Trendbreaking influences of built form upon travel in UK cities: evidence from new quantifications of within- and between-built-form variations

Kaveh Jahanshahi
The Martin Centre for Architecture and Urban Studies, University of Cambridge
1-5 Scroope Terrace, Cambridge, CB2 1PX
Tel: +44-7737-987812; Email: kj299@cam.ac.uk

Ying Jin
The Martin Centre for Architecture and Urban Studies, University of Cambridge
1-5 Scroope Terrace, Cambridge, CB2 1PX
Tel: +44-1223-760112 Email: ying.jin@aha.cam.ac.uk

Word count: 4,750 words text + 9 tables/figures x 250 words (each) +31 references = 7,000 words

TRR Paper number: 16-4708

27/Feb/2016

ABSTRACT

Using a consistent time series of 11 years (2002-2012, 1,137,259 sample trips) and all main variables from the UK National Travel Survey (NTS), this paper develops a novel structural equation model (SEM) with random intercepts to gain new insights into the influences of the built form characteristics upon travel behaviour. Simultaneous modelling of car ownership as an endogenous variable in determining influences on travel outcomes and allowing SEM intercepts to vary across built form categories, we control for car ownership endogeneity whilst measuring within and between built form variations through incorporating random intercepts in the model. The new models unambiguously decompose the self-selection/spatial sorting of households effects from the influence of the built form characteristics. This research further investigates trend breaking influences by generating model outputs pre-2007 and post-2007 using a multi-time-period structure. Our new quantifications are built on more robust statistical theories than hitherto and thus imply significant implications on how to interpret the impact of recent transformations in UK's inner cities. The method can also be used to produce a regular and timely update on any shifts in the influences on travel, and provide feedbacks for land use planning and integrated built form/transport interventions.

Keywords: structural equation modelling, random effect modelling, UK National Travel Survey, accessibility, mobility, car ownership

1. INTRODUCTION

During the past two decades many of the dense urban areas in UK have seen momentous changes in population growth and car traffic decline. One of the key drivers of change has been enhanced integration between transport investments with appropriate physical design of urban land use and built forms (*1*). The inner city areas of UK's large cities, such as Manchester, Leeds, Birmingham, Bristol and London have seen population rise accompanied by a significant fall in per person car ownership and car use. Among the successful examples, it is often the case that a higher population rise is associated with a sharper reduction in car use. For instance, during 1995-2013 inner London saw a population rise by 20% whilst per person car ownership drop by 10% and road traffic drop by 15%; the number of private cars entering the core business area in morning rush hours has dropped by 60%. This is now being experimented in the core cities of lower density areas, such as Cambridge and Milton Keynes. The new models of urban growth allows more jobs to be created in the most productive locations whilst achieving lower carbon travel, which may be cogent to developing new urban development strategies for a large number of cities around the world.

However, there has been considerable debate regarding what has caused car ownership and urban car traffic to fall, and whether this trend would continue. In this paper we put forward a new modelling method incorporating structural equation modelling with random intercepts analysis, which we hope to shed a fresh light on the debate, particularly regarding the many and varied interdependencies and endogeneities among the key influences on travel decisions. The analyses and modelling are supported by a very large dataset from the UK National Travel Survey (NTS), which is one of the most comprehensive travel surveys in the world.

2. LITERATURE REVIEW

Recent literature have targeted the challenges in dealing with interdependencies and endogeneities among the influences of travel, which are often at the center of the debates e.g. over the effectiveness of interventions in land use in contemporary cities (*2-4*). The nature and magnitudes of the influences are expected to shift substantially through time. However, persisting data difficulties have made it demanding to investigate cross-sectionally, let alone obtaining regular updates on how such influences evolve.

Significant progress has been made since the early 2000s to map the influences through structural equation modelling (SEM). The literature is particular focused on the interdependencies among travel patterns, attitudes, built form characteristics and car ownership (*3-15*).

Residential self-selection and sorting effects have attracted a lot of attention, i.e. whether neighbourhood design independently influences travel behaviour or whether preferences for travel options affect residential choice (*11*). Using a self-administered twelve-page survey of 1682 respondents from eight neighbourhoods in Northern California, Handy et al (*5; 16*) and Cao et al (*3; 11*) analyse the factors affecting car ownership. The responses regarding neighbourhood characteristics, neighbourhood preferences, travel attitudes are examined along with the socioeconomic profiles both cross-sectionally and as quasi-panel, which shows that the correlations between neighbourhood characteristics and car ownership is primarily the result of self-selection.

By contrast, Giuliano and Dargay (*6*), Silva et al (*7*) and Silva et al (*4*) are among the notable few who test car ownership as an endogenous variable that is itself subject to strong influences of the built form as well as residents' socioeconomic profiles; after controlling for self-selection effects, they find that land use characteristics does significantly affect car ownership levels as well as travel behaviour.

