on causality. Knowledge of the magnitude of the effects of different built environments on travel behavior in different contexts contributes to more realistic expectations of the effects that can be achieved by different interventions (Naess, 2014b). This will assist in determining the economic and wider societal impacts of future intervention strategies.

2 Background

2.1 Overview of residential self-selection literature

Residential self-selection has received much scientific attention. Most papers have focused on travel behavior in general (e.g., Susilo, 2015; Cao, 2014, 2009; Scheiner, 2014; Wang & Lin, 2014; Chatman, 2014; Zhang, 2014; Cao & Ettema, 2014; Cao, Xu, & Fan, 2010; Scheiner, 2010; Bhat & Eluru, 2009; Pinjari, Bhat, & Hensher, 2009; Cao et al., 2009a; 2009b; Naess, 2009), with others examining

---

[1]It is important to note that whereas cross-sectional studies investigate the correlation between the built environment and travel behavior, intervention studies generally investigate the association between an intervention (a deliberate change in the built environment) and changes in travel behavior.

active travel in particular (e.g., Cao, 2015; Yu & Zhu, 2015; Schoner & Cao, 2014), physical activity (e.g., Baar, Romppel, Igel, Brähler, & Grande, 2015; Van Dyck, Cardon, Deforche, Owen, & De Bour-deaudhuij, 2011; Boone-Heinonen et al., 2010), or other issues including residential choice, transport emissions, and car or bicycle ownership (e.g., He & Zhang, 2014; Bhat, Paleti, Pendyala, Lorenzini, & Konduri, 2013; Hong & Shen, 2013; Biying, Zhang, & Fujiwara, 2012; Chen & Lin, 2011; Pinjari, Eluru, Bhat, Pendyala, & Spissu, 2008). It is beyond the scope of this paper to discuss all these studies in detail. Therefore, in this section, we focus on how the awareness of potential residential self-selection bias is handled. To date, most studies addressing attitudinal residential self-selection have used different (modelling) techniques (Mokhtarian & Cao, 2008; Cao et al., 2009a): direct questioning, statistical control, instrumental variable models, sample selection models, joint discrete choice models, structural equation models (SEM), mutually dependent discrete choice models, and longitudinal designs. Both of these studies recommend the use of longitudinal structural equation modelling with a control group. However, most existing studies are cross-sectional, and control for residential self-selection statistically. More recently, the importance of intervention studies over other, purely observational, longitudinal studies has been acknowledged.

The main difference between intervention studies and longitudinal studies without an intervention is, as the name suggests, that intervention studies involve one or more specific interventions whose effect is systematically determined. In contrast to other longitudinal studies, which can only determine the association between changes in the built environment and changes in travel behavior, intervention studies allow stronger causal inference as the environmental change can more clearly be seen to precede the change in travel behavior.

The importance of intervention studies is shown by a systematic review on the relationship between the built environment and physical activity—including walking and cycling—by McCormack and Shi-ell (2011). They evaluated 33 studies, of which 20 used statistical approaches to control for self-selection and 13 were intervention studies. They concluded that whereas the built environment was significantly associated with physical activity in cross-sectional studies using statistical control for residential self-selection, more rigorous intervention studies provided less support for this relationship and sometimes showed insignificant or even counterintuitive results. For example, proximity to public transport was found to be associated with physical activity in cross-sectional studies such as Chatman (2009), which concluded that access to a light, but not a heavy, rail service within 800 m of the home was associated with a higher frequency of walking or cycling. In a quasi-experimental study, however, a new rail stop neither resulted in differences in ridership nor was associated with physical activity (Brown & Werner, 2009). Similarly, the development of a light rail transit corridor was not associated with the realization of recommended levels of walking and physical activity by local residents (MacDonald, Stokes, Cohen, Kofner, & Ridgeway, 2010). Although this evidence is not conclusive, partly as the locations of these studies differ, this review demonstrated that studies whose designs supported stronger causal inference tended to show different results from those with more limitations in reducing self-selection bias.

2.2   Existing conceptualization of the relationship between the built environment, travel behavior, and attitudes for cross-sectional studies

The finding that intervention studies have yielded more mixed evidence than cross-sectional studies, even those that have controlled for residential self-selection, suggests the importance of how the relationship between the built environment, travel behavior, and attitudes is conceptualized.[2] Several pa-

---

[2]The differences in results between cross-sectional and intervention studies may have several explanations. One (and the most common) explanation is that changing the environment does not result in changes in behavior, despite the fact that cross-sectional studies suggest an association. Another is that cross-sectional studies often use large data sets, whereas intervention studies generally involve smaller sample sizes. Cross-sectional studies are therefore more prone to accept an alternative hypothesis and reject the null hypothesis, whilst the null hypothesis is correct in reality, i.e., a type-1 error. Intervention studies may not have enough power to show a statistically significant effect, although present in reality, a type-2 error.

pers on residential self-selection are (mostly) conceptual (e.g., Cao, 2015; van Wee & Boarnet, 2014; Chatman, 2014; Zhang, 2014; Naess, 2014a, 2014b, 2009; Chen & Lin, 2011; Bohte, Maat, & van Wee, 2009; Mokhtarian & Cao, 2008). Cao et al. (2009a) provides a systematic set of conceptualizations, reproduced in Figure 1.[3] In the first conceptualization (1.A), attitudes are an antecedent to travel behavior as well as the built environment. This implies that any measured association between the built environment and travel behavior is at least partly a result of sharing "a parent" (i.e., both are caused by a shared predictor (attitudes) and they are correlated as a consequence). The second conceptualization (1.B) shows that travel behavior influences attitudes, which, in turn, determine the residential location (i.e., built environment). The third conceptualization (1.C) illustrates the reverse scenario in which the built environment influences attitudes that in turn affect travel behavior. Only the fourth conceptualization (1.D) shows the built environment influencing travel behavior directly without attitudes affecting this relationship. We added three conceptualizations to the four presented by Cao et al. (2009a). In the fifth conceptualization, attitudes influence travel behavior, but all other relationships are not causal (1.E). In 1.F, attitudes are an antecedent of both the built environment and travel behavior. In addition, there is a causal relationship between the built environment and travel behavior. In the seventh and final conceptualization, attitudes influence the built environment, which consequently influences travel behavior (1.G). In theory, every line can be drawn four ways (correlation, causation in one direction, causation in the other direction, and bidirectional causations) resulting in a total of 81 (3⁴) possible conceptualizations for cross-sectional studies. Any conceptualization in which attitudes have a casual influence (either one way or bidirectional) on the built environment represents an instance of residential self-selection (in Figure 1 these are conceptualizations 1.A, 1.B, 1.F and 1.G). Any conceptualization in which attitudes are an antecedent (parent) of both the built environment as well as travel behavior is vulnerable to residential self-selection bias in ascertaining the relationship between the latter two variables. In Figure 1, this is visualized in conceptualizations 1.A and 1.F, the latter representing the "classical example of residential self-selection".
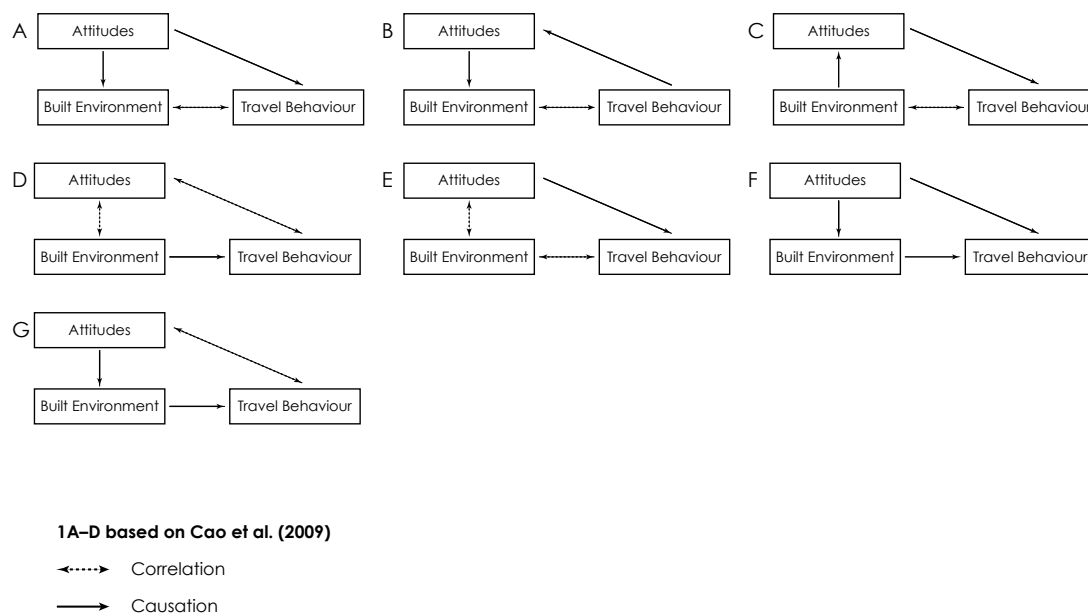

