Transportation Research Part B 68 (2014) 31–51

Contents lists available at ScienceDirect

Transportation Research Part B

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

A joint count-continuous model of travel behavior with selection based on a multinomial probit residential density choice model

Chandra R. Bhat [a,b,*], Sebastian Astroza [a,1], Raghuprasad Sidharthan [c,2], Mohammad Jobair Bin Alam [d,3], Waleed H. Khushefati [d,3]

[a] The University of Texas at Austin, Department of Civil, Architectural and Environmental Engineering, 301 E. Dean Keeton St. Stop C1761, Austin, TX 78712, United States
[b] King Abdulaziz University, Jeddah 21589, Saudi Arabia
[c] Parsons Brinckerhoff, 999 3rd Ave, Suite 3200, Seattle, WA 98104, United States
[d] King Abdulaziz University, Department of Civil Engineering, P.O. Box 80204, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Article history:
Received 14 August 2013
Received in revised form 4 February 2014
Accepted 9 May 2014

Keywords:
Multivariate dependency
Self-selection
Treatment effects
Maximum Approximate Composite Marginal Likelihood
Land-use and built environment
Travel behavior

ABSTRACT

This paper formulates a multidimensional choice model system that is capable of handling multiple nominal variables, multiple count dependent variables, and multiple continuous dependent variables. The system takes the form of a treatment-outcome selection system with multiple treatments and multiple outcome variables. The Maximum Approximate Composite Marginal Likelihood (MACML) approach is proposed in estimation, and a simulation experiment is undertaken to evaluate the ability of the MACML method to recover the model parameters in such integrated systems. These experiments show that our estimation approach recovers the underlying parameters very well and is efficient from an econometric perspective. The parametric model system proposed in the paper is applied to an analysis of household-level decisions on residential location, motorized vehicle ownership, the number of daily motorized tours, the number of daily non-motorized tours, and the average distance for the motorized tours. The empirical analysis uses the NHTS 2009 data from the San Francisco Bay area. Model estimation results show that the choice dimensions considered in this paper are inter-related, both through direct observed structural relationships and through correlations across unobserved factors (error terms) affecting multiple choice dimensions. The significant presence of self-selection effects (endogeneity) suggests that modeling the various choice processes in an independent sequence of models is not reflective of the true relationships that exist across these choice dimensions, as also reinforced through the computation of treatment effects in the paper.

* Corresponding author at: The University of Texas at Austin, Department of Civil, Architectural and Environmental Engineering, 301 E. Dean Keeton St. Stop C1761, Austin, TX 78712, United States. Tel.: +1 512 471 4535; fax: +1 512 475 8744.

E-mail addresses: bhat@mail.utexas.edu (C.R. Bhat), sastroza@utexas.edu (S. Astroza), srprasad@utexas.edu (R. Sidharthan), malam@kau.edu.sa (M.J.B. Alam), wkhushefati@kau.edu.sa (W.H. Khushefati).

[1] Tel.: +1 512 471 4535; fax: +1 512 475 8744.
[2] Tel.: +1 206 382 5289; fax: +1 206 382 5222.
[3] Tel.: +966 2 6402000x51339; fax: +966 2 6952179.