It would seem that in four aspects the SEM approach still has under-tapped potential. First, most studies reveal insights into the influences on distance travelled, but so far relatively few do so on travel time; this limits the understanding of influences on travel accessibility and leaves an apparent gap on the efficiency of travel and mobility. Secondly, although existing studies collectively suggest that significant endogenous interactions exist among the influencing factors like travellers' socio-economic and demographic profiles, residents' self-selection and spatial sorting, land use, built form and to some extent car ownership, few if any studies have examined this whole range of influences in one model. Thirdly, there are some potentially important interactions that have been left un-investigated, such as the influences between different trip purposes or travel outcomes (i.e. travel distance and time). For instance, would longer commuting be offset by shorter shopping journeys? Fourthly, few existing studies could easily provide regular updates going forward without major data efforts, with the exception of population census-based longitudinal work which could discern some variables regarding travel outcomes.

To respond to the above research gaps, Jahanshahi et al (*17*) use the NTS dataset to develop new, better integrated structural equation models (SEMs) to uncover the influences of latent land use characteristics, indirect influences on car ownership, interactions among trip purposes as well as residents' self-selection and spatial sorting. That paper highlights the contributions in the UK of land use planning and urban design in restraining travel demand in the 2000s. However, it also shows that there may still be a large mobility disadvantage among the fastest growing segments of workers, particularly in dense urban areas.

However, to develop policy and urban design responses it would be necessary to have clearer insights into the specific influences of land use and built form in different types of areas. The approach adopted by Jahanshahi et al (*17*) in modelling land use and built form as a latent variable is unable to provide such specific quantifications.

In this context the idea of random effects analysis, which has so far been applied in transport studies with mixed discrete choice models, may be relevant. A good example is Bhat and Gossen (*18*) where a mixed multinomial logit model is developed for the type of weekend recreational activity episodes that individuals pursue. Their model allows both unobserved heterogeneity in preferences across individuals, and correlation across unobserved utility components of the alternatives. The former is accommodated through defining a random intercept which can vary across individuals, and the latter through decomposing the error terms into two components: one standard iid and one to induce heteroskedasticity and correlations across unobserved utility components of the alternatives.

Spissu et al (*19*) study six categories of discretionary activity participation to understand influences on the inter-personal and intra-personal variability in weekly activity engagement. They develop a mixed multiple discrete-continuous extreme value (MMDCEV) model to dismantle inter-individual and intra-individual variations. In their formulation, the intercept has three separate components: the first one is to capture the average effects across individuals of unobserved variables on the base-line utility for a particular choice alternative (i.e. a fixed component across the individuals); the second one is to capture the heterogeneity across individuals due to unobserved individual attributes that are not correlated across alternatives; the third one is to adopt a mechanism to generate individual level correlation across unobserved utility components of the alternatives. The second and third components together are equivalent to a random intercept which capture inter-individual (or between level) variations.

Notwithstanding the rapid progress in the use of random intercept models in mixed travel choice modelling, we are not aware of any studies which combine such random effects with a SEM. We note that the random intercept approach has been widely used in other disciplines such as education and health (*20-23*). This is because of its capability for simultaneously controlling potential endogeneities through SEM and measuring the macro level variations when individuals are nested within more aggregate units (such as schools in educational studies or built form clusters in our context).

It would therefore appear of both theoretical and policy interested to incorporate a random intercept model in SEM for examining the more complex and controversial aspects regarding influences on travel behaviour. In particular, the new model would allow variations in modelled intercepts across built form categories. This not only captures the heterogeneity across built form due to unobserved built form attributes, but also, in a second level of the model, allows the quantification of the influences of observed household socioeconomic profiles - i.e. both self-selection and spatial sorting effects.

3. METHOD OF ANALYSIS

In this paper we construct two-level Structural Equation Model (SEM) to benefit from a comprehensive range of variables available from the NTS data set, including household socioeconomic characteristics, car ownership status, built form characteristics of the household residential areas, weekly travel distances and travel times by trip purposes, etc. The model builds upon our recent work (*17*) which adopts a single level SEM with a land use latent variable to control for self-selection and spatial sorting effects (see right hand diagram in Figure 1). For this paper, the land use latent variable, which has been used in Jahanshahi et al (*17*) is replaced with random intercepts which can vary across built form categories for each regression equation of the SEM (cf left hand diagram in Figure 1). The random intercepts provide more precise quantification of the share of self-selection and spatial sorting effects vis-à-vis land use characteristics in explaining the influences on travel behaviour.

The left hand diagram in Figure 1 presents graphical description of SEM with random intercept. The upper path diagram postulates the structure of the model *within* built form clusters. The within model is built on our prior experience in NTS analysis (*17; 24-28*), we have settled upon a conceptual path diagram that consists of the following two types of explanatory variables: (1) a long list of demographic and socioeconomic characteristics as exogenous variables; (2) car ownership as a mediating, endogenous variable that is subject to influences from (1). This enables us to control for car ownership endogeneity. For dependent variables, we exploit the fact that the NTS records weekly travel distance and trip time by trip purpose. We set up one model respectively for each of the two travel outcomes, within which the amounts of travel by trip purpose are defined as separate dependent variables – this will allow us to see if any complementary or substitutive effects exist among the trip purposes.