1A–D based on Cao et al. (2009)

◄┄┄► Correlation

───► Causation

Figure 1: Conceptualization of the relationship between the built environment, travel behavior and attitudes for cross-sectional studies

---

[3]Some meanings differ depending on the relationship discussed.

2.3   Reflections on existing conceptualizations and their implication

Existing continued uncertainty regarding the structure of these causal relationships complicates the determination of causal effects, as different causal structures require different statistical approaches to determine true estimates of effect sizes. For example, whether attitudes should be controlled for in cross-sectional studies by being included as covariates in multivariable models depends on the true relationships among attitudes, the built environment, and travel behavior. If attitudes are "only" a competing predictor of travel behavior (See Figure 2, conceptualization A), adding them to statistical models is recommended because this would increase the predictive power of the model and would not change the effect size estimate for the built environment.[4] However, if attitudes (partly) mediate the effect of the built environment on travel behavior (See Figure 2, conceptualization B and C), although adjustment for attitudes would likewise increase the predictive power of the model, it would also introduce bias. The effect size estimate for the built environment would be limited to the direct effect (which is zero in 2.C), and would not include the indirect effect now captured by the estimated effect of attitudes. In this situation, we do not recommend statistical adjustment for the effect of attitudes in a regression analysis or similar, as one should not control for a mediator if one aims to determine the "true" effect of one variable on another (Shrier & Platt, 2008). In the classic case of residential self-selection bias shown in Figure 1.F, attitudes influence travel behavior, (partly) mediated by the built environment. In this case, controlling for attitudes may reveal the independent effect of the built environment on travel behavior in addition to being a mediator of the relationship between attitudes and travel behavior.[5] Although indirect effects may be captured in the SEM, data collection should be aligned with the conceptualization (i.e., the cause should precede the mediator), which, in turn, should precede the outcome, for which longitudinal data are required.

---

[4]In the case that attitudes and travel behavior are correlated, not including attitudes will change the effect size of the built environment. The exclusion of attitudes would result in endogeneity bias.

[5]The exact modelling depends on the research question. If the built environment is only a mediator, it should not be included in the model at all. A better approach to find the effect of the built environment on travel behavior may be conducting an intervention study (see Section 3 onward).
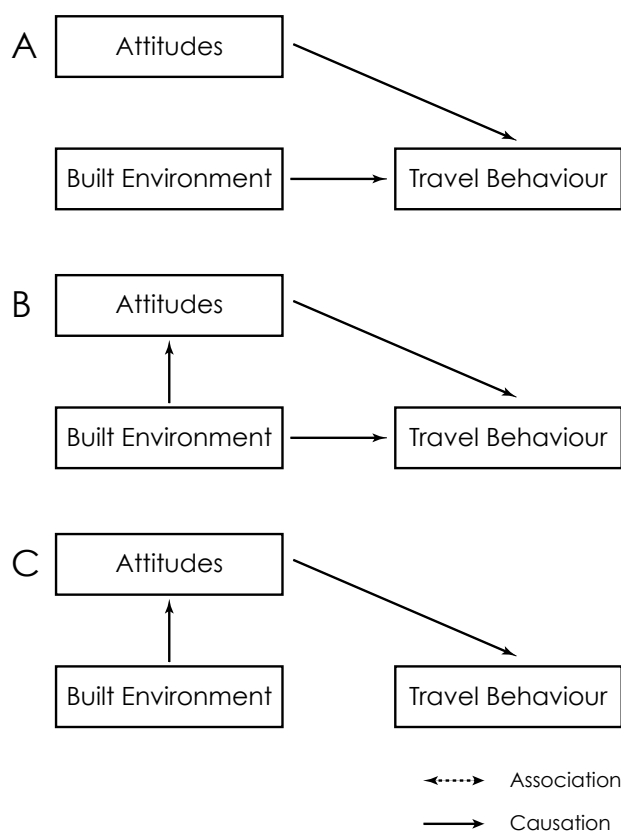
Figure 2: Visualization of attitudes as a competing exposure and as a mediator

3 Intervention studies

Many authors have come to the conclusion that longitudinal data are essential to determine causal relationships and "true" effect size estimates (e.g., Boone-Heinonen et al., 2011; Cao et al., 2009a) and we agree with this conclusion. However, in the absence of randomized controlled trials, which in the case of the built environment may be impractical or unethical—for example, we cannot randomly assign individuals to a residential area—the strongest evidence for causal inference will come from quasi- and natural experimental studies (Craig et al., 2012). The main difference between these types of study is that in quasi-experimental studies, the researcher is in control of the intervention, whereas in natural experimental studies, they may not. For example, in a natural experimental study, a municipal authority may decide to construct new infrastructure and will be in control of this intervention, whereas an independent researcher will have control only of the study design and data collection, and not of the intervention itself. Quasi-experimental studies, on the other hand, may be very similar to real experiments except for the lack of one essential element: the random assignment of individuals to treatment condition. Although most of our arguments apply to both research designs, in this section, we focus on natural experimental studies as these are currently more commonly encountered in this field.

3.1   Determining causal effects in natural experimental studies

There are multiple ways to strengthen causal inference in attributing outcomes to interventions in natural experimental studies (Craig et al., 2012). These include (Craig et al., 2012):

1. Matching: comparing exposed and non-exposed groups matched on important characteristics;
2. Regression adjustment: adjusting for measured differences in characteristics between those who did and did not receive an intervention in multivariable regression models;
3. Propensity scores: using the likelihood of being exposed to an intervention given a set of covariates. This can be used to match individuals/observations, to stratify individuals/observations, to weight for probability of treatment, or as a covariate in analysis.