http://dx.doi.org/10.1016/j.trb.2014.05.004

1. Introduction

A question that has received particular attention within the broad land use-transportation literature is whether any effect of the built environment on travel demand is causal or merely associative (or some combination of the two; see, for example, Bhat et al., 2009; van Wee, 2009). Commonly labeled as the residential self-selection problem, the underlying issue is that the data available to assess the potential effects of land-use on travel patterns is typically of a cross-sectional nature. In such observational data, the residential location of households and the travel patterns of household members are jointly observed at a given point in time. Thus, the data reflects household residential location preferences co-mingled with the travel preferences of the households. On the other hand, from a policy perspective, the emphasis is on analyzing whether (and how much) a neo-urbanist design (compact built environment design, high bicycle lane and roadway street density, good land-use mix, and good transit and non-motorized mode accessibility/facilities) would help in reducing motorized vehicle miles of travel (VMT). To do so, the conceptual experiment that reveals the "true" effect of the built environment (BE) features of the residential location on travel patterns is the one that randomly locates households in residential locations. The problem then, econometrically speaking, is that the analyst has to extract out the "true" BE effect from a potentially non-randomly assigned (to residential locations) observed cross-sectional sample. If the non-random assignment can be completely captured by observed non-travel characteristics of households and the BE (such as, say, poor households locating in areas with low housing cost), then a conventional travel model accommodating the observed non-travel characteristics of households and the BE characteristics would suffice to extract the "true" BE effect on travel. However, it is quite possible (if not likely) that there are some antecedent characteristics of households that are not observed by the analyst and that impact both residential location choice and travel behavior. For instance, a household whose members have an overall auto inclination and a predisposition to enjoy private travel may locate itself in a conventional neighborhood (low population density, low bicycle lane and roadway street density, primarily single use residential land use, and auto-dependent urban design) and undertake substantial auto travel, while a household whose members dislike driving and prefer non-motorized and transit forms of travel may seek out neo-urbanist neighborhoods so they can pursue their activities using non-motorized and transit modes of travel. Ignoring such self-selection effects in residence choices can lead to a "spurious" causal effect of neighborhood attributes on travel, and potentially lead to misinformed BE design policies.

Many different approaches have been proposed in the literature to account for residential self-selection effects, a detailed review of which is beyond the scope of this paper (the reader is referred to Bhat and Guo, 2007; Pinjari et al., 2007; Mokhtarian and Cao, 2008; Bohte et al., 2009; van Wee, 2009; Van Acker et al., 2011, 2012). In this paper, we accept the limitations of traditional cross-sectional surveys and attempt to control for self-selection effects through econometric instrumental variable techniques, and/or parametric distribution assumptions regarding the unobserved factors. Many earlier efforts in the transportation literature have used such an approach, which can also be used in combination with other approaches (see Chatman, 2009; Pinjari et al., 2011; de Abreu e Silva et al., 2012). In doing so, we provide important empirical extensions of earlier works as well as methodological innovations, as discussed in the next section.

1.1. The current paper in the context of earlier studies

As discussed by Bhat and Guo (2007), there are several challenges in analyzing the effects of BE measures on travel behavior, even beyond the issue of residential self-selection, including the multi-dimensional nature of the BE and travel behavior. In terms of travel behavior, the different dimensions include motorized and non-motorized vehicle ownership by type, number of tours and stops, time-of-day, route choice, and travel mode choice. The net impact on overall VMT patterns will depend on the aggregation across the effects on individual travel dimensions. However, most earlier studies on the effect of BE measures on travel, while considering residential self-selection, focus directly (and solely) on the effect on vehicle miles of travel (see Zhang et al., 2012; Salon et al., 2012; Cao and Fan, 2012, which are but a few recent examples). There have also been studies that consider residential self-selection and focus on BE effects on specific travel dimensions, such as auto ownership, vehicle type, trip frequencies, bicycle ownership, activity durations, and mode choice, though these have been relatively few and have focused on each dimension individually (see Bhat and Eluru, 2009; Handy and Krizek, 2012 for detailed reviews). On the other hand, BE measures may have opposite effects on different dimensions characterizing the VMT components. For instance, a neo-urbanist design at the residence end may decrease trip lengths, but also increase the number of auto trips. As a result, a BE variable may appear to have no effect on VMT, though that may be because of opposite effects on different components constituting VMT. This is of relevance for policy, because the emissions per mile can be higher if a neo-urbanist design increases the number of auto trips, which may more than compensate for the emissions decrease because of a VMT decrease (see Sperry et al., 2012). Thus, there is a need to understand the differential effects of BE on different travel dimensions, rather than simply examine an aggregate effect on VMT or on an individual dimension of VMT. Further, the travel dimensions need to be modeled jointly because, as elucidated by Van Acker et al. (2012) and Paleti et al. (2013), self-selection need not be only through residential choice. For example, an auto-disinclined household may own fewer motorized vehicles, make fewer auto tours, as well as drive shorter distances using the car as the mode of transportation. As a consequence, any effect of the number of motorized vehicles on auto travel and VMT will be moderated by the auto-disinclined nature of the household. If some of the attributes associated with the auto-disinclined nature of the household are unobserved, there is self-selection in auto travel and VMT based not only on residential choice but also based on the number of motorized vehicles owned. This self-selection needs to be considered to obtain accurate estimates of BE effects