



The built environment typologies in the UK and their influences on travel behaviour: new evidence through latent categorisation in structural equation modelling

Kaveh Jahanshahi & Ying Jin

To cite this article: Kaveh Jahanshahi & Ying Jin (2016) The built environment typologies in the UK and their influences on travel behaviour: new evidence through latent categorisation in structural equation modelling, *Transportation Planning and Technology*, 39:1, 59-77, DOI: [10.1080/03081060.2015.1108083](https://doi.org/10.1080/03081060.2015.1108083)

To link to this article: <http://dx.doi.org/10.1080/03081060.2015.1108083>



© 2015 The Author(s). Published by Taylor & Francis.



Published online: 25 Nov 2015.



Submit your article to this journal [↗](#)



Article views: 427



View related articles [↗](#)



View Crossmark data [↗](#)

The built environment typologies in the UK and their influences on travel behaviour: new evidence through latent categorisation in structural equation modelling

Kaveh Jahanshahi  and Ying Jin

Department of Architecture, Martin Centre for Architectural and Urban Studies, University of Cambridge, Cambridge, UK

ABSTRACT

This paper uses a new latent categorisation approach (LCA) in structural equation modelling (SEM) to gain fresh insights into the influence of the built environment characteristics upon travel behaviour. So far as we are aware, this is the first LCA-SEM application in this field. We use all the main descriptors of the built environment in the UK National Travel Survey data in the analysis whilst accounting for the high correlations among the descriptors – this is achieved through defining a categorical rather than continuous latent variable for the built environment characteristics. This novel approach to defining a tangible typology of the built environment in the UK is capable of making the analytical results more cogent to formulating new, proactive land use planning and urban design measures as well as monitoring the outcomes of on-going planning and transport interventions. Since travel survey data are regularly collected across a large number of cities in the world, our approach helps to guide the design of future travel surveys for those cities in a way that enhances the analysis and monitoring of the impacts of planning and transport policies on travel choices.

ARTICLE HISTORY

Received 9 March 2015

Accepted 18 September 2015

KEYWORDS

Built environment typologies; travel demand modelling; UK National Travel Survey; structural equation modelling; latent categorical analysis; travel behaviour

1. Introduction

In this paper, we aim to formulate and test a new model that can more precisely measure the effects of the built environment upon travel demand through a novel extension to structural equation modelling (SEM). We model the built environment characteristics as a *categorical* latent variable by employing latent categorisation approach (i.e. latent class analysis- LCA) within a SEM framework. We name it a LCA-SEM approach. This approach goes beyond the existing methods using *continuous* latent variables; it enables us to quantify the influence of the built environment on travel behaviour in a tangible way – as a result, the findings has the potential to be translated into advice on policy interventions and guidance for land use planning and urban design. The statistical analysis is

CONTACT Kaveh Jahanshahi  kj299@cam.ac.uk

© 2015 The Author(s). Published by Taylor & Francis.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

placed under a SEM framework to control systematically for the effects of self-selection and spatial sorting through incorporating a comprehensive range of demographic and socio-economic variables of households and individuals as attributes describing their residential areas; we also incorporate controls for the interactions among different purposes of travel. Without those controls in the SEM, the findings would be seriously biased.

We use an extensive National Travel Survey (NTS) data set from the UK, which has the appropriate variables and sample size to support the SEM approach. To engage directly with the current policy concerns of equitable access to job opportunities and employee productivity growth, our tests are focused on travel by working adults under the retirement age; the tests are repeatable for other types of individuals. The UK NTS has been collecting an extensive set of information regarding journeys made within the country by all members of sampled households. Its purpose is to provide annual updates on personal travel and monitor changes in travel behaviour over time. The survey methodology has been continuously improved over decades recording the characteristics of the journeys made, and carefully selected personal, household and circumstantial variables that are believed to relate to or influence travel behaviour. The list of the variables is arguably the most comprehensive in travel surveys around the world, and over the years the survey has built up an impressive sample size.

The NTS has already provided valuable insights into how the UK residents travel and the data set has allowed the recorded travel patterns to be linked with the personal, household and circumstantial variables when inferring the key influences of travel behaviour. However, the characteristics of trip making and the personal, household and circumstantial variables are often highly intercorrelated, notably through endogeneity (e.g. residents' self-selection and spatial sorting), which has so far restricted the range and depth of the insights that may be gleaned from the data set. For instance, the multiple descriptors of the built environment characteristics available in the NTS data are also highly correlated to the extent that often only one of the descriptors could be used in regression-based analyses.

2. Literature review

Although the intellectual and practical interests in the complex built environment influences on travel has a long history (notably, Mitchell and Rapkin 1954; Cervero 1996; Cervero and Kockelman 1997; Banister 1997; Newman and Kenworthy 1999; Crane 2000; Ewing and Cervero 2001; Stead 2001), it is understandable that a comprehensive mapping of the effects is still emerging. First of all, the empirical data sets that include a wide range of relevant variables are difficult to assemble. Secondly, the analytical challenges that arise from model specification issues such as endogeneities among variables cast doubt on many estimates (Boarnet 2004; Cao, Mokhtarian, and Handy 2007a; Silva, Morency, and Gouliasc 2012). Thirdly, the economic, social, cultural and physical circumstances within which travel is undertaken are shifting substantially through time; regular and timely updates on the effects – which could provide fundamental insights into the changing travel behaviour – prove particularly difficult to achieve given the data and analytical challenges just mentioned.

Whilst data collection and assembly are largely dependent on funding, skills and the perceived payback, remarkable progress has been made in model specification in recent

years. In particular, there is a growing body of literature that aims to isolate the built environment effect after controlling the endogeneities among different factors such as the interdependencies¹ between travel patterns, travel attitudes, built environment characteristics and car ownership (Handy, Cao, and Mokhtarian 2005; Van Acker, Witlox, and Van Wee 2007; Gao, Mokhtarian, and Johnston 2008; Mokhtarian and Cao 2008; Bohte, Maat, and van Wee 2009; Cao, Mokhtarian, and Handy 2009; Sun et al. 2009; Cervero and Murakami 2010; Silva, Morency, and Gouliasc 2012; Sun et al. 2012; Zegras, Lee, and Ben-Joseph 2012).

Residential self-selection or sorting effect is one of the endogeneities, which has attracted a great deal of attention. As outlined by Cao, Mokhtarian, and Handy (2007b), the question is whether neighbourhood design independently influences travel behaviour or whether preferences for travel options affect residential choice. Using a self-administered 12-page survey of 1682 respondents from eight neighbourhoods in Northern California, Cao, Mokhtarian, and Handy (2007a, 2007b) and Handy, Cao, and Mokhtarian (2005, 2006) analyse the factors affecting car ownership. The respondents were questioned about their neighbourhood characteristics, neighbourhood preferences and travel attitude. The data are used to explore the role of the self-selection effect in explaining travel patterns. Notably, Cao, Mokhtarian, and Handy (2007a) examine the influences of neighbourhood characteristics, neighbourhood preferences, travel attitudes and socio-demographics on car ownership in both a cross-sectional and a quasi-panel context. The findings from cross-sectional analysis show that the correlation between neighbourhood characteristics and car ownership is primarily the result of self-selection. Apart from the SEM approach, some recent studies have adopted other modelling techniques such as latent class and random effect modelling through discrete choice analysis (Walker and Li 2007; Liao et al. 2015; Prato 2015) or propensity scoring and direct matching (McDonald and Trowbridge 2009) to control for endogeneities. Notably, Liao et al. (2015) examine the residential preferences for compact development in the State of Utah whilst controlling for heterogeneity in residential location choice arising from household socio-economic backgrounds and attitudes. Using LCA within a discrete choice framework, they classify individuals into latent classes based on their socio-demographic characteristics and attitudes towards the natural and social environments, travel mode and environmental protection. Their results suggest strong associations between location choice and socio-demographic status and attitudes. They recommend the use of SEMs as a more suitable technique to further gauge the endogenous linkages between socio-demographics, attitudes and residential preferences in future studies.