Measured characteristics often vary between intervention and control groups in natural experimental studies, and matching is therefore often not a feasible option. The use of propensity scores and matching individuals with different propensities still requires the covariates to be balanced across treatment and comparison groups. This often requires large sample sizes, which may not always be feasible either. We therefore continue our argument on the assumption that most studies are likely to use regression adjustment.

The effects of interventions in the built environment are usually multi-faceted. In the short term, individuals may change their travel choices in respect of certain destinations, the frequency of travel, and the mode and route selected. In the medium and long term, residential locations may change, as well as workplace locations and the choice of destinations for other activities. We focus on changes in mode choice, because this is often the most important change in travel behavior from a policy perspective.

3.2   Timeline of intervention studies

Figure 3 shows a timeline of a typical intervention study. It begins with a "baseline" situation (A). At a certain moment in time, which is often hard to pinpoint, ideas of an intervention begin to develop.[6] For the purposes of illustration, we focus on an intervention that consists of the construction of new physical transport infrastructure. The period between initial planning and the actual decision to build is period B. In natural experimental studies, the researcher is not in control of the intervention, or even (usually) of the initiation of discussions concerning a possible intervention. Thus, the collection of research data most likely begins after the decision to build, indicated by period C. Data collection before the decision is possible, but unlikely in situations where this would entail significant cost and there is significant risk of the intervention ultimately not being implemented. Period D represents the interval between the start of the study (i.e., collection of "pre-intervention" data) and the actual start of construction, whereas period E represents the actual period of construction. It is possible that one or more intermediate waves of data collection may take place during this period. At the end of the construction period, the new infrastructure opens, and after some further interval (period F), the post-intervention data are collected. Period G denotes the period after the final data have been collected. It is possible that two or more data collections take place after the intervention.

---

[6] It should be noted that an intervention may not be random, and is also not necessarily exogenous. In other words, the location of the intervention may be a result of the existing local built environment or residents, or even invoked by its residents. The effect of the intervention may in the latter case be larger than in the entire population. However, for clarity of the argument, the paper assumes that the intervention location is as-if random.
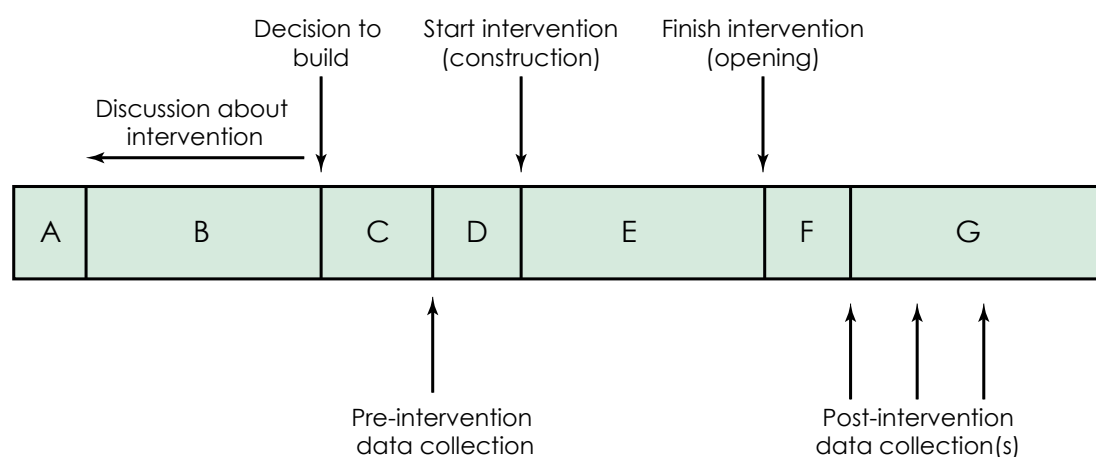
Figure 3: Timeline of an intervention study

4   Residential self-selection in intervention studies

As explained above, intervention studies are often proposed as a solution to the problems of causal infer-
ence in general and residential self-selection bias in particular. In this section, we argue that residential
self-selection may still introduce bias in intervention studies. Residential relocation of participants is par-
ticularly problematic, but other changes, such as destination choices, can also introduce bias. We limit
our discussion to residential relocation. We discuss this in intervention studies, followed by an examina-
tion of the relationship between the built environment, travel behavior, and attitudes in these studies,
before presenting an extended conceptualization. Residential relocation, as well as other changes, may
also be a (desired) effect of an intervention. However, if we wish to determine the effect of an interven-
tion in the built environment on travel behavior, residential relocation of participants may introduce
bias, especially if moving is also an effect of the intervention.

4.1   Residential relocation during intervention studies

Residential relocation may introduce bias in effect size estimates of the relationship between the built
environment and travel behavior and thereby affect the causal inference from a study. Residential reloca-
tion may occur at any moment in the timeline of a study (Figure 2). The most obvious "problematic"
periods from a researcher's point of view are periods D, E, and F, as these occur between the pre- and
post-intervention data collections. These relocations produce two complications for researchers:

1. Relocation in these periods complicates the assignment of a correct measure of exposure to the
   intervention to individuals. The level of exposure or whether an individual is exposed to an
   intervention or not should be determined before the intervention to meet the requirement of
   temporality. Moving during the study complicates the assignment of an individual to an inter-
   vention or control group, as the individual may live in the exposed area for part of the study and
   in the control area for the rest. If a continuous measure of exposure is used, relocation compli-
   cates the determination of the "correct" level of exposure.
2. Although residential relocation may occur for various reasons, including change in employment
   or household structure, it may also be (partly) driven by attitudes whilst being aware of the likely
   post-intervention situation. This is a form of residential self-selection and may introduce bias.
   In this paper, we mainly focus on the second complication, but both complicate the determination

of the independent effects of changes in the built environment. Even moving before the intervention may be problematic because residential self-selection may still occur in period C and, to a lesser extent, period B. Some people may relocate in anticipation of the planned or potential changes, despite the uncertainty about whether they will take place or not. These individuals do not move home during the period of data collection, and thus a "valid" measure of exposure can be determined for them. However, they may be more inclined than others to change their travel behavior in response to the intervention because their residential location may have been influenced by the extent to which their attitudes are congruent with the anticipated built environment after the intervention. Moving in periods A or G does not direct threaten the validity of a causal estimate, although moving in period G may indicate a time-lagged intervention effect.

4.2   Moving in intervention studies and the relationships between the built environment, attitudes, and travel behavior: An extended conceptualization

This section discusses and extends the conceptualization of the relationships between the built environment, travel behavior, and attitudes in relation to residential relocation in intervention studies.

Figures 4 and 5 present the extended conceptualizations following from these observations. These extend the work of Cao et al. (2009a) and others, and have four central concepts: the outcome (labelled "change in travel behavior"), the exposure (labelled "intervention: change in the built environment"), the alternative explanation of a possible effect (competing exposure) (labelled "moving"), and "attitudes" as well as "change in attitudes." These conceptualizations may be extended even further, but are limited in this paper to how attitudes may affect causal inference in intervention studies. Our thinking is influenced by causal diagrams and directed acyclic graphs (DAGs) (see e.g., Shrier & Platt, 2008; Greenland, Pearl, & Robins, 1999). DAGs can be used to provide a compact graphical abstraction of the relationships one tries to investigate, including potential confounders and competing exposures. The graphical representation of causal effects between variables may provide insights into the potential reduction of or increase in bias with the inclusion of additional variables in an analysis. An arrow connecting two variables indicates causation. Causal diagrams are complete insofar as all common causes of the outcome and intervention should be included. However, they are incomplete in that it may not include every single competing exposure, or mediator (Scheines, 1997). This approach relates to the definition of causality of Suppes (1970), which we follow. This simplification highlights the key variables and their roles, which in turn, helps informs decisions about how they might be handled in statistical analyses to produce the most likely "true" effect sizes for the relationship of interest.