Regression equations of car ownership, travel distance and travel time by purpose have random intercepts which vary across built-environment clusters. This *between* level variation can be left unrestricted (we call this **Model A**) or can be conditional on exogenous variables at built-environment cluster level as is shown in lower diagram in Figure 1 (we name this **Model B**).

We use the 14 population density levels and 8 area types as defined in NTS (for details, see TABLE 1 below) to form 98 built form cluster categories with the range of 24 to over 16000 residences. The smallest group is Inner London with a population density of 1 to 5 persons per hectare which has only 24 residing individuals and the largest one is the rural area with less than 1 individual per hectare with over 16000 individuals. . Naturally there are a few density/area type combinations (such as 'high density rural areas') which are omitted from the analysis.
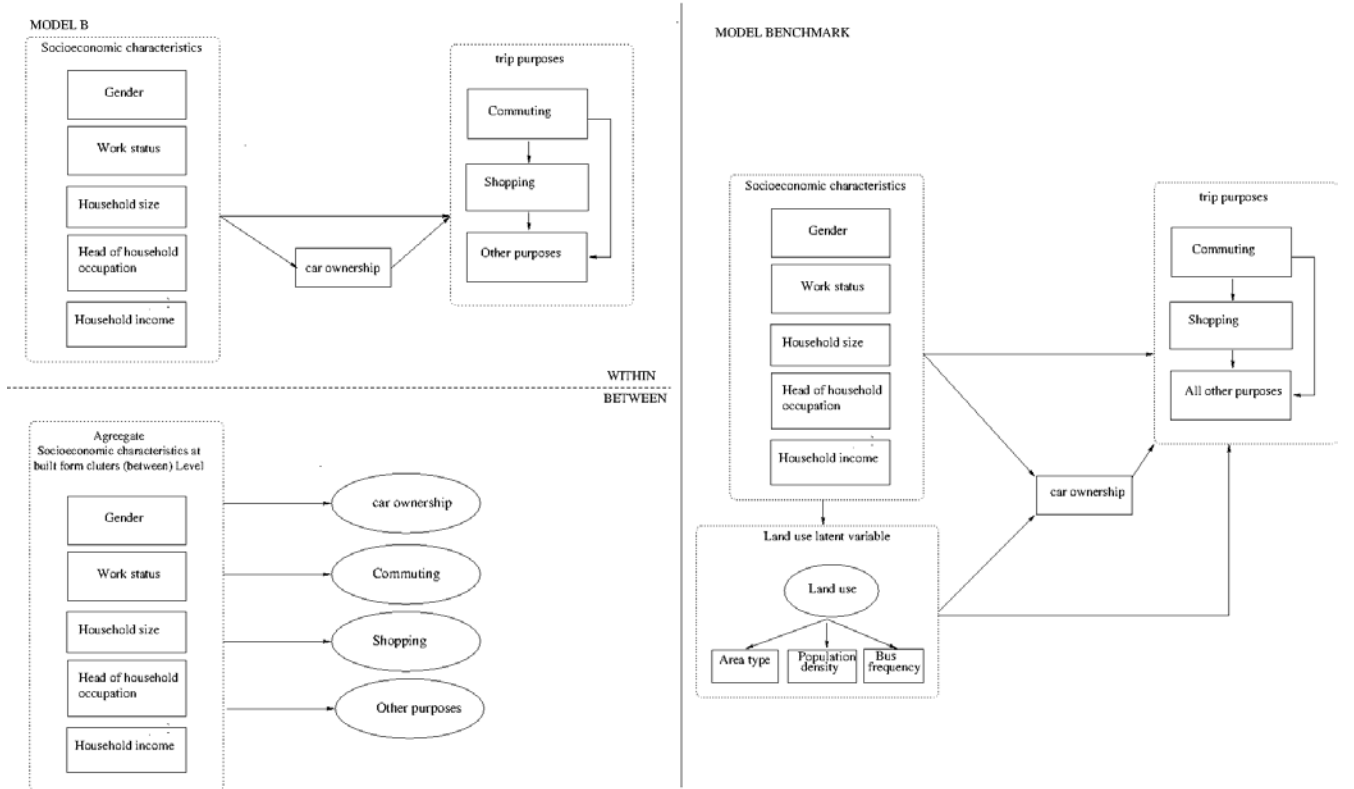
Figure 1 Comparing random intercept SEM (Model B) with single level SEM presented in Jahanshahi et al (*17*) (Model Benchmark)

Note: Eclipse represents unobserved/latent variables- Left: random intercepts, Right: Land use latent variable .