Silva, Morency, and Gouliasc (2012) is one of a limited few examples, which have examined car ownership as an intervening variable in influencing total kilometre travelled and trip frequency. In addition, they control for self-selection effects by modelling concentration, density and diversity as a function of socio-economic attributes in their SEM framework. Their results suggest that beside socio-economic self-selection effect, built environment variables significantly affect travel behaviour like commuting distance and car ownership.

Cervero and Murakami (2010) represent an important landmark in tackling both the data and model specification challenges through assembling a very large data set from 370 US urban areas around the year 2003 and employing an extensive SEM to examine the effects of density, diversity, destination accessibility and design on vehicle miles

travelled (VMT), building on analyses of the first three Ds in Cervero and Kockelman (1997). They analyse a complex web of interactions among built environment characteristics, average household income and travel demand, where travel demand is represented as VMT, percentage of commute trip by private car and rail passenger miles per capita. Their findings, after evaluating the interrelation between road density and population density, suggest that the largest reduction in vehicle travel distance comes from the combination of compact design and below-average roadway provision.

The study of temporal changes is so far focused on better quantification of the effects from quasi-panel data sets. Cao, Mokhtarian, and Handy (2007b) use a quasi-longitudinal data of movers (688 respondents who changed their residential locations over the previous year) to extend their former cross-sectional SEM analysis of the interdependencies between socio-economic factors and built environment characteristics. Their study is able to identify a small though causal effect of some built environment elements (i.e. perceived spaciousness and living in diverse-land-use areas) on car ownership. This finding is in contrast with the cross-sectional analysis of Cao, Mokhtarian, and Handy (2007a) where the correlation between neighbourhood characteristics and car ownership is found primarily to be the results of self-selection.

Adopting a quasi-longitudinal SEM approach, Aditjandra et al. (2012) report similar conclusions of the impact of neighbourhood design (e.g. accessibility, safety and attractiveness) upon the amount of private car travel after controlling for self-selection. Using Tyne and Wear metropolitan area as their case study, this is one of the first studies of this kind which has used British metropolitan data. It is also a recent study which has controlled for the endogeneity of car ownership in influencing travel, suggesting that neighbourhood design affects travel behaviour through their influence on car ownership.

Using an age-period-cohort-residential area model, Sun, Waygood, and Huang (2012) analyse the influence of five separate generation cohorts on automobility: household car ownership, the automobile mode share and the auto travel time in Osaka metropolitan area in Japan. Their analyses suggest that the life style expectations, attitudes and values represented by cohorts along with characteristics of residential area and age, have a large impact on household car ownership and auto use.

In summary, a large number of existing studies have investigated the influences on car ownership and travel distance, whereas the prevailing data difficulties meant that the existing studies tend to focus on one or several of the possible influences out of the bundle of known factors (such as travellers' socio-economic and demographic profiles, accessibility, car ownership and built environment characteristics), but very rarely the whole bundle. In addition, we are not aware of any study which has employed LCA-SEM to classify built environment into distinct categories based on built environment and socio-demographic characteristics of the residents in order to investigate the variations in influences on travel. Categorising geographical locations can better quantify the built environment effect to inform built environment and transport policies and models.

In this context, it would seem that the UK NTS data set has a great deal more to offer than hitherto explored. To date, only a handful of studies have related travel patterns to the extensive range of the NTS variables (see Stead and Marshall 2001; Stead 2001; Dargay and Hanly 2004; Jahanshahi, Williams, and Hao 2009; Jahanshahi, Jin, and Williams 2015); none except the last one have made use of the improved time series of survey results since 2002. Methodological limitations tend to be the main reason that has held back a

fuller exploitation of the comprehensive list of NTS variables. In this context, we develop here a latent categorical analysis (LCA) in a SEM.

3. Methodology

SEM is an approach to testing complex, multivariate data and differentiating direct and indirect effects using a combination of statistical data and qualitative causal assumptions. The definition of SEM was first articulated by the geneticist Wright (1921), the economist Haavelmo (1943) and the cognitive scientist Simon (1953), and was formally defined by Pearl (2000) using a calculus of counterfactuals. SEM has gained increasing acceptance in a wide range of fields including transport and urban studies (Golob 2003; Van Acker, Witlox, and Van Wee 2007; Cao, Mokhtarian, and Handy 2007b; Gao, Mokhtarian, and Johnston 2008; Weis and Axhausen 2009; Lin and Yang 2009; Cervero and Murakami 2010; Schmöcker, Pettersson, and Fujii 2011).

SEM requires the modeller to provide a conceptual model in the form of a path diagram, which hypothesises causal effects. It then tests the model on specific data to determine how valid the hypotheses are. The modeller can reconfigure the conceptual model through varying the variables and paths based on statistical fit and overall model performance.

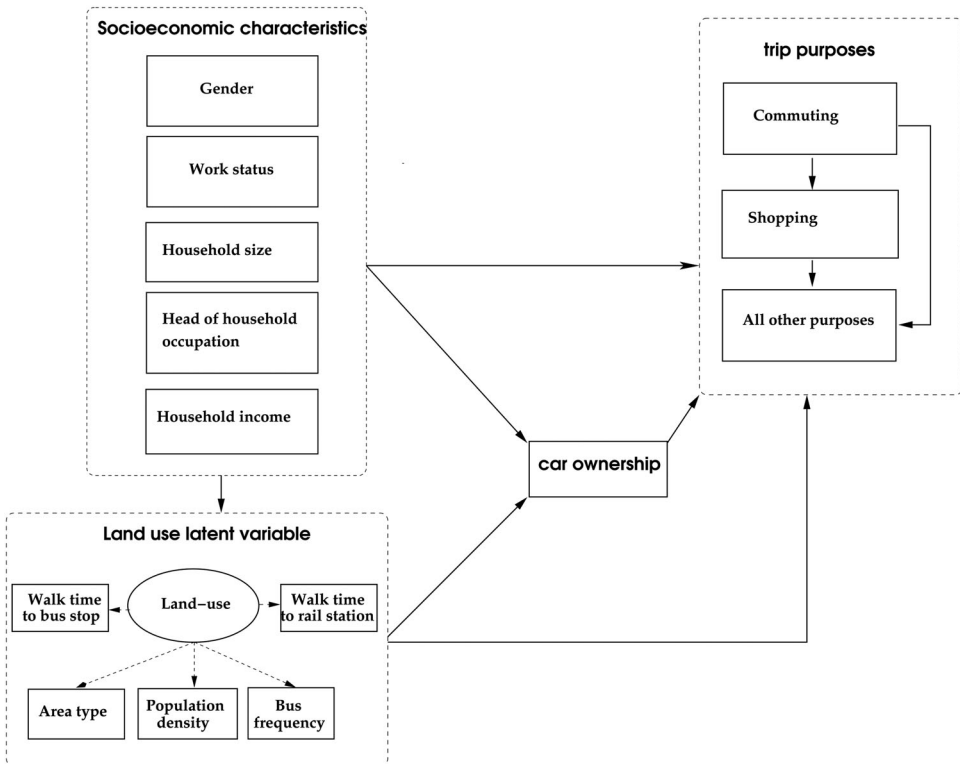


Figure 1. The conceptual structural equation model (SEM) for influences on travel.