Non-movers
In intervention studies, the cause (an intervention in the built environment) precedes the anticipated effect (a change in travel behavior) and therefore fulfils the requirement of temporality.[7] Intervention studies are also designed (or, at least, analyzed) as an experiment in which one or more characteristics of an exposure are purposefully altered to test the effect on the outcome. Potential bias due to residential self-selection is reduced because attitudes are not an antecedent (parent) to both the travel behavior and the built environment. The effect of the intervention on changes in travel behavior can therefore be determined relatively easily for non-movers (Figure 4).

Two relationships are still uncertain in these conceptualizations. One is the position of attitudes. It is conceivable that the intervention drives a change in attitudes, which consequently results in a change in travel behavior (conceptualization 4.A). Alternatively, a change in travel behavior as a result of the intervention may consequently lead to a change in attitudes (conceptualization 4.B). As suggested

---

[7] That is assuming that the anticipation of the intervention does not result in a change in travel behaviour before the intervention is introduced.

by Naess (2014) and others, bi-directional relationships are also possible. As such, conceptualizations 4.C and 4.D visualize such bidirectional relationships between attitudes and travel behavior. It seems inconceivable that such bi-directional relationships may exist between other pairs of variables (without being noticeable), namely between changes in the built environment and changes in travel behavior or between changes in the built environment and changes in attitudes; these are therefore not included in the conceptualization.

These conceptualizations assume that the intervention is not targeted at specific areas or individuals. However, some interventions may be undertaken in areas in which greater effects are anticipated (sometimes referred to as "selection bias" or "confounding by indication"). For example, a new railway station is more likely to be opened in an area where existing access to railway stations is low and a large uptake of train use is anticipated.

All conceptualizations for non-movers allow one to determine the intervention effect of a change in the built environment on travel behavior in a similar way. The differences shown in the conceptualization will mainly complicate one important aspect in intervention studies: the timing of the post-intervention data collection(s). Especially for conceptualization 4.C and 4.D, but perhaps also for conceptualizations 4.A and 4.B, the intervention effect may increase over time, and selecting the most optimal moment(s) to collect data for the evaluation is challenging.
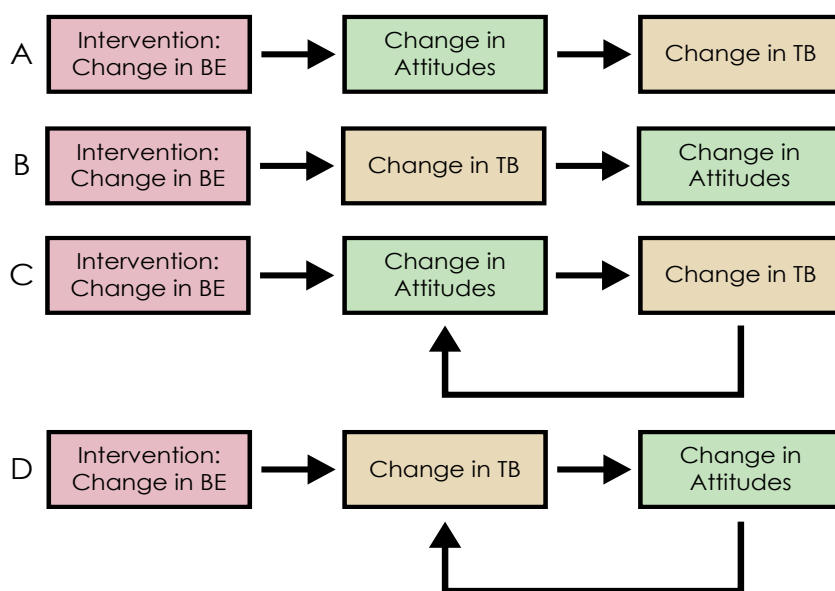

Figure 4: Conceptualization of the relationship between the built environment, travel behavior, and attitudes for non-movers in quasi- and natural experimental studies*

*BE: built environment. TB: travel behavior

Movers

For those who have relocated, the potential conceptualizations are more numerous because relocation creates an opportunity to move to an area matching the mover's attitudes whilst considering the anticipated or actual effects of the intervention. Both the move itself, as well as the changes in the built environment as a consequence of this relocation, may influence a change in travel behavior [8].

There are two possible relationships between attitudes and residential relocation. The first is that neither the decision to relocate nor the choice of a new residential area is driven by attitudes. This implies that the new residential location is chosen independent of personal and household attitudes and the relocation is driven by other reasons (e.g., change in household composition, house prices, schools, etc.). The second is that the relocation is at least partly driven by attitudes in that individuals decide to relocate to a new location that more closely matches their attitudes. If this relocation occurs at least partly because of the intervention, this is a form of residential self-selection during the study. These relationships and consideration will inform the extended conceptualizations (Figure 5). In the first row, the conceptualizations 4.C and 4.D are extended. In the following rows, different relationships are conceptualized between the constructs. As a result, Figure 5 shows a similar order of constructs in each column, and a similar arrangement of arrows in each row.

Similar to the conceptualizations of the non-movers, the cause (an intervention in the built environment) precedes the anticipated effect (a change in travel behavior) and therefore fulfils the requirement of temporality. Similar too are the variations on how the intervention may affect travel behavior, with the possibility of a bidirectional relationship between a change in attitudes and a change in travel behavior (all conceptualizations in Figure 5, except the conceptualizations in row B (i.e., 5.B1–B4)).

In row A, the conceptualizations capture the idea that attitudes drive the decision to relocate, and the relocation, in turn, influences either travel behavior (conceptualizations 5.A1 and 5.A3) or a change in attitudes (conceptualizations 5.A2 and 5.A4).

In row B of the conceptualizations, the idea presented is that there are no bidirectional effects between changes in attitudes and changes in travel behavior (i.e., similar to conceptualizations 4.A and 4.B).

The conceptualizations in row C capture another causal structure: that attitudes influence the choice to move, but the new environment does not result in changes in travel behavior (conceptualization 5.C1 and 5.C3) or changes in attitudes (conceptualizations 5.C2 and 5.C4). For example, someone who holds positive attitudes towards cycling may move to a location with easy access to good cycling facilities, but may not change their mode choice either because they were already travelling by bicycle and continue to do so (but, now, more easily), or because they continue to use other modes of transport because they still do not consider cycling the best option.

In the conceptualizations in row D, attitudes do not influence the selection of the new residential location. Moving is still a competing exposure, but it did not originate from the attitudes of the movers.