The model notations are straightforward. Let y_{pij} be the p-th dependent variable (i.e. car ownership status, travel distance, travel time by trip purpose) for individual i in built-environment cluster j. We proceed as in Muthen (29) by defining an underlying normally distributed latent variable y^{*}_{pij}. For travel distance and travel time by purposes which are normally distributed, we have $y_{pij} = y^{*}_{pij}$. The car ownership variable is a binary variable with probit distribution, we can parameterize its function as shown in equation 1.

$$y_{pij} = \begin{cases} 1 & y^{*}_{pij} > 0 \\ 0 & Otherwise \end{cases} \qquad (1)$$

The two level model can be constructed as:

$$y^{*}_{pij} = y_{wpij} + y_{bpij} \qquad (2)$$

where y_{wpij} and y_{bpij} are individual level (within level) and built-environment cluster level (between level, i.e. random intercept) components of y^{*}_{pij} respectively. Both y_{wpij} and y_{bpij} are assumed to be normally distributed.

Equation 3 shows the model notation at individual level (within-level model)

$$Y_{wij} = BY_{wij} + \Gamma_{w} X_{wij} + \varepsilon_{wij} \qquad (3)$$

where Y_{wij} and X_{wij} are the vector of within level dependent (y_{wpij}) and independent variables (x_{wmij})-i.e. individual and household level socioeconomic variables in our model. The random intercepts (i.e. y_{bpij}) can be modelled unrestricted which follow the notation in

equation 4 (we call this Model A in section 5) or as a function of between level independent variables (i.e. Model B in section 5) which is shown in equation 5.

$$Y_{bij} = \alpha_b + \varepsilon_{bij} \tag{4}$$

$$Y_{bij} = \alpha_b + \Gamma_b X_{bij} + \varepsilon_{bij} \tag{5}$$

where Y_{bij} and X_{bij} are the vectors of model random intercepts (i.e. y_{bpij}) and socioeconomic characteristics at built-environment clusters (i.e. x_{bnij}) respectively.

$B, \Gamma_w, \alpha_b, \Gamma_b$ are the vectors and matrices of slopes and regression parameters to be estimated. ε_{wij} and ε_{bij} are zero mean normally distributed independent vector variables. For identification purposes, the variance of the error term (ε_{wpij}) associated with car ownership binary variable is fixed to 1. The model is estimated by a maximum likelihood estimator using an EM algorithm (30) with the latent variable Y_{bij} (i.e. random intercept) treated as missing data., The observed data likelihood is:

$$L = \int \psi_b(Y_b) \prod_i f_{wi}(Y_{wi}) d_{Y_b} \tag{6}$$

where $f_{wi}(Y_{wi})$ is the likelihood function of within level variable, being marginalized over between-level random intercept.

The model is estimated in the MPLUS software (https://www.statmodel.com, accessed 15 May 2015). MPlus provides a number of specific estimation options that are suitable for our purposes, such as the capability for estimating multi-level SEM over multiple time periods and for estimating continuous, censored, binary, ordered categorical (ordinal), unordered categorical (nominal), counts, or a combination of them.

In practice, there are two general limitations in estimating the random intercept models. First, the degree of freedom of the between level model is a function of the number of clusters (i.e. 98 built form categories here). This can limit the number of between-level parameters to be estimated. Second, the complexity of between level model can increase the dimensions of the numerical integration resulting in slow convergence. Such considerations have been factored into model design to ensure feasible and easy to understand model results.

4. NTS DATA

For this paper we use the NTS data for 2002-2012 which forms a consistent time series of eleven years. There are in total 1,137,259 trips and 9.9 million passenger miles travelled for commuting, shopping and other journeys by employed adults. 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 the data is representative of travel of an average week for the UK population as a whole.

The NTS data is organised in nested related tables of households, individuals, long distance journeys prior to the survey, days within the survey week, journeys made during the survey week, the stages of these journeys and vehicles (*31*). Based on previous NTS analysis (*17; 24; 25; 27; 28*), we have selected all the main attributes for households, individuals and their trip-making that have shown to be suitable in previous work (TABLE 1).

The NTS data is subject to what the survey consider significant and practical to collect within the available resources and without causing an undue increase in attrition rates.

TABLE 1 NTS data: Definitions of variables selected for SEM analysis

Data for	Classifications of respective variables			
Households	Household size: 1 Adult; >1 adult.	Annual income: <£10,000; £10,000-14,999; £15,000-19,999; £20,000-24,999; £25,000-29,999; £30,000-34,999; £35,000-39,999; £40,000-49,999; £50,000-59,999; ≥£60,000.	Household head occupation: Manual; Skilled manual; White collar clerical; Professional.	Car ownership: No access to car; Own or access to one or more than one car.
Employed adults (linked to households)	Gender: Male; Female.		Work status: Full time (FT); Part time (PT).	
Journeys (linked to adults)	Journey purpose (for outbound purpose): Home-based commuting(HBW); All shopping (Sh); All other purposes (Oth).		Journey distance: Miles/trip	Journey time: Minutes/trip
Land use characteristics at household location (post code unit level)	Area type: Rural areas; Urban areas 3,000-10,000 population; Urban areas 10,000-25,000 population; Urban areas 25,000-250,000 population; Urban areas >250,000 population; Metropolitan areas outside London; Outer London; Inner London.		Population density (persons/hectare): <1; 1-4.99; 5-9.99; 10-14.99; 15-19.99; 20-24.99; 25-29.99; 30-34.99; 35-39.99; 40-44.99; 45-49.99; 50-59.99; 60-74.99; ≥75.	