The conceptual model, which is developed in our recent work (Jahanshahi, Jin, and Williams 2015), is proposed in Figure 1. We include in the SEM (a) a set of explanatory variables of the main socio-economic characteristics of the individuals and their households, (b) the built environment characteristics of households' residential areas modelled as the measurement indicators of built environment latent variable and (c) household car ownership. We have chosen three dependent variables, each measuring the amount of travel distance, respectively, in commuting, shopping and all other purposes. The same approach may be applied to quantify the effects of the built environment on travel time or trip frequency.

Here, we have expanded the conventional SEM formula provided in Jahanshahi, Jin, and Williams (2015) by employing conditional LCA where we model built environment as a categorical latent variable with socio-demographic characteristics of residents as controlling covariates.

LCA involves a set of observed variables, which are called indicators (i.e. in our case Area Type, Population Density, Bus Frequency and Walk Time to Bus Stops and Railway Stations in Figure 1). The indicators form the basis for estimating latent variables such as the Land Use latent variable in Figure 1. The LCA approach shares the same conceptual aim with Explanatory Factor Analysis (EFA; Jahanshahi, Jin, and Williams 2015): Both LCA and EFA are to estimate latent variables from observed indicators. However, the estimated latent variable is continuous for EFA and discrete (or categorical) for LCA – LCA gives rise to a *latent class* model because the latent variable is discrete; latent class is characterised by a pattern of conditional probabilities that indicate the chance that the variables take on specific values. When it comes to interpretation of results, EFA focuses on grouping contributing variables (such as the contribution of land use area type, density and public transport access), and can be considered as a variable-centred approach. By contrast, LCA focuses on grouping survey respondents or cases facing distinct patterns of the contributing variables into classes, and is thus a respondent-centered approach (Wang and Chen 2012).

The statistical estimations are carried out using the Mplus software (Muthen and Muthen 2007) in two stages:

Firstly, we use conditional LCA to cluster individuals who reside in similar geographical location by estimating simultaneously individuals' built environment class membership and their socio-economic background; secondly, the SEM is used to account for the inter-correlations among the built environment classes, the residents' socio-economic characteristics, their car ownership status and the interactions among different journey purposes in the quantification of the direct and indirect influences on the amount of travel carried out for each journey purpose. The second stage estimation is performed conditional on the class membership which is estimated in the first.

To formulate the first stage, let Y_{ij} be the j th indicator variable (i.e. population density, area type, etc.) of the built environment latent categorical variable, C_i , for individual i . As all our indicators are ordered categorical variables, we can formulate the link function by defining an underlying continuous variable, Y^*_{ij} such that

$$Y_{ij} = s|C_i = c \Leftrightarrow \tau_{c_j,s} < Y^*_{ij} < \tau_{c_{j,s+1}} \quad (1)$$

where C_i is the latent categorical variable (i.e. built environment), which takes values between $1, \dots, c$, and $\tau_{c_j,s}$ are a set of threshold parameters.

Conditional on regressors X (e.g. our socio-economic characteristics), we can then present the link function as

$$Y_{ij}^*|_{C_i=k,x_i} = \nu_{kj} + K_{kj}X_i + \varepsilon_{ij} \quad (2)$$

The normal distribution assumption for ε_{ij} is equivalent to a probit regression for categorical variable Y_{ij} on X_i with the following probability function:

$$\Pr(Y_{ij} = s|c_i = k) = \Phi[(\tau_{kj,s+1} - \nu_{kj} - K_{kj}X_i)] - \Phi[(\tau_{kj,s} - \nu_{kj} - K_{kj}X_i)] \quad (3)$$

The class membership probability conditional on X is given by multinomial logistic regression with the following formula:

$$\Pr(C_i = k|X_i) = \frac{\exp(\alpha_k + \gamma_k X_i)}{\sum_{s=1}^k \exp(\alpha_s + \gamma_s X_i)} \quad (4)$$

The joint probability of indicators or observed-data likelihood is then given by

$$\Pr(Y_{i1} \dots Y_{ij}) = \prod_i \sum_{k=1}^c \Pr(C_i = k) \prod_j \Pr(Y_{ij} = s|c_i = k) \quad (5)$$

EM algorithm is then used for estimating the parameters and class membership where the latent variable C_i is treated as missing data. We first compute the posterior distribution for the latent variable. The posterior conditional joint distribution is calculated as

$$\Pr(C_i = k|*) = \frac{\Pr(C_i = k) \prod_j \Pr(Y_{ij} = s|c_i = k)}{\sum_{k=1}^c \Pr(C_i = k) \prod_j \Pr(Y_{ij} = s|c_i = k)} \quad (6)$$

which is estimated given the parameters.

Given the class membership, model parameters are then estimated through maximising Equation 5. The model is solved iteratively until reaching convergence.

Equations 7–9 specify the SEM, which is estimated within each latent class for the second stage of our modelling. The subscript for latent class membership is dropped here for simplicity

$$Y_{ij} = \nu_j + K_j X_{ij} + \epsilon_{ij} \quad (7)$$

where Y_{ij} refers to the i th respondent and j th vector of a dependant variable (e.g. travel distance for commuting to work) and X_{ij} is the vector of all individual level covariates. ν_j and K_j are the vectors of intercepts and the matrices of regression parameters correspondingly.

ϵ_{ij} is a vector of residuals with a mean of zero and covariance Θ . Where the j th observed dependent variable, Y_{ij} , is a normally distributed continuous variable (e.g. the distance travelled by journey purpose), the residual variable ϵ_{ij} is assumed normally distributed. For a dichotomous variable Y_{ij} (i.e. car ownership), a normality assumption for ϵ_{ij} is equivalent to the probit regression for Y_{ij} on X_{ij} .²

The observed-data likelihood is given by

$$\prod_{ij} f_{ij}(Y_{ij}) \quad (8)$$

where f_{ij} is the likelihood function for Y_{ij} .

The expected log-likelihood is then maximised with respect to model parameter estimation:

$$\sum_{ij} \log(f_{ij}(Y_{ij})) \quad (9)$$

To avoid the trap in a local maxima for the log-likelihood, we use many different sets of starting values in the iterative maximisation procedure to ensure that the maximised value of the likelihood function is replicated.

Because the NTS is a very large data set, we consider the coefficients to be statistically significant only when the estimated coefficients have a $\geq 99\%$ confidence interval (i.e. the respective p -values are $\leq 1\%$).

4. Data

Substantial changes were made to the NTS organisation and method just before 2002 (Hayllar et al. 2005). For this paper, we therefore use the NTS data for 2002–2010, which forms a consistent time series of 9 years. The commuting, shopping and other journeys by working adults, which are used in the SEM model tests, consist of 933,296 trips and 8.2 million passenger miles travelled in the 9-year sample.³ For each journey, the NTS provides a household weight to account for non-response and a trip weight for the drop-off in the number of trips recorded by respondents during the course of the survey week, uneven recording of short walks by day of the week and the short-fall in reporting long distance trips. This is to ensure that the data are representative of travel of an average week for the UK population as a whole.

As outlined in the NTS technical report (2013),⁴ NTS data were organised into multiple levels: households, individuals, vehicles, long distance journeys made in the seven days before the placement interview or the Travel Week, whichever date was the earliest, days within the Travel Week, journeys made during the Travel Week and the stages of these journeys. In our analysis, we have used five of the linked attribute tables (i.e. up to the journey level), which are required for estimating average travel distance, as shown in Table 1.

Table 2 presents the headline averages of travel distance per week, which provide a benchmark for the analysis of the findings.

Figure 2 is the specific path diagram of our SEM model. The diagram is based on the conceptual model (Figure 1). Similar to linear regression models, for each categorical variable, one of the categories is used as the reference category. The estimated coefficients for all other categories are then evaluated relative to the reference one. In Figure 2, the reference categories are shown in parentheses. For instance, the middle level income group ‘Income level of 25k–50k’ is chosen as the reference category for the lower and higher income categories.