The final possibilities (rows E and F) demonstrate attitudes as an interaction effect (effect modifier or moderator). In row E, attitudes are a moderator of the effect of moving on travel behavior (conceptualizations 5.E1 and 5.E3). In other words, attitudes may strengthen or attenuate the effect on travel behavior of the residential relocation. This relates to the concept of consonance and dissonance, i.e. the extent to which the built environment corresponds with one's attitudes (Festinger, 1957; Schwanen & Mokhtarian, 2004, 2005a, 2005b; Kamruzzaman, Baker, & Turell, 2015; Kamruzzaman, Baker, Washington, & Turell, 2013a, 3013b; Manaugh & El-Gneidy, 2015). The cited studies suggest that individuals whose attitudes correspond with their residential built environment tend to have different travel patterns from those who are mismatched. It is also possible that attitudes moderate the effect of the "intervention" in the built environment on travel behavior (conceptualizations 5.F1 and 5.F4) through the same phenomena of consonance and dissonance.

Two other considerations should be borne in mind. First, a change in the built environment due to residential relocation may or may not result in a change in travel behavior. If moving is driven by attitudes (regarding transport), this can be seen as residential self-selection. However, there is only a po-

---

[8]Both moving itself and the consequent change in the BE are denoted in the conceptualizations by "moving". We have chosen to do so for simplicity, as it is less relevant whether it is the move itself or the changes in BE that may result in a change in TB, it is the fact that it is a competing exposure, potentially driven by attitudes.

tential residential self-selection bias in a study that aims to determine the effect of the built environment on travel behavior, if this relocation also affects travel behavior. Second, the effects of relocation on travel behavior may not be restricted to those evoked by changes in the built environment. Moving, as a major life event in its own right or as a proxy of another life event such as childbirth, may elicit a reconsideration of travel options and changes in travel behavior.
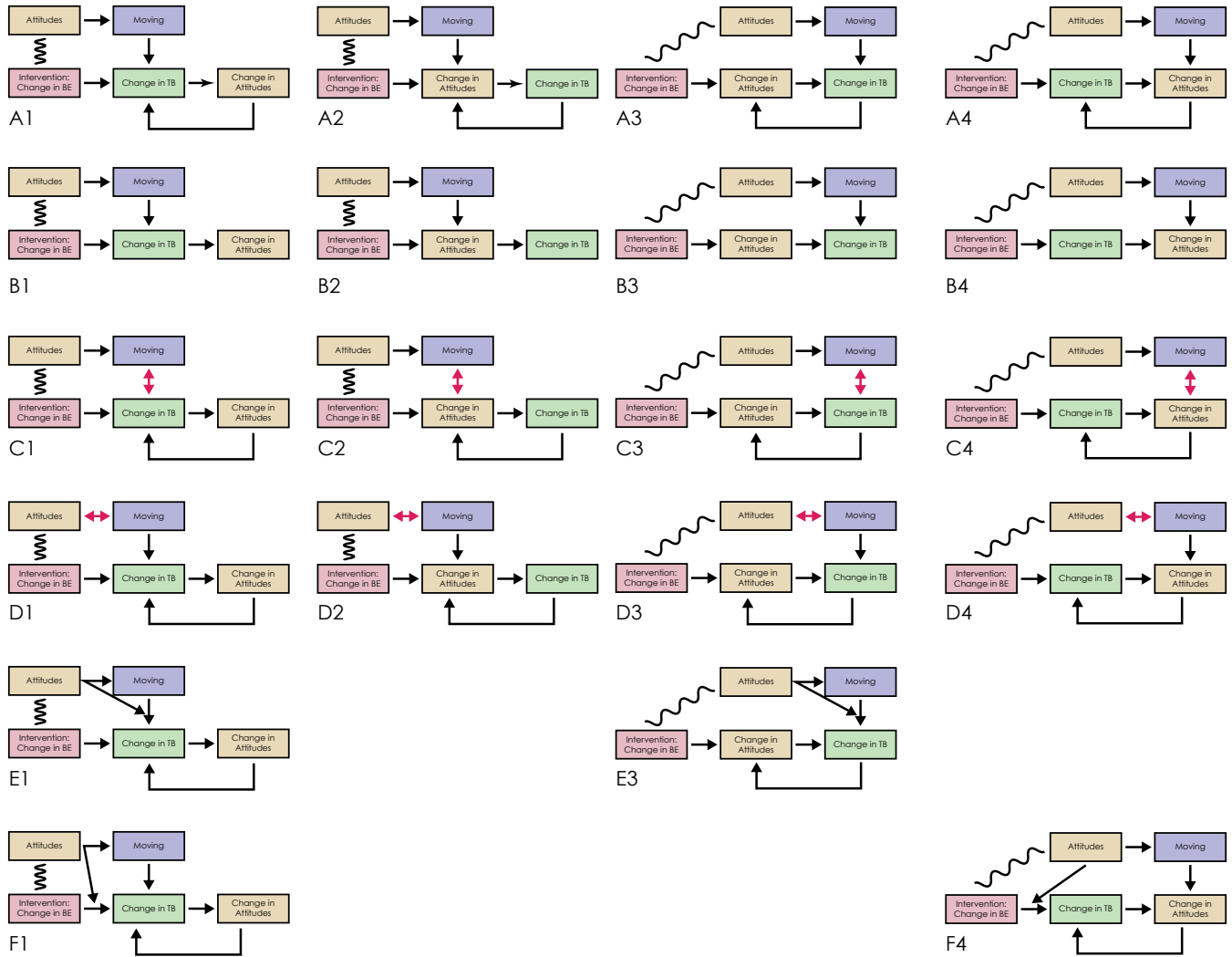

Figure 5: Conceptualization of the relationship between the built environment, travel behavior, and attitudes for movers in quasi- and natural experimental studies*

*BE: built environment. TB: travel behavior. A black single-headed arrow represents causation; a red double-headed arrow represents correlation (at most). The sinusoidal line indicates that the attitudes that drive the move are related to the intervention (i.e., if attitudes drive the relocation, this is to be closer to or farther away from the intervention).

4.3   Reflections on the extended conceptualizations and their implications

These conceptualizations depict different possible causal pathways, and there may be no single perfect conceptualization/causal pathway that applies for everyone. It is important to be aware that the concepts "attitude" and "change in attitudes" are not independent and are, at a minimum, statistically coupled.

Thus, our discussion below on attitudes reflects both the attitudes at baseline ("attitudes" in Figure 5), as well as "change in attitudes". In our example, this only applies to participants who moved during the study, as the intervention effect is relatively easy to determine for non-movers.

Attitudes have different roles in the different conceptualizations. Attitudes may act as an antecedent to a competing exposure (moving) (all conceptualizations in Figure 5 in rows A, B, E and F), a mediator of the intervention effect (conceptualizations 5.A2, 5.A3, 5.B2, 5.B3, 5.C2, 5.C3, 5.D2, 5.D3, and 5.E3 and via the bidirectional relationship also 5.A1, 5.A4, 5.C1, 5.C4, 5. D1, 5.D4, 5.E1, 5.F1, and 5.F4), a moderator of the intervention effect (conceptualizations 5.F1 and 5.F4), a moderator of a competing exposure (conceptualizations 5.E1 and 5.E3), or a combination of these functions. The uncertainty surrounding the causal structure complicates the determination of causal effects, as different causal structures may require different statistical approaches.

All conceptualizations in rows A, B, C, E, and F are vulnerable to residential self-selection, and in conceptualizations in rows A, B, E, and F, this potential residential self-selection may result in a residential self-selection bias. The latter issue does not arise if row C represents the correct conceptualization, because the move does not result in a change in travel behavior due to assumptions underlying the structures in that row.