5. MAIN FINDINGS

As explained in Section 3, we first run a SEM that incorporates random intercepts with unrestricted variance (Model A). We then run a second SEM where the random intercepts are modelled as a function of household socioeconomic profiles; this is essentially a random intercept model with second level determinants (Model B). For comparison with the latest SEM model using the NTS data, we run an updated version of the model in Jahanshahi et al (*17*), using the same dataset for 2002-2012 – this single level SEM serves as a benchmark and we name it **Model Benchmark**.

We examine the model results in three steps. First, we examine the model goodness of fit. Second, we explore effects of introducing random intercepts for car ownership, travel distance and travel time through Model A. Third, we examine the random intercepts as a function of socioeconomic profiles of the households and differentiate the influences of self-selection and spatial sorting from those specifically associated with the built form categories. Finally, we examine any inter-temporal effects by examining the data in two periods – before and after 2007 – to see how the influences have evolved over time.

5. 1 Model goodness of fit

TABLE 2 compares the goodness of fit statistics of three alternative models. It shows that Models A and B fit much better to the observed data – the sharp reductions of AIC, BIC and ABIC are a little unexpected but a vindication of the introduction of the random intercepts. Modelling the random intercepts as a function of household socioeconomic variables continues to improve the fit, although the majority of gains are achieved when the random intercepts are introduced under Model A.

TABLE 2 Goodness of fit: Models A and B vs a Benchmark Model without Random Intercepts

	AIC	BIC	ABIC
Travel distance model			
Model A	1,624,785	1,625,511	1,625,266
Model B	1,624,510	1,625,338	1,625,059
Model Benchmark	2,472,702	2,473,785	2,473,419
Travel time model			
Model A	803,521	804,255	804,007
Model B	803,361	804,189	803,910
Model Benchmark	1,657,683	1,658,766	1,658,401

5. 2   The influences of household socio-economic profiles

First, we examine the influences of household socioeconomic profiles as estimated by the random intercept model relative to Model Benchmark.

TABLE 3 shows the influences of household socioeconomic profiles upon car ownership after controlling for the built form characteristics of the residential areas – the model coefficients (including the nonsignificant coefficients) are reported from the travel distance model, and the equivalent set of coefficients from the travel time model are reassuringly similar and not reported here. A large, positive coefficient indicates a strong influence for not owning or retaining regular access to a car in the household. The coefficients from Model A are in fact quite similar to the Model Benchmark. As with regression models for categorical variables, one category per set is left out so that the coefficient estimation can treat it as the reference category (which is reported in the right most column of TABLE 3)

TABLE 3 Influence of household socioeconomic profile on car ownership after controlling for built form categories

Household socioeconomic variables	Model A coefficient	Model Benchmark coefficients	Reference variable for model
Male	0.017	0.026	Female
Full time working	0.022	0.018	Part time working
1 adult households	0.478***	0.484***	>1 adult households
Manual workers	0.366***	0.373***	White collar clerical
Skilled manual workers	-0.030	-0.035	White collar clerical
Professionals	-0.150***	-0.135***	White collar clerical
Household income less £10k	0.545***	0.540***	Income 25-30k
Household income £10k to £15k	0.469***	0.476***	Income 25-30k
Household income £15k to £20k	0.340***	0.341***	Income 25-30k
Household income £20k to £25k	0.128***	0.134***	Income 25-30k
Household income £30k to £35k	-0.062	-0.065	Income 25-30k
Household income £35k to £40k	-0.111***	-0.102***	Income 25-30k
Household income £40k to £50k	-0.152***	-0.130***	Income 25-30k
Household income £50k to £60k	-0.238***	-0.225***	Income 25-30k
Household income more than £60k	-0.319***	-0.264***	Income 25-30k

*** significant with 99% confidence interval; as a rule we only report model coefficients that are significant with 99% confidence interval, because we have more than 90,000 individuals in the dataset and thus a large degree of freedom for within-level model. The same principle is applied throughout this paper.

TABLE 4 presents the influences of household socioeconomic profiles and car ownership upon travel distances and travel times for each of the three trip purposes from Model A. Again the Model A coefficients are generally very similar to those from the Model Benchmark. In TABLE 4 we present the coefficients in miles and minutes – e.g. a coefficient of 10.8 for variable 'Male' implies that all being equal a male worker commutes 10.8 more miles per week relative to females. Looking across the variables, it is not surprising that workers from lower income occupations and lower household incomes travel less. For commuting (Panel 4a), the most striking difference is between full- and part-time workers.