Table 1. A list of linked NTS data tables that are used in this paper.

| Data table | Data contents used for the analysis |
|-----------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Household | Household related variables – numbers of resident adults [1 adult, 2+ adults], annual income [less than £25k (IncomeLess25k), £25k to £50k, more than £50k (IncomeOver50k)], head of household occupation [manual, skilled manual (SkillManual), white collar clerical, professional (Prof)], frequency of local buses [level 1 for less than one a day progressing through to level 5 for at least 1 every quarter hour], walk time to bus stop [6 minutes or less, 7 to 13 minutes, 14 to 26 minutes, 27 to 43 minutes, 44 minutes or more], walk time to rail station [6 minutes or less, 7 to 13 minutes, 14 to 26 minutes, 27 to 43 minutes, 44 minutes or more], car ownership [no car, 1+ car] |
| Individual Journey | Individual related variables – gender [male, female], work status [full time (FT), part time (PT)] Variables specific to each journey made – trip purposes from, trip purposes to, travel time, travel distance, number of trips. We modelled three outbound travel purposes: Home-based work (HBW), Home-based and non-home-based Shopping (Sh) and all Other home-based and non-home-based purposes categorised as other trips (Oth)] |
| Postcode sector unit (Psu.) | Variables specific to the postcode sector unit in which the household is located – area type [from level 1 for rural areas progressing through to level 5 for London, the top metropolitan area], population density [level 1 for lowest density, i.e. under 10 persons/hectare, progressing through to level 10 the highest which is ≥ 50 persons/hectare] |

Table 2. Average travel distance per person per week: working adults.

| Period | Home-based commuting | Shopping | Other purposes | All |
|--------------|----------------------|----------|----------------|-------|
| 2002–2010 | 30.3 | 11.3 | 72.9 | 114.4 |
| 2002–2006 | 30.9 | 11.7 | 75 | 117.6 |
| 2008–2010 | 29.2 | 10.6 | 69.5 | 109.3 |
| Difference | -1.7 | -1.0 | -5.5 | -8.3 |
| % Difference | -0.1 | -0.1 | -7% | -7% |

Note: The data in this table represent outbound travel by working adults during a 7-day week. They exclude any return trips and any travel by people other than working adults. The distances are in miles per week.

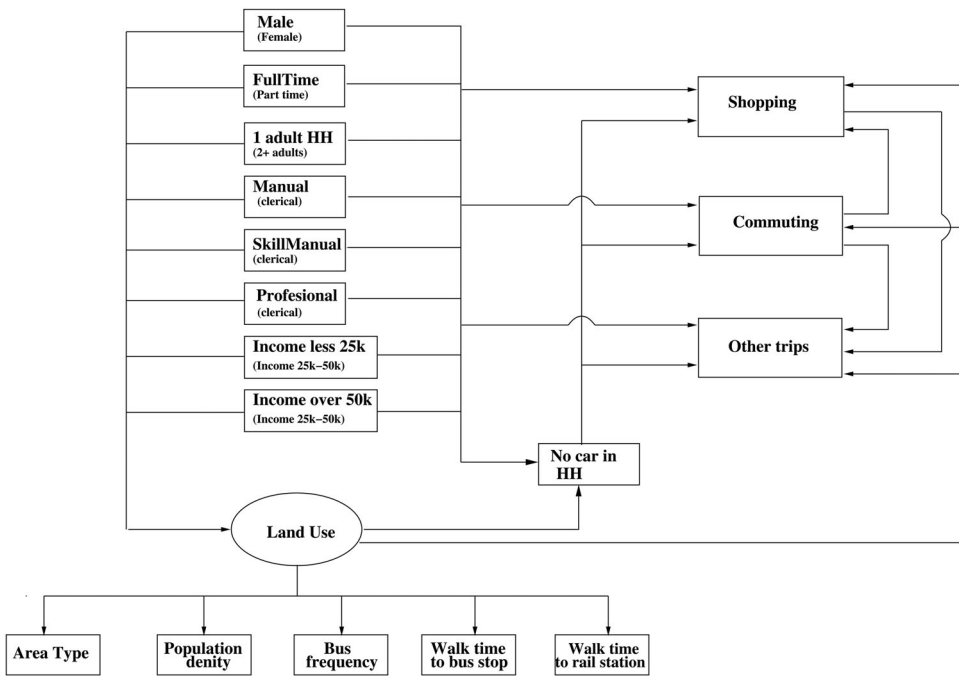


Figure 2. The SEM structure for testing the NTS data.

Downloaded by [University of California, Los Angeles (UCLA)], [Noah Hastings] at 18:53 21 February 2016

5. Main findings

A SEM test is characterised by its extensive range of outputs, with reams of tables. To present succinctly, we summarise the main findings in three steps. First, we present the latent built environment classes, their definition and unconditional and conditional probabilities for individuals to be in each class. Second, we compare the socio-economic characteristics of residents within the built environment latent classes. Finally, within each built environment class, we explore influences on travel distance by journey purpose after controlling for interactions among journey purposes as well as endogeneities arising from self-selection, spatial sorting and car ownership.

5.1. Latent classes of the built environment in the UK

The basic approach to categorisation of latent classes of the built environment is to run the LCA using NTS variables that describe the relevant characteristics of the areas the respondents live in. We have developed an extended, conditional LCA model, in which we include the demographic and socio-economic characteristics as covariates (cf. Figure 2). This involves a simultaneous estimation of the influence of the residents' demographic and socio-economic profiles so that the effects arising from spatial sorting are accounted for.

Our LCA is built on the EFA for continuous latent variable analysis in Jahanshahi, Jin, and Williams (2015). In the EFA, five built environment attributes namely 'area type', 'population density', 'frequency of local buses', 'walk time to bus stop' and 'walk time to rail station' are found to have large loading factors, sufficient to be considered as the defining characteristics of the built environment. The LCA that defines built environment as discrete categorical classes (as opposed to defining a continuous latent variable for the built environment in EFA) has similarly found those five attributes to have large loading factors. The availability of five attributes with large loading factors can allow us to define up to three distinct built environment classes with the sufficient degree of freedom for model estimation.

Our conditional LCA identifies three latent built environment classes with an entropy of 0.832.⁵ This suggests that the latent classes are very well defined. A cross-tabulation of the most likely latent class membership (row) by latent class (column) in Table 3 corroborates the high entropy value.

Panel 4a of Table 4 shows the unconditional and conditional probabilities of individuals in each latent class. Based on the estimated model, Classes 1–3 contain, respectively, 18%, 54% and 27% of all working adults.