5   Considerations and suggestions for future research

Longitudinal studies are often proposed to be an effective method of limiting potential bias from residential self-selection. However, as the conceptualizations in Figure 5 show, residential self-selection may still be present in intervention studies. There may be more than one "true conceptualization" (e.g., different groups of people may be more aligned with different specific conceptualizations), and as such, methods such as latent class SEM that aim to disentangle membership of such groups may be useful. We need to make it good practice to estimate different causal structures, and among the best ones, analyze the varying conclusions they lead to with respect to the questions of interest.

Nonetheless, an acknowledgement of the different roles of each concept influences how we should model the relationship between an intervention in the built environment and a change in travel behavior. The uncertainty about the causal structure complicates the determination of causal effects, as the different causal structures require different statistical approaches (see e.g., Shrier & Platt, 2008; Tu & Gilthorpe, 2012)

- We wish to control for competing exposures (or proxies of competing exposures). Controlling for competing exposure is done for two main reasons. First, by controlling for a competing exposure the overall variance explained will increase. Second, and more importantly given the aim of the study we are describing, it removes a potential bias in the effect size estimate. This bias is present when a competing exposure is not equally distributed over the population compared to the intervention exposure.
- We do not wish to control for mediators. Controlling for mediators introduces bias, as one will likely underestimate the effect of the intervention on the outcome as part of the variance will be explained by the mediator.
- In order to improve analyses of the impacts of interventions on travel behavior, we suggest the following considerations.

Research design:
- Planning of the data collection to reduce the possibility of bias
- Collect data on residential relocation before the intervention;
- Collect additional data (e.g. repeated cross-sectional recruitment, panel data, additional recruitment during later waves of data collection, and additional qualitative studies);
- Disentangle the causal relationships further using qualitative methods.

Statistical analysis:
- Repeat analyses with and without the inclusion of movers;
- Repeat analyses by controlling and not controlling for residential relocation.

We elaborate on these below. Although some of these suggestions may seem obvious, they are not widely and carefully considered in most current intervention studies.

5.1 Research design considerations

The considerations for research design are intended to obtain a better understanding of which conceptualization is most likely to correspond with the true causal structure(s), as well as to provide data that permit tests to understand the extent to which residential self-selection may be present in a natural experimental study.

5.1.1 Planning of the data collection

The moment of residential relocation has implications for its potential impact on the study. Moving between the two waves of data collection (periods D, E, and F in Figure 2) directly threatens the validity of the effect size estimates and the possibility of determining a causal relationship. Minimizing the durations of periods C, D, and F may reduce the number of participants who relocate during the study and thereby limit this threat. We recommend keeping period D as brief as possible. Period C would ideally also be as short as possible to ensure that pre-intervention data collection truly represents a baseline and does not capture any anticipatory behavior change. However, the optimal length of period F is more difficult to pinpoint, as a more condensed period F also reduces the opportunity for participants to change their behavior and consequently for the detection of an intervention effect, which may especially be the case if there is a bidirectional effect between a change in attitudes and change in travel behavior. Therefore, multiple post-intervention measures are optimal, despite greater costs in time and money.

5.1.2 Collect data on residential relocation before the intervention

It is important to note that moving may be an effect of the intervention itself. However, particularly if moving is a result of the intervention and is driven by transport attitudes, the uncertainty of the true causal relationships and the consequent modelling become substantially more complicated.

Although moving outside the period of data collection may, at first glance, appear to have little impact, it may still affect the results. For example, moving in periods B or C may be a result of residential self-selection as individuals may act in anticipation of the intervention. If such relocation is (partly) driven by attitudes (such as in all the conceptualizations except those in row D in Figure 5), then those who have recently relocated to the intervention area may be more likely to adjust their travel behavior as a result of the intervention than existing residents. This may pose threats to the accuracy of the effect size estimates. Thus, we suggest collecting information on the date when participants moved to their current residence and when they made the decision to move, as well as the location of their former residence and its built environment characteristics [9], as some relocations may have been made over a only short distance for practical reasons (e.g., change in family size), whilst remaining in the same neighborhood. This would allow researchers to compare residents who have recently moved into the area with longer-term residents controlling for potential confounding factors. We also recommend collecting data on their former residence and its built environment characteristics for those that move during the study. This would allow us to conduct the statistical analyses proposed in Section 5.2.2.

---

[9] These could also be derived by geographic information system (GIS).

5.1.3   Collect additional data

Intervention studies generally recruit participants before the intervention is completed, and individuals are often assigned to exposed/intervention and control groups. The surface of exposed areas is often smaller than the surface of the control areas, which may also correspond with a larger population in the control area. The recruitment strategy may specifically target the exposed area to secure a large enough sample size, but moving may cause problems. Even if we assume that moving occurs at similar rates in the intervention and control areas (this is not a given because the housing tenure may differ in the areas, and it also assumes the construction of the intervention is not too burdensome), given the differences in the size of the areas it is more likely that a given participant—independent of where they lived at the time of recruitment—will relocate into the control area than into the exposed area.

Consequently, the exposed group is more likely to "lose members" over time and cohort (panel) data will capture moving into the exposed area to only a limited extent. This presents an interesting situation. The intervention is likely to attract "lovers" (individuals whose attitudes are congruent with the intervention) and repel "haters" (individuals whose attitudes do not correspond with the intervention). In a cohort without additional recruitment after the pre-intervention data collection, and with a larger control area than an exposed area, the sample will mostly include "lovers" who moved in periods B and C into the exposed area, and "haters" who moved to the control area in periods B–F. Therefore, "haters" are more likely to be included in the population than "lovers" and this may consequently result in an underestimation of the effects of the intervention.

We therefore suggest additional data collection to ascertain whether "lovers" are included in studies corresponding with their prevalence in the population. One option is a repeated cross-sectional data collection to monitor whether the average characteristics and behavior in the intervention area (and possibly the control area) change over time. A second option is additional recruitment, especially in the exposed area, during later waves of data collection. This may allow a case-control study comparing individuals who have relocated into the exposed area with those who lived in the area before the intervention to determine whether movers into the area are more likely to be "lovers" than non-movers are. A third option is the use of qualitative research to understand the behavior of "lovers" and "haters" of the intervention.

5.1.4   Disentangle the causal relationships further using qualitative methods

Whereas the determination of correct effect size estimates requires a quasi- or natural experimental quantitative approach, complementary qualitative research can help to further disentangle the relationships between moving, attitudes, the built environment, and travel behavior. This will contribute to greater certainty about the nature of the relationships, which, in turn, will help in selecting the appropriate statistical approach matching the conceptualization. It may also show other relationships between the constructs, which may inform future quantitative research.

A qualitative study of movers (see 5.1.3.) may also be useful in disentangling the causal relationships, more fully capturing the intervention effects including medium- and long-term impacts, and informing future research designs and statistical analyses.

5.2   Statistical considerations

The statistical considerations are aimed at improving estimates of the effect sizes of the built environment and understanding how to handle potential bias from residential self-selection whilst not introducing additional biases.