For shopping (Panel 4b), not having a car implies shorter travel distances (by 3.4 miles). Males on average travel 3.1 miles and spend 11.7 minutes less than females. Working full time implies less shopping travel, although the influence is well less than half the influence of gender.

For other travel (Panel 4c), not having a car imply shorter travel distance and less travel time. Males on average travel 14.6 miles and spend 15.4 minutes more than females.

TABLE 4 Fixed influences on travel distance and times arising from traveler profiles

Direct influence	Model A - Travel Distance (miles)	Model A - Travel Time (minutes)	Model Benchmark - travel distance (miles)	Model Benchmark - travel time (minutes)	Reference variable
Panel 4a. Direct influences on commuting					
Male	10.8***	12.4***	10.7***	12.7***	Female
Full time working	16.6***	41.9***	16.5***	41.7***	Part time working
1 adult households	3.2***	-1.6	3.2***	not significant	>1 adult households
Semi- or unskilled manual workers	-3.3***	-3.6**	-3.2***	-3.6***	White collar clerical
Skilled manual workers	-4.8***	-13.5***	-4.7***	-13.2***	White collar clerical
Professionals	2.4***	0.00	2.3***	not significant	White collar clerical
Household income less £10k	-1.7	-3.7	not significant	not significant	Income 25-30k
Household income £10k to £15k	-4.7***	-6.4***	-4.7***	-6.5***	Income 25-30k
Household income £15k to £20k	-2.2***	-0.7	-2.2***	not significant	Income 25-30k
Household income £20k to £25k	-0.6	-1.3	not significant	not significant	Income 25-30k
Household income £30k to £35k	0.8	0.6	not significant/	not significant	Income 25-30k
Household income £35k to £40k	2.0***	2.2	2.0***	not significan	Income 25-30k
Household income £40k to £50k	4.4***	5.4***	4.5***	6.3***	Income 25-30k
Household income £50k to £60k	5.0***	5.4***	5.1***	7.1***	Income 25-30k
Household income more than £60k	7.5***	9***	7.6***	11.8***	Income 25-30k
No car in household	-4.7***	19.6***	-4.4***	20.3***	With car in household
Panel 4b. Direct influences on shopping					
Male	-3.1***	-11.7***	-3.0***	-11.6***	Female
Full time working	-1.1***	-5.2***	-0.9***	-5.1***	Part time working
1 adult households	0.8***	2.8***	0.8***	2.8***	>1 adult households
Semi- or unskilled manual workers	-1.3***	-3.4***	-1.4***	-3.5***	White collar clerical
Skilled manual workers	-1.0***	-3.8***	-1.1***	-3.8***	White collar clerical
Professionals	0	-1	not significant	not significant	White collar clerical
Household income less £10k	-0.98***	0.1	-0.9***	not significant	Income 25-30k
Household income £10k to £15k	-0.91	0.4	-1.0***	not significant	Income 25-30k
Household income £15k to £20k	-0.78	-0.1	-0.8***	not significant	Income 25-30k
Household income £20k to £25k	0.25	1.9	not significant	not significant	Income 25-30k
Household income £30k to £35k	0.34	0.8	not significant/	not significant	Income 25-30k
Household income £35k to £40k	0.13	0.3	not significant/	not significant	Income 25-30k
Household income £40k to £50k	0.57	1	not significant/	not significant	Income 25-30k
Household income £50k to £60k	0.49	0.9	not significant/	not significant	Income 25-30k
Household income more than £60k	0.4	-0.2	not significant/	not significant	Income 25-30k
No car in household	-3.4***	-0.4	-2.7***	not significant	With car in household
Panel 4c. Direct influences on other purposes combined					
Male	14.6***	15.4***	14.3***	15.2***	Female
Full time working	2.1	-14.3***	2.4***	-14.5***	Part time working
1 adult households	20.8***	43.5***	21.0***	44.1***	>1 adult households
Semi- or unskilled manual workers	-19.7***	-43.8***	-19.7***	-44.6***	White collar clerical
Skilled manual workers	-18.8***	-44.1***	-19.2***	-44.9***	White collar clerical
Professionals	14.0***	20.0***	13.5***	19.5***	White collar clerical
Household income less £10k	-9.44***	-14.9***	-9.1***	-14.4***	Income 25-30k
Household income £10k to £15k	-10.87***	-17.9***	-10.9***	-18.1***	Income 25-30k
Household income £15k to £20k	-6.8***	-11.1***	-6.7***	-11.4***	Income 25-30k
Household income £20k to £25k	-3.92***	-4.4	-3.8***	not significant	Income 25-30k
Household income £30k to £35k	3.00	5.9	not significant/	not significant	Income 25-30k
Household income £35k to £40k	4.37***	9.9***	4.4***	10.0***	Income 25-30k
Household income £40k to £50k	7.60***	13.6***	7.7***	14.3***	Income 25-30k
Household income £50k to £60k	13.85***	22.3***	14.0***	22.9***	Income 25-30k
Household income more than £60k	24.0***	38.6***	24.2***	40.9***	Income 25-30k
No car in household	-22.0***	-28.1***	-19.1***	-22.7***	With car in household

*** significant with 99% confidence interval

The interactions between different purposes of travel are also very similar across the SEMs, showing significant negative influence of commuting time on shopping travel time as well as significant influence of commuting and shopping on other purposes, in the context of both travel distance and time.