Conditional probabilities further reveal the patterns of the latent classes benchmarked by the specific characteristics of the built environment (Panel 4b of Table 4). For example, residents in Latent Class 1 consists of, respectively, those from the medium urban, big urban, metropolitan and London area types (of, respectively, 2.2%, 15.8%, 16.2 and

Table 3. Average latent class probabilities for residents' most likely latent class membership (row) by latent class of the built environment (column).

| | Class 1 | Class 2 | Class 3 |
|--------------------|---------|---------|---------|
| Class 1 membership | 0.917 | 0.083 | 0 |
| Class 2 membership | 0.045 | 0.919 | 0.036 |
| Class 3 membership | 0 | 0.061 | 0.939 |

Table 4. Unconditional and conditional probabilities for the three-class built environment LCA model.

| Indicators | Latent class | | |
|--------------------------------------------------|-------------------------------------|---------------------------------|--------------------------------|
| | 1 – London dominated (N = 13853) | 2 – Medium urban (N = 40874) | 3 – Rural areas (N = 20301) |
| <i>Panel 4a: Unconditional probabilities</i> | 0.18 | 0.54 | 0.27 |
| <i>Panel 4b: Conditional probabilities</i> | | | |
| 4b-1: Area type | | | |
| Rural | 0 | 0.003 | 0.720 |
| Small urban | 0 | 0.080 | 0.179 |
| Medium urban | 0.022 | 0.468 | 0.078 |
| Big urban | 0.158 | 0.231 | 0.022 |
| Metropolitan | 0.162 | 0.201 | 0.001 |
| London | 0.658 | 0.015 | 0 |
| 4b-2: Population density (person/hectare) | | | |
| Under 10 | 0.003 | 0.200 | 0.949 |
| 10–14.99 | 0.021 | 0.125 | 0.027 |
| 15–19.99 | 0.019 | 0.134 | 0.019 |
| 20–24.99 | 0.020 | 0.119 | 0.005 |
| 25–29.99 | 0.039 | 0.122 | 0 |
| 30–34.99 | 0.048 | 0.089 | 0 |
| 35–39.99 | 0.053 | 0.080 | 0 |
| 40–49.99 | 0.164 | 0.096 | 0 |
| 50–59.99 | 0.168 | 0.021 | 0 |
| over 60 | 0.465 | 0.013 | 0 |
| 4b-3: Bus frequency | | | |
| Less than once a day | 0 | 0.008 | 0.206 |
| At least once a day | 0 | 0 | 0.027 |
| At least once every hour | 0.005 | 0.128 | 0.432 |
| At least once every 30 minutes | 0.131 | 0.462 | 0.283 |
| At least once every 15 min | 0.864 | 0.401 | 0.051 |
| 4b-4: Walk time to bus stops | | | |
| 44 min and more | 0 | 0 | 0.021 |
| 27–43 min | 0 | 0.001 | 0.021 |
| 14–26 min | 0.007 | 0.013 | 0.057 |
| 7–13 min | 0.072 | 0.078 | 0.108 |
| 6 min or less | 0.921 | 0.908 | 0.793 |
| 4b-5: Walk time to rail station | | | |
| 44 min and more | 0.093 | 0.336 | 0.665 |
| 27–43 min | 0.176 | 0.207 | 0.103 |
| 14–26 min | 0.355 | 0.292 | 0.129 |
| 7–13 min | 0.224 | 0.105 | 0.058 |
| 6 min or less | 0.150 | 0.060 | 0.044 |

65.8%), with no one from rural or small urban (see Panel 4b–1). The members of this class also reside in the densest areas (see Panel 4b–2) and benefit from the most frequent buses and highest level of accessibility to public transport (see Panel 4b–3 to 4b–5). The clear dominance of London residents in this latent class prompts us to label it ‘London dominated’. Similarly, the dominance of medium urban in Latent Class 2 (of 46.8% of the residents in this class) and the dominance of rural in Latent Class 3 (of 72% of residents) give rise to the labels ‘Medium urban’ and ‘Rural areas’, respectively. The individuals in Class 3 reside in the least dense area with the least convenient access to public transport. Those in Class 2 sit between Class 1 and Class 3 in terms of population density, bus frequency and public transport access.

A comparison across the three columns of latent classes gives us an insight into the distribution of residents within a NTS area type across the latent classes. For instance, for the London area type, 93.7% of the residents there belong to Latent Class 1.⁶ This composition by NTS area type is presented in Figure 3.

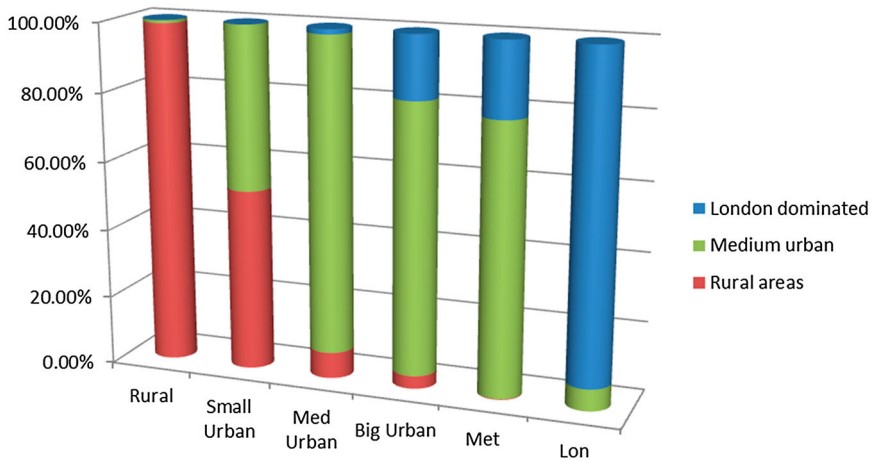


Figure 3. (Colour online) Composition of built environment latent classes by NTS land use area type.

5.2. Spatial sorting of residents among latent built environment classes

The second step of the analysis is to understand how the latent built environment class membership interacts with the demographic and socio-economic profiles of the residents – self-selection and spatial sorting of the residents of different demographic and socio-economic profiles often has a material bearing on where they live. This is carried out through the estimation of the covariates in the LCA.

The results of this analysis of the covariates are reported in terms of odds ratios with one of the latent classes designated as a reference class. This is shown in Table 5 where Latent Class 2 (Medium urban) is chosen as the reference class. For residents of a particular demographic or socio-economic characteristic, an odds ratio for a given class of built environment that is higher than 1 indicates that those residents are more likely to live in that class of built environment than in the reference class areas. Similarly, an odds ratio less than 1 implies the reverse. For instance, the odds ratio for being male is 1.077 for the ‘London dominated’ class, and this means that male workers are 7.7% more likely to live in the ‘London dominated’ areas than the ‘Medium urban’ areas.⁷ The magnitudes of the odds ratios indicate the strength of that difference. For instance, further down in Table 5 the odds ratio of skilled manual workers suggest that they are 15.8% more likely to live in ‘Rural areas’ and 43.1% less likely to live in the ‘London dominated’ areas than in the ‘Medium urban’ areas.

Not surprisingly, the results in Table 5 suggest that relative to the Medium urban class, working adults who reside in the ‘London dominated’ areas are more likely to be male, coming from one adult households, and with full-time working patterns; professionals and skilled manual workers are more likely to be found in the ‘Rural areas’ class. As for household income profiles, the ‘London dominated’ class has 56.5% more high-income households (with income >50k per year) than the ‘Medium urban’; the ‘Rural areas’ by contrast has 17.6% more high-income households than in ‘Medium urban’.

5.3. Influences on distance travelled

Table 6 shows the influence on distance travelled for different purposes across the latent built environment classes. The incorporation of the LCA provides a unique opportunity to

Table 5. Odds ratios of demographic and socio-economic covariates.

| Covariates | Built environment latent classes | | |
|-----------------------------------|----------------------------------|----------------------------------|-----------------|
| | 1 – London dominated | 2 – Medium urban | 3 – Rural areas |
| Male | 1.077*** | Used as a reference latent class | 1.077*** |
| Full-time working | 1.115*** | | 0.87*** |
| 1 adult households | 1.61*** | | 0.866*** |
| Semi- or unskilled manual workers | 0.807*** | | 0.978 |
| Skilled manual workers | 0.569*** | | 1.158*** |
| Professionals | 0.797*** | | 1.294*** |
| Household income less £25k | 1.055 | | 0.969 |
| Household income more than £50k | 1.565*** | | 1.176*** |

Note: Base or reference group is Class 2 (medium urban class).