5.2.1 Repeat analyses with and without the inclusion of movers

At first sight, an apparently simple solution may be to exclude individuals who relocated during the study, perform the analyses on non-movers alone, eliminate all conceptualizations in Figure 5, and work with the conceptualizations in Figure 4. Indeed, the exclusion of individuals who have relocated removes most potential residential self-selection. However, this approach ignores the fact that residential relocation is generally not equally distributed over the (study) population. Relocation is restricted by micro-level restrictions (such as income) (van Ham & Feijten, 2008) and may be more prevalent among younger people and those without a permanent job. This implies that excluding movers may reduce the likelihood of certain groups being included in the study, and therefore introduce a form of selection bias and a subtle change in the research question the study is addressing. It follows that including and excluding individuals who have relocated during the study may result in different inferences, with neither providing an accurate estimate of the true effect size for the intervention. We therefore suggest repeating analyses with and without individuals who relocated during the study. It is important to note that neither model is necessarily superior—as one may suffer from a residential self-selection bias, and the other from a recruitment bias—but they are complementary and together may strengthen the case for causal inference. If the two estimates are very different and the confidence intervals do not overlap, it is likely that either residential self-selection or another form of selection bias is present.

5.2.2 Repeat analyses with and without controlling for residential relocation

Residential relocation may be a competing exposure (potential cause) for the outcome of a change in travel behavior (see all conceptualizations in Figure 5 in rows A, B, D, E, and F). As such, when movers are retained in the analysis, we recommend statistical adjustment for residential relocation by adding a covariate that either indicates whether a person has relocated or serves as a proxy for this (such as a variable that capture changes in the built environment consequent on relocation).

However, not all conceptualizations support adding a variable representing moving into statistical models. Conceptualizations 5.A1–4, 5.B2–3, 5.E1, 5.E3, 5.F1, and 5.F4 illustrate that attitudes may, in addition to being a predictor for moving (i.e. a competing exposure), be a mediator of the effect of the intervention on a change in travel behavior (either directly or as a bidirectional relationship (feedback loop) with change in travel behavior). As attitudes are a parent of moving, moving could be seen as a proxy of attitudes. Therefore, in the case that these conceptualizations are assumed true, controlling for moving may confound the relationship between the intervention and changes in travel behavior, and bias effect size estimates.

Thus, we suggest estimating models with and without adjustment for moving, on a similar basis to the argument for estimating models with and without the inclusion of participants who have moved, as neither model is necessarily superior.

The adjustment could be performed by introducing a dummy variable (having moved/not having moved), but more complex measures could be explored including a variable that not only captures the move but also the moment of moving (based on the data collection as suggestion in 5.1.2), or by including a measure of change in exposure that the move resulted in. The dummy variable option is the easiest to explore. The latter option, depending on the way in which the level of exposure to an intervention is measured, is a more complex, but perhaps more fruitful way forward. This may (partly) address both complications of moving in intervention studies (see Section 4.1) simultaneously.

6   Conclusion

This paper has argued that intervention studies alone are not sufficient to control for the potential bias of residential self-selection in estimating the travel behavior impacts of changes in the built environment. By extending the existing conceptualization, we have shown how residential self-selection may still pose threats in quasi- and natural experimental studies. The conceptualizations presented may help in clarifying the causal relationships involved and selecting the most appropriate statistical methods for analysis.

We acknowledge that residential self-selection constitutes only one of the biases that may need to be addressed, and reducing attrition bias and selection bias are at least as important. Nevertheless, to reduce self-selection bias, we suggest repeating analyses with and without the inclusion of individuals who have relocated during the study, and with and without statistical control for relocation. Additional quantitative and qualitative data may be necessary to obtain more accurate effect size estimates and understanding of the causal structures. Some of these suggestions may seem trivial, but a more careful consideration of the potential conceptualizations is essential to improve our understanding of the determinants of travel behavior. This will ultimately contribute to more accurate estimates of the quantifiable effects of interventions in the built environment, and thereby to more robust planning and policymaking.

Acknowledgements

References

Baar J., Romppel, M., Igel, U., Brähler, E., & Grande, G. (2015). The independent relations of both residential self-selection and the environment to physical activity. *International Journal of Environmental Health Research, 25*(3), 288–298.

Bhat, C. R., & Eluru, N. (2009). A copula-based approach to accommodate residential self-selection effects in travel behavior modelling. *Transportation Research Part B: Methodological, 43*(7), 749–765.

Bhat, C. R., Paleti, R., Pendyala, R., Lorenzini, K., & Konduri, K. (2013). Accommodating immigration status and self-selection effects in a joint model of household auto ownership and residential location choice. *Transportation Research Record, 2382,* 142–150.

Biying, Y., Zhang, J., & Fujiwara, A. (2012), Analysis of the residential location choice and household energy consumption behavior by incorporating multiple self-selection effects. *Energy Policy, 46,* 319–334.

Bohte, W., Maat, K., & van Wee, B. (2009). Measuring attitudes in research on residential self-selection and travel behavior: A review of theories and empirical research. *Transport Reviews, 29*(3), 325–357.

Boone-Heinonen, J., Guilkey, D. K., Evenson, K. R., & Gordon-Larsen, P. (2010). Residential self-selection bias in the estimation of built environment effects on physical activity between adolescence and young adulthood. *International Journal of Behavioral Nutrition and Physical Activity, 7,* 70.

Brown B. B., & Werner, C. M. (2011). The residents' benefits and concerns before and after a new rail stop: Do residents get what they expect? *Environment & Behavior, 43,* 789–806.

Cao, X. J. (2015). Examining the impacts of neighborhood design and residential self-selection on active travel: A methodological assessment. *Urban Geography, 36*(2), 236–255.

Cao, X. J. (2014). Residential self-selection in the relationship between the built environment and travel behavior: Introduction to the special issue. *Journal of Transport and Land Use, 7*(3), 1–3.

Cao, X. J. (2009). Disentangling the influence of neighborhood type and self-selection on driving behavior: An application of sample selection model. *Transportation, 36*(2), 207–222.

Cao, X. J., & Ettema, D. (2014). Satisfaction with travel and residential self-selection: How do preferences moderate the impact of the Hiawatha light rail line? *Journal of Transport and Land Use, 7*(3), 93–108.

Cao, X. J., Mokhtarian, P. L., & Handy, S. L. (2009a). Examining the impacts of residential self-selection on travel behavior: A focus on empirical findings, *Transport Reviews, 29*(3), 359–395.

Cao, X. J., Mokhtarian, P. L., & Handy, S. L. (2009b). The relationship between the built environment and nonwork travel: A case study of Northern California. *Transportation Research A, 43*(5), 548–559.

Cao, X. J., Xu, Z., & Fan, Y. (2010). Exploring the connections among residential location, self-selection, and driving: Propensity score matching with multiple treatments. *Transportation Research Part A: Policy and Practice, 44*(10), 797–805.

Craig, P., Cooper, C., Gunnell, D., Haw, S., Lawson, K., Macintyre, S., … Thompson, S. (2012) Using natural experiments to evaluate population health interventions: New Medical Research Council guidance. *Journal of Epidemiology and Community Health, 66,*1182–1186.

Chatman, D. G. (2009). Residential choice, the built environment, and nonwork travel: Evidence using new data and methods. *Environment and Planning A, 41*(5), 1072–1089.

Chatman, D. G. (2014). Estimating the effect of land use and transportation planning on travel patterns: Three problems in controlling for residential self-selection. *Journal of Transport and Land Use, 7*(3), 47–56.

Chen, C., & Lin, H. (2011). Decomposing residential self-selection via a life-course perspective. *Environment and Planning A, 43*(11), 2608–2625.