The comparison of the model results with Model Benchmark shows that whilst Model A have become far more statistically robust after introducing the random intercepts, its actual coefficient values have only changed fairly slightly. However, Model A provides the foundation for exploring the variations across the built form clusters which we now turn to below.

5. 3   Self-selection vs intrinsic built form effects

The simple formulation of the random intercept model under Model A outputs the mean and variance of the intercepts. Model B defines the random intercepts as a function of household socioeconomic profiles. This is an effective way to control for self-selection and spatial sorting of the residents when quantifying the intrinsic influences of the built form.

TABLE 5 first presents the mean and variance of car ownership random intercept in Panel 5a. Panel 5b further reports statistically significant influences of the socioeconomic variables on the random intercept for car ownership, and the residual variance.

The substantial overall variance of the car ownership intercept in panel 5a (0.216) confirms that the level of car ownership varies across built form categories. Further, after controlling for the influences of socioeconomic profiles, the residual variance is considerably reduced – to 0.046 and 0.034 respectively for travel distance and travel time. The ratios of the residual variances to the overall variances (0.046/0.216= 21% and 0.034/0.216=16%) are the share of influence of the built form categories that is not explained by self-selection and spatial sorting of the households. In other words, 79% (1-21%) of the variance as shown in the travel distance model and 84% (1-16%) in the travel time model is explained by household socioeconomic profiles.

Out of the range of socioeconomic variables, the most statistically significant variables are the proportions of 1 adult households and of professional and skilled manual workers (panel 5b). In other words, around 80% of the observed outcomes of the much lower proportions of car owners residing in dense urban areas could be attributed to the fact that there are more 1 adult households living there who have a considerably lower levels of car ownership, and the fact that households with skilled manual and professional workers who have high van/car ownership tend to live in less dense, rural areas. We have already suspected this from the findings in Jahanshahi et al (*17*), but this is the first time that we are able to provide an unambiguous quantification of the effects.

TABLE 5 Between-built form variations in car ownership and the main underlying influences

Influences on intercept		Travel distance model	Travel time model
Panel 5a car ownership random intercept threshold and variance- Model A			
Intercept			
	Threshold	1.17***	1.17***
	Variance	0.216***	0.216***
Panel 5b Influences on car ownership random intercept- Model B			
Percentage of			
	1 adult households	6.00***	5.9***
	Skilled manual workers	-3.38***	-2.62***
	Professional	-1.99***	-1.55**
	Household income less £25k	Not significant	-1.20*
Residual variance		0.046***	0.034***

*** significant with 99% interval ,** 95%, * 90% -

TABLE 6 presents similar information for commuting (panel 6a), shopping (panel 6b), and combined other travel purposes (panel 6c). Within each panel, the upper parts (i.e. panel 6a-1, 6b-1, and 6c-1) shows the means and overall variances, meanwhile the lower parts present the statistically significant influences on the variations of the intercepts for the travel distance and travel time outcomes. The ratio of the residual variance from Model B to the overall variance from Model A implies the percentage of travel outcome variations across the built form categories that are not explained by the household socioeconomic profiles. The outputs from TABLE 6 show that after controlling for the household socioeconomic profiles, respectively 54%, 43% and 53% of the travel distance variations in commuting, shopping, and other travel can be attributed to intrinsic built form characteristics. Similarly, the percentages of travel time variations that can be attributed to built form characteristics are 75%, 43% and 77% respectively for commuting, shopping and other travel.

TABLE 6 Between level influences on travel

Influences on intercept		Travel distance model coefficient	Travel time model coefficient
Panel 6a modelling commuting random intercept			
Panel 7a-1 commuting random intercept mean and variance _Model A			
Intercept	Mean	7.8***	50.6***
	Variance	10.3***	330.5***
Panel 6a-2 Influences on commuting random intercept- Model B			
Proportion of	1 adult households	-35.06**	18.49***
	Skilled manual workers	Not significant	-6.38**
	Full time workers	-43.70***	Not significant
Residual variance		5.6***	246.5***
Panel 6b modelling shopping random intercept			
Panel 6b-1 shopping random intercept mean and variance-Model A			
Intercept	Mean	12.45***	48.7***