***Significant within 99% CI, **significant within 95% CI, *significant within 90% CI.

decompose precisely the influences both for each of the demographic and socio-economic variables and across the different built environment classes. Furthermore, to identify the additional insights of incorporating a categorical built environment variable in the SEM model, we compare results from our new model with those from a constrained SEM where the model parameters do not vary across the built environment classes. This constrained SEM is typical of the existing models that do not account for the specific influences of the built environment characteristics.

To aid intuitive interpretation of the model outputs, in [Table 6](#) we first define a reference group of residents who are female, part time working in white collar clerical occupations from a car-owning household with more than one adults and a household income of 25–50k per year. The first line of the model outputs in Panel 6a reports how this group differ in their average weekly commuting distances among the three built environment classes through the model intercept values: those live in the ‘London dominated’ areas travel 10.4 miles per week, in ‘Medium urban’ 9.6 miles and in ‘Rural areas’ 13.59 miles. Similarly, the first lines under Panels 6b and 6c in [Table 6](#) show that for shopping and other travel purposes, the more rural the area, the longer the distances travelled which is intuitive. As expected, the reference group residents commute well below the working adult average of 30.3 miles per week for all classes of areas, but for shopping and other travel (for which the average weekly distances travelled are, respectively, 11.3 and 72.9 miles) they travel shorter than the average in more urban areas and longer in the rest (cf. [Table 2](#)).

The model intercepts and coefficients can help us quantify the levels of influences of the demographic and socio-economic variables in the context of the land use latent classes. Whilst an intercept represents the average travel distance of the Reference Group, the coefficients indicate how much influence a change in the demographic and socio-economic profiles has. The general patterns of small coefficients for the London-dominated class (i.e. relative to its model intercept), and the large ones for the other two land use latent classes indicates that the influence of the built environment on travel is relatively strong in the London-dominated class; this influence is much weaker in areas of the other two classes relative to that of demographic and socio-economic profiles.

For instance, the coefficient for high-income households (households with income more than £50k) in the London-dominated class is 2.1, which shows that by virtue of the higher income, such commuters travel 2.1 km more relative to the Reference

Table 6. Direct influences on travel distance (in miles) arising from traveller profiles.

| Direct influence | Constrained model | 1 – London dominated | 2 – Medium Urban | 3 – Rural areas | Class 1 vs. Class 3 Wald test p-value |
|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------|------------------|-----------------|---------------------------------------|
| Panel 6a. Direct influences on commuting | | | | | |
| Model intercept for the reference group, which is represented by a female, part time working white collar clerical worker from a car-owning household with more than one adults and a household income of 25–50k per year | | 10.39*** | 9.60*** | 13.59*** | |
| Male | 10.66*** | 6.31*** | 11.84*** | 10.84*** | 0.000 |
| Full-time working | 16.8*** | 12.83*** | 15.96*** | 20.54*** | 0.000 |
| 1 adult households | 2.88*** | –0.08 | 3.64*** | 4.87*** | 0.004 |
| Semi- or unskilled manual workers | –3.13*** | –0.35 | –3.11*** | –5.33*** | 0.001 |
| Skilled manual workers | –4.4*** | 0.01 | –3.87*** | –7.73*** | 0.000 |
| Professionals | 2.68*** | 3.1*** | 2.3*** | 2.71** | 0.787 |
| Household income less £25k | –4.32*** | –2.32*** | –4.53*** | –5.18*** | 0.023 |
| Household income more than £50k | 4.45*** | 2.1*** | 5.2*** | 4.71*** | 0.043 |
| No car in household | –4.6*** | –2.46*** | –5.79*** | –9.25*** | 0.000 |
| Panel 6b. Direct influences on shopping | | | | | |
| Model intercept for the reference group, which is represented by a female, part time working white collar clerical worker from a car-owning household with more than one adults and a household income of 25–50k per year | | 7.75*** | 12.41*** | 20.36*** | |
| Male | –3.13*** | –1.79*** | –2.7*** | –4.99*** | 0.000 |
| Full-time working | –0.98*** | –0.58*** | –0.7 | –1.5*** | 0.074 |
| 1 adult households | 0.69*** | 0.79*** | 0.84*** | 0.43 | 0.570 |
| Semi- or unskilled manual workers | –1.37*** | –0.42 | –1.54*** | –1.47** | 0.176 |
| Skilled manual workers | –1.12*** | 0.02 | –1.26*** | –1.43** | 0.028 |
| Professionals | –0.02 | 0.16*** | 0 | –0.25 | 0.511 |
| Household income less £25k | –0.56*** | –0.28*** | –0.64*** | –0.28 | 0.989 |
| Household income more than £50k | 0.07 | –0.28** | 0.01 | 0.47 | 0.207 |
| No car in household | –3.83*** | –2.58*** | –4.41*** | –7.48*** | 0.000 |
| Panel 6c. Direct influences on other purposes combined | | | | | |
| Model intercept for the reference group, which is represented by a female, part time working white collar clerical worker from a car-owning household with more than one adults and a household income of 25–50k per year | | 44.37*** | 55.99*** | 79.40*** | |
| Male | 15.03*** | 7.12*** | 15.55*** | 19.03*** | 0.000 |
| Full-time working | 2.25** | 0.85 | 1.6 | 4.31*** | 0.1881 |
| 1 adult households | 20.33*** | 18.67*** | 20.59*** | 23.74*** | 0.224 |
| Semi- or unskilled manual workers | –19.72*** | –14.11*** | –16.67*** | –29.05*** | 0.000 |
| Skilled manual workers | –20.04*** | –15.64*** | –17.15*** | –27.55*** | 0.000 |
| Professionals | 13.82*** | 6.43** | 15.08*** | 15.14*** | 0.031 |
| Household income less £25k | –10.13*** | –7.78*** | –9.18*** | –12.14*** | 0.143 |
| Household income more than £50k | 16.88*** | 14.23*** | 15.44*** | 21.45*** | 0.043 |
| No car in household | –26.06*** | –16.06*** | –32.13*** | –47.42*** | 0.000 |

***Significant within 99% CI, **significant within 95% CI, *significant within 90% CI.

Group's intercept of 13.59 km, or 20.2% more. By contrast, commuters from high-income households in medium urban and rural areas travel, respectively, 54.2% (coefficient 5.2 divided by intercept 9.6) and 34.7% (4.71/13.59) more. This pattern is mirrored by the commuting distances for commuters from households with less than 25k income per year. Similarly, households with no cars in London travel only 23.7% less (–2.46/10.39),

whilst those in medium urban and rural areas, respectively, 60.3% ($-5.79/9.6$) and 68.1% ($-9.25/13.59$) less.

The rest of the model results provide opportunities to compare the journey distances both within each column (i.e. holding the built environment class constant and decompose the influences of demographic, socio-economic and car ownership characteristics) and across the columns for each row (i.e. to identify the influence of the built environment, given a particular demographic, socio-economic and car ownership profile). Note that the values for the demographic, socio-economic and car ownership variable rows are additive within each column, which allows the readers to work out the specific distances travelled for an arbitrary type of resident. The results are intuitively correct and they provide a substantially more robust set of quantifications of the influences upon distance travelled by working adults. For instance, existing models suggest that those households with no cars tend to travel much shorter distances than those with cars. However, when we take account of the latent built environment classes, then we see considerable variability than suggested by the existing models: in the 'London dominated' areas, those with cars only commute slightly more (2.46 miles per week or 8% of the national average) than those without cars. In 'Rural areas', the corresponding value is 3.7 times higher or 9.25 miles more per week.