Ewing, R., & Cervero, R. (2010). Travel and the built environment. *Journal of the American Planning Association, 76*(3), 265–294.

Festinger, L. (1957). *A theory of cognitive dissonance.* Stanford, CA: Stanford University Press.

Greenland, S., Pearl, J., & Robins, J. M. (1999). Causal diagrams for epidemiologic research. *Epidemiology, 10*(1), 37–48.

Handy, S. L., Boarnet, M., Ewing, R., & Killingsworth, R. (2002). How the built environment affects physical activity: Views from urban planning. *American Journal of Preventative Medicine, 23*(2S), 64–73.

Handy, S. L., Cao, X., & Mokhtarian, P. L. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research D, 10*(6), 427–444.

He, X., & Zhang, L. (2014). Quantifying the self-selection effects in residential location choice with a structural equation model. *Transportation Research Record, 2453,* 153–161.

Hill, A. B. (1965). The environment and disease: Association or causation? *Proceed Royal Society of Medicine, London, 58,* 295–300.

Hong, J., & Shen, Q. (2013). Residential density and transportation emissions: Examining the connection by addressing spatial autocorrelation and self-selection. *Transportation Research Part D: Transport and Environment, 22,* 75–79.

Kamruzzaman, M., Baker, D. C., & Turrell, G. (2015). Do dissonants in transit oriented development adjust commuting travel behavior? *European Journal of Transport and Infrastructure Research, 15*(1), 66–77.

Kamruzzaman, M., Baker, D. C., Washington, S. P., & Turrell, G. (2013a). Residential dissonance and mode choice. *Journal of Transport Geography, 33,* 12–28.

Kamruzzaman, M., Baker, D. C., Washington, S. P., & Turrell, G. (2013b). Does residential dissonance impact residential mobility? *Transportation Research Record, 2344,* 59–67.

MacDonald, J., Stokes, R., Cohen, D., Kofner, A., & Ridgeway, G. (2010). The effect of light rail transit on body mass index and physical activity. *American Journal of Preventive Medicine, 39,* 105–112.

Manaugh, K., & El-Geneidy, A. (2015) The importance of neighborhood type dissonance in understanding the effect of the built environment on travel behavior. *Journal of Transport and Land Use, 8,* 245–257.

McCormack, G. R., & Shiell, A. (2011). In search of causality: A systematic review of the relationship between the built environment and physical activity among adults. *International Journal of Behavioral Nutrition and Physical Activity, 8*(1), 125.

Mokhtarian, P. L., & Cao, X. (2008). Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research B, 43*(3), 204–228.

Naess, P. (2016). Built environment, causality and urban planning. *Planning Theory and Practice 17*(1), 52–71.

Naess, P. (2014a). Tempest in a teapot: The exaggerated problem of transport-related residential self-selection as a source of error in empirical studies. *Journal of Transport and Land Use, 7*(3), 57–79.

Naess, P. (2014b). Response to van Wee and Boarnet. *Journal of Transport and Land Use, 7*(3), 87–92.

Naess, P. (2009). Residential self-selection and appropriate control variables in land use: Travel studies. *Transport Reviews, 29*(3), 293–324.

Parascandola, M., & Weed, D. L. (2001). Causation in epidemiology. *Journal of Epidemiology and Community Health, 55,* 905–912.

Pinjari, A. R., Bhat, C. R., & Hensher, D. A. (2009). Residential self-selection effects in an activity time-use behavior model. *Transportation Research Part B: Methodological, 43*(7), 729–748.

Pinjari, A. R., Eluru, N., Bhat, C. R., Pendyala, R. M., & Spissu, E. (2008). Joint model of choice of residential neighborhood and bicycle ownership accounting for self-selection and unobserved heterogeneity. *Transportation Research Record, 2082,* 17–26.

Saelens, B. E., Sallis, J. F., & Frank, L. D. (2003). Environmental correlates of walking and cycling: Findings from the transportation, urban design, and planning literatures. *Annals of Behavioral Medicine, 25,* 80–91.

Scheiner, J. (2014). Residential self-selection in travel behavior: Toward an integration into mobility biographies. *Journal of Transport and Land Use, 7*(3), 15–28.

Scheiner, J. (2010). Social inequalities in travel behavior: Trip distances in the context of residential self-selection and lifestyles. *Journal of Transport Geography, 18(*6), 679–690.

Scheines, R. (1997). An introduction to causal inference. In V. R. McKim & S. P. Turner (Eds.), *Causality in crisis? Statistical methods and the search for causal knowledge in the social sciences (*pp 185–199). Notre Dame, IN: University of Notre Dame Press.

Schoner, J., & Cao, X. (2014). Walking for purpose and pleasure influences of light rail, built environment, and residential self-selection on pedestrian travel. *Transportation Research Record, 2464,* 67–76.

Schwanen, T., & 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, & Mokhtarian, P. L. (2005b). What if you live in the wrong neighborhood? The impact of residential neighborhood type dissonance on distance travelled. *Transportation Research Part D: Transport and Environment, 10*(2), 127–151.

Schwanen, T., & Mokhtarian, P. L. (2004). The extent and determinants of dissonance between actual and preferred residential neighborhood type. *Environment and planning B: Planning and Design, 31*(5), 759–784.

Shrier, I., & Platt, R. W. (2008). Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology, 8,* 70.

Suppes, P. (1970). *A probabilistic theory of causality.* Amsterdam: North-Holland Publishing Company.

Susilo, Y. O. (2015). The influence of parent's perceptions and residential self-selection to the children's travel modes at single parent households. *Transport and Sustainability, 7,* 43–64.

Tu, Y.-K., & Gilthorpe, M. S. (2012). *Statistical thinking in epidemiology.* Boca Rentoon, Fl: CRC Press, Taylor and Francis group.

van de Coevering, P., Maat, K. & van Wee, B. (2015). Multi period research designs for identifying causal effects of built environment characteristics on travel behavior. *Transport Reviews, 35*(4), 512–532.

van Dyck, D., Cardon, G., Deforche, B., Owen, N., & De Bourdeaudhuij, I. (2011). Relationships between neighborhood walkability and adults' physical activity: How important is residential self-selection? *Health and Place, 17*(4), 1011–1014.

van Ham, M., & Feijten, P. (2008). Who wants to leave the neighborhood? The effect of being different from the neighborhood population on wishes to move. *Environment and Planning A, 40*(5), 1151.

van Wee, B. (2009). Self-selection: A key to a better understanding of location choices, travel behavior and transport externalities? *Transport Reviews, 29*(3), 279–292.

van Wee, B., & Boarnet, M. (2014). Reaction to the paper Tempest in a Teapot: The exaggerated problem of transport-related residential self-selection as a source of error in empirical studies. *Journal of Transport and Land Use, 7(*3), 81–86.

Wang, D., & Lin, T. (2014). Residential self-selection, built environment, and travel behavior in the Chinese context. *Journal of Transport and Land Use, 7(3),* 5–14.

Yu, C.-Y., & Zhu, X. (2015). Impacts of residential self-selection and built environments on children's walking-to-school behaviors. *Environment and Behavior, 47*(3), 268–287.

Zhang, J. (2014). Revisiting residential self-selection issues: A life-oriented approach. *Journal of Transport and Land Use, 7*(3), 29–25.