	Variance	6.5***	6.6***	***
Panel 6b-2 Influences on shopping random intercept-Model B				significan
Proportion of	Full time workers	-16.7***	Not significant	
	1 adult households	-14.6**	-31.1***	t with
	Skilled manual workers	22.7***	24.0***	
	Professional	17.36***	33.3***	99%
	Household income less £25k	Not Significant	21.2**	
Residual variance		2.8***	2.8***	interval
Panel 6c modelling Other purposes random intercept				
Panel 6c-1 Other purposes random intercept mean and variance-Model A				,** 95%,
Intercept	Mean	55.19***	192.1***	
	Variance	125.6***	133.0***	* 90% .
Panel 6c-2 Influences on other purposes' random intercept- Model B				
Proportion of	1 adult households	-121.4***	Not significant	S
	Professional	105.2***	140.9**	imilar to
	Household income more than £50k	-66.5***	Not significant	the
Residual variance		66.7***	102.3	influenc
				es on

car ownership model, the significant household socioeconomic influences on travel outcomes are 1 adult households, skilled manual workers, professional workers, etc, which tend to have substantial self-selection and spatial sorting tendencies.

5. 4 Within and between built form influences pre- and post-2007

We further extend the analyses through subdividing the NTS data into two subsets: 2002-2006 and 2008-2012. We exclude the year 2007 because the financial crisis had already crept in for some sectors but not others in the UK during that year. The purpose is to see if there had been any significant shifts in the influences over time both within and between built form categories.

The model estimation is carried out through a multi-group SEM where the influences are allowed to vary between the two time periods. This roughly doubles the number of unknown parameters to be estimated. To optimise the model runs, we run the multi-period model in two stages: First, we run the model only for car ownership (cf TABLE 7) to identify any highly significant between level socioeconomic variables. We then run the travel distance and travel time models for each travel purposes with the best between level car ownership model, (refer to TABLE 8 below).

Car ownership model shows no significant coefficient changes pre- and post-2007. However, the variance for car ownership intercept (between built form variation) has increased by 36% after 2007 (panel 7a). This corroborates the findings in Jahanshahi et al (*17*) and the Model Benchmark which show an increase of 22% in the effect of land use latent variable on car ownership. Controlling specifically for variations across built form categories in a multi-period model shows a stronger growth in forgoing car ownership in more dense and urbanized areas after 2007.

Panel 7b shows that this increased influence on car ownership in dense urban areas is not due to household socioeconomic profiles. In other words, the role of self-selection effect is minimal in explaining this increase, and the influences would appear to have come from changing built form characteristics.

15

TABLE 7 Summary of significantly changed influences pre- and post-2007 for between level model car ownership

Influences on intercept	Coefficients and (p-value) pre 2007	Coefficient and (p-value) post 2007	% change in coefficient values
Panel 7a car ownership free random intercept variance			
Intercept variance	0.196***	0.266***	36%***
Panel 7b Influences on car ownership random intercept			
Proportion of			
1 adult households	5.21***	6.41***	Not significant
Skilled manual workers	-2.87***	-3.10***	Not significant

*** significant with 99% interval ,** 95%, * 90% .

Travel distance and travel times show no significant between built form changes across the two periods. The analysis of within built form influences (cf TABLE 8) confirms the findings by Jahanshahi et al (*17*) and Model Benchmark. First, the gender gap appears to be narrowing after 2007 for travel distance and time. Our analyses of the NTS data show that for shopping trips, it is the females who have reduced their travel distance; for commuting and other trips, it is the males' reduced travel that narrowed the gap.

Second, in line with the trends above, the gap in travel distance between the low and the middle income group is no longer significantly different from 0 after 2007. This is because the rate of drop in travel distance for middle income group has been higher vis-à-vis that for the low income.

Thirdly, the gap in travel time of manual workers to work appears to have disappeared after 2007. This is a two sided effect involving on the one hand increases in manual workers commuting time and decrease in commuting time of other occupation groups.

As a final note, it is useful to note that the majority of the influences remain remarkably stable over time. For instance, the large differences between full- and part-time working in terms of commuting distance and time have not changed, in spite of the rapid rise in part-time and free-lancing work, and in the spread of ICT usage.

TABLE 8 Summary of significantly changed influences pre- and post-2007 (intra-built form level)

Influences	Coefficients and (p-value) pre 2007	Coefficient and (p-value) post 2007	% change in coefficient values
Panel 8a Travel distance analysis			
Male->HBW	11.47***	9.86***	-14%***
Male->Sh	-3.48***	-2.66***	24%***
Male-> Oth	16.49***	8.26***	-50%***
IncomeLess25k->HBW	-2.76***	-0.78	-
Skilled Manual->Sh	-1.42***	-0.48	-
Panel 8b Travel time analysis			
Manual->HBW	-4.47***	-1.61	-
Male->Sh	-14.27***	-9.61***	33%***
Male-> Oth	15.48***	1.23	-

*** significant with 99% confidence interval

6. CONCLUSION

In summary the new models have revealed fresh insights into the enhanced influences on travel by transformations in the built form in the UK since 2007. Overall, around 80% of the variations in car ownership