6. Conclusions

This paper uses a new conditional LCA in SEM to gain new insights into the influences of the built environment characteristics upon travel behaviour through the use of the UK NTS data for 2002–2010. Conditioning on demographic, socio-economic and car ownership characteristics of the households and individuals recorded in the NTS, the LCA reveals three distinct built environment categories in the UK: London dominated, Medium Urban and Rural areas. The latent classes are defined based on a specific combination of the built environment characteristics, which provides the insights into their joint influences upon travel decisions.

The LCA-SEM area categorisation reveals profound variations across geographic areas in the joint influences of demographic, socio-economic, car ownership and built environment profiles on distances travelled, with a much firmer grip on the endogeneity effects such as self-selection, spatial sorting and car ownership status. Our findings confirm that the built environment characteristics remain an important influence upon the distances travelled even after controlling for the endogeneities. This is evidenced by strong variations in our model intercepts in addition to the variations in influences upon travel distance across built environment latent classes.

For instance, although no-car owning households tend generally to travel shorter distances, the influence of car ownership upon travel is not quite the same across all areas. Significant variations in influences also exist for the majority of socio-economic characteristics and on all travel purposes. Broadly speaking, in the London-dominated class (which include 18% of the UK population) the influence of the built environment on travel is strong relative to demographic, socio-economic and car ownership profiles – here the built environment contributes significantly to the shaping of travel choices; in the Rural Areas class (27% of population), the influence of built environment is weak relative to the demographic, socio-economic and car ownership profiles. Surprisingly, although the Medium Urban areas look in many ways similar to the London-dominated ones in

physical built-upness, its built environment has just as a weak influence as the Rural areas. This indicates that the main challenges for professionals working towards sustainable transport solutions are to do with developing effective planning and design measures in the Medium urban areas (which contains 54% of population and may have already developed many of the land use planning measures to influence travel), in order to enhance the influence of the built environment on travel choices.

The main new contribution of this extended LCA-SEM model here is that the built environment as per the NTS descriptors can now be identified as tangible categories that directly relate to people's daily experiences, which makes the model cogent for monitoring the evolution of the urban and rural areas as they are transformed for better sustainability, and for identifying new interventions in land use planning and urban design to enhance the policy impacts on sustainable travel through shaping specific built environment typologies. Since travel survey data are regularly collected across a large number of cities in the world, this approach also helps to guide the design of those surveys in a way that can contribute to the analysis and monitoring of the impacts of planning and transport policies on travel choices.

Notes

1. Here we wish to highlight the bi-directional influences between built environment and travel. While this paper mainly examines the influences of the built environment on travel behaviour, it should be noted that travel behaviours can also influence the built environment over time.
2. For more information on modelling categorical data in SEM and MPLUS, see Muthén (1984).
3. For comparison all the commuting, shopping and other journeys in the NTS sample for all people (both working adults and others) total 1.84 million trips and 13.5 million passenger miles travelled for 2002–2010. The total return journeys in the sample, which are not used in the LCA-SEM model, total 1.36 million trips and 9.7 million passenger miles travelled for the same period.
4. The report can be found at https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/337263/nts2013-technical.pdf.
5. Entropy is measured on a 0 to 1 scale with the value of 1 indicating the individuals are perfectly classified into latent classes, and a value that is greater than 0.8 indicates a well-defined categorisation (Wang and Wang, 2012).
6. $(0.658 \times 13853) / (0.658 \times 13853 + 0.015 \times 40874 + 0.00 \times 20301)$ using data in Panel 4b-1 of Table 4.
7. This result is different to that produced by Jahanshahi, Jin, and Williams (2015) where built environment is modelled as a continuous latent variable – their results in that paper indicate that male workers tend to commute from less dense and more rural locations with less frequent bus services, which is counterintuitive. This highlights the benefits of modelling built environment as a categorical as opposed to a continuous latent variable.

Acknowledgements

The authors wish to acknowledge unflinching advice and comments from Ian Williams of Ian Williams Services and Mplus software developers. We also acknowledge the NTS data provided by the UK Data Archive, which is supplemented by the UK Department for Transport. The usual disclaimers apply and the authors alone are responsible for any views expressed and any errors remaining.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported by the EPSRC Centre for Smart Infrastructure and Construction at Cambridge University [grant number EP/K000314/1]; EPSRC Doctoral Training Grant [grant number EP/P505445/1]. Kaveh Jahanshahi acknowledges the support of an EPSRC Doctoral Training Grant (EPSRC reference EP/P505445/1) and Ying Jin acknowledges the funding support from EPSRC Centre for Smart Infrastructure and Construction at Cambridge University (EPSRC reference EP/K000314/1).

ORCID

Kaveh Jahanshahi  <http://orcid.org/0000-0003-3232-4600>

References

- Aditjandra, Paulus Teguh, Xinyu (Jason) Cao, and Corinne Mulley. 2012. "Understanding Neighbourhood Design Impact on Travel Behaviour: An Application of Structural Equations Model to a British Metropolitan Data." *Transportation Research Part A: Policy and Practice* 46 (1): 22–32.
- Banister, D. 1997. "Reducing the Need to Travel." *Environment and Planning B — Planning & Design* 24 (3): 437–449. doi:10.1068/b240437.
- Boarnet, Marlon G. 2004. *The Built Environment and Physical Activity: Empirical Methods and Data Resource*. Washington, DC: Transport Research Board.
- Bohte, W., K. Maat, and B. van Wee. 2009. "Measuring Attitudes in Research on Residential Self-Selection and Travel Behaviour: A Review of Theories and Empirical Research." *Transport Reviews* 29 (3): 325–357. doi:10.1080/01441640902808441.
- Cao, Xinyu, Patricia L. Mokhtarian, and Susan L. Handy. 2007a. "Cross-sectional and Quasi-Panel Explorations of the Connection between the Built Environment and Auto Ownership." *Environment and Planning A* 39 (4): 830–847. doi:10.1068/a37437.
- Cao, Xinyu, Patricia L. Mokhtarian, and Susan L. Handy. 2007b. "Do Changes in Neighborhood Characteristics Lead to Changes in Travel Behavior? A Structural Equations Modeling Approach." *Transportation* 34 (5): 535–556. doi:10.1007/s11116-007-9132-x.
- Cao, Xinyu, Patricia L. Mokhtarian, and Susan L. Handy. 2009. "The Relationship between the Built Environment and Nonwork Travel: A Case Study of Northern California." *Transportation Research Part A – Policy and Practice* 43 (5): 548–559. doi:10.1016/j.tra.2009.02.001.
- Cervero, Robert. 1996. "Jobs-Housing Balance Revisited – Trends and Impacts in the San Francisco Bay Area." *Journal of the American Planning Association* 62 (4): 492–511. doi:10.1080/01944369608975714.
- Cervero, Robert, and Kara Kockelman. 1997. "Travel Demand and the 3Ds: Density, Diversity, and Design." *Transportation Research Part D: Transport and Environment* 2 (3): 199–219. doi:10.1016/S1361-9209(97)00009-6.
- Cervero, Robert, and Jin Murakami. 2010. "Effects of Built Environments on Vehicle Miles Traveled: Evidence from 370 US Urbanized Areas." *Environment and Planning A* 42 (2): 400–418.
- Crane, Randall. 2000. "The Influence of Urban Form on Travel: An Interpretive Review." *Planning Literature* 15 (1): 3–23. doi:10.1177/08854120022092890.
- Dargay, J., and M. Hanly. 2004. "Land Use and Mobility." Paper presented at the World Conference on Transport Research, Istanbul, Turkey. doi:http://discovery.ucl.ac.uk/1236/.
- Ewing, Reid, and Robert Cervero. 2001. "Travel and the Built Environment: A Synthesis." *Transportation Research Record: Journal of the Transportation Research Board* 1780: 87–114. doi:10.3141/1780-10.

- Gao, Shengyi, Patricia L. Mokhtarian, and Robert A. Johnston. 2008. "Exploring the Connections Among Job Accessibility, Employment, Income, and Auto Ownership Using Structural Equation Modeling." *Annals of Regional Science* 42 (2): 341–356. doi:10.1007/s00168-007-0154-2.
- Golob, Thomas F. 2003. "Structural Equation Modeling for Travel Behavior Research." *Transportation Research, B – Methodological* 37: 1–25.
- Haavelmo, Trygve. 1943. "The Statistical Implications of a System of Simultaneous Equations." *Econometrica* 11 (1): 1–12. doi:10.2307/1905714.
- Handy, S., X. Y. Cao, and P. Mokhtarian. 2005. "Correlation or Causality between the Built Environment and Travel Behavior? Evidence from Northern California." *Transportation Research Part D – Transport and Environment* 10 (6): 427–444. doi:10.1016/j.trd.2005.05.002.
- Handy, S., X. Y. Cao, and P. L. Mokhtarian. 2006. "Self-selection in the Relationship between the Built Environment and Walking – Empirical Evidence from Northern California." *Journal of the American Planning Association* 72 (1): 55–74. doi:10.1080/01944360608976724.
- Hayllar, Oliver, Paul McDonnell, Christopher Mottau, and Dorothy Salathiel. 2005. *National Travel Survey 2003 & 2004 Technical Report*. London: UK Department for Transport.
- Jahanshahi, Kaveh, Ying Jin, and Ian Williams. 2015. "Direct and Indirect Influences on Employed Adults' Travel in the UK: New Insights from the National Travel Survey Data 2002–2010." *Transportation Research Part A: Policy and Practice* 80: 288–306. doi:10.1016/j.tra.2015.08.007.
- Jahanshahi, Kaveh, Ian Williams, and Xu Hao. 2009. "Understanding Travel Behaviour and Factors Affecting Trip Rates." Paper presented at European Transport Conference, Leeuwenhorstm, the Netherlands.
- Liao, Felix Haifeng, Steven Farber, and Reid Ewing. 2015. "Compact Development and Preference Heterogeneity in Residential Location Choice Behaviour: A Latent Class Analysis." *Urban Studies* 52 (2): 314–337. doi:10.1177/0042098014527138.
- Lin, Jen-Jia, and An-Tsei Yang. 2009. "Structural Analysis of How Urban Form Impacts Travel Demand: Evidence from Taipei." *Urban Studies* 46 (9): 1951–1967. doi:10.1177/0042098009106017.
- McDonald, Noreen, and Matthew Trowbridge. 2009. "Does the Built Environment Affect When American Teens Become Drivers? Evidence from the 2001 National Household Travel Survey." *Journal of Safety Research* 40 (3): 177–183. doi:10.1016/j.jsr.2009.03.001.
- Mitchell, R. B., and C. Rapkin. 1954. *Urban Traffic: A Function of Land Use*. New York: Columbia University Press.
- Mokhtarian, Patricia L., and Xinyu Cao. 2008. "Examining the Impacts of Residential Self-selection on Travel Behavior: A Focus on Methodologies." *Transportation Research Part B: Methodological* 42 (3): 204–228. doi:10.1016/j.trb.2007.07.006.
- Muthén, Bengt. 1984. "A General Structural Equation Model with Dichotomous, Ordered Categorical, and Continuous Latent Variable Indicators." *Psychometrika* 49 (1): 115–132.
- Muthen, Linda K., and Bengt O. Muthen. 2007. *Mplus User's Guide*. 5th ed. Los Angeles, CA: Muthén & Muthén: 197–200.
- Newman, Peter, and Jeffrey R. Kenworthy. 1999. *Sustainability and Cities: Overcoming Automobile Dependence*. Island Press. <https://islandpress.org/book/sustainability-and-cities>.
- Pearl, Judea. 2000. *Causality: Models, Reasoning and Inference*. Vol. 29. Cambridge University Press. <http://www.amazon.com/Causality-Reasoning-Inference-Judea-Pearl/dp/0521773628>.
- Prato, Carlo Giacomo. 2015. "Latent Lifestyle and Mode Choice Decisions When Travelling Short Distances." General Papers IATBR, Sindsor, 2015-07-13.
- Schmöcker, Jan-Dirk, Pierre Pettersson, and Satoshi Fujii. 2011. "Comparative Analysis of Proximal and Distal Determinants for the Acceptance of Coercive Charging Policies in the UK and Japan." *International Journal of Sustainable Transportation*. 6 (3): May–June, 2012: pp. 156–173.
- Silva, João de Abreu e, Catherine Morency, and Konstadinos G. Gouliasc. 2012. "Using Structural Equations Modeling to Unravel the Influence of Land Use Patterns on Travel Behavior of Workers in Montreal." *Transportation Research Part A: Policy and Practice* 46 (8): 1252–1264.
- Simon, H. A. 1953. "Causal Ordering and Identifiability." In *Studies in Econometric Method*, edited by W.C. Hood and J.C. Koopmans, 49–74. New York, NY: Wiley.

- Stead, D. 2001. "Relationships between Land Use, Socioeconomic Factors, and Travel Patterns in Britain." *Environment and Planning B-Planning & Design* 28 (4): 499–528. doi:10.1068/b2677.
- Stead, Dominic, and Stephen Marshall. 2001. "The Relationships between Urban Form and Travel Patterns. An International Review and Evaluation." *European Journal of Transport and Infrastructure Research* 1 (2): 113–141.
- Sun, Yilin, E. Waygood, Kenichiro Fukui, and Ryuichi Kitamura. 2009. "Built Environment or Household Life-Cycle Stages – Which Explains Sustainable Travel More?" *Transportation Research Record: Journal of the Transportation Research Board* 2135: 123–129. doi:10.3141/2135-15.
- Sun, Yilin, E. Waygood, and Zhiyi Huang. 2012. "Do Automobility Cohorts Exist in Urban Travel?" *Transportation Research Record: Journal of the Transportation Research Board* 2323: 18–24. doi:10.3141/2323-03.
- Van Acker, V., F. Witlox, and B. Van Wee. 2007. "The Effects of the Land Use System on Travel Behavior: A Structural Equation Modeling Approach." *Transportation Planning and Technology* 30 (4): 331–353. doi:10.1080/03081060701461675.
- Walker, Joan L, and Jieping Li. 2007. "Latent Lifestyle Preferences and Household Location Decisions." *Journal of Geographical Systems* 9 (1): 77–101. doi:10.1007/s10109-006-0030-0.
- Wang, Tingting, and Cynthia Chen. 2012. "Attitudes, Mode Switching Behavior, and the Built Environment: A longitudinal Study in the Puget Sound Region." *Transportation Research Part A: Policy and Practice* 46 (10): 1594–1607.
- Wang, Jichuan, and Xiaoqian Wang. 2012. "Structural Equation Modeling: Applications Using Mplus." Probability and Statistics. Wiley <http://www.wiley.com/WileyCDA/WileyTitle/productCd-1119978297.html>.
- Weis, Claude, and Kay W. Axhausen. 2009. "Induced Travel Demand: Evidence from a Pseudo Panel Data Based Structural Equations Model." *Research in Transportation Economics* 25 (1): 8–18. doi:10.1016/j.retrec.2009.08.007.
- Wright, Sewall. 1921. "Correlation and Causation." *Journal of Agricultural Research* 20 (7): 557–585.
- Zegras, Christopher, Jae Seung Lee, and Eran Ben-Joseph. 2012. "By Community or Design? Age-restricted Neighbourhoods, Physical Design and Baby Boomers' Local Travel Behaviour in Suburban Boston, US." *Urban Studies* 49 (10): 2169–2198. doi:10.1177/0042098011429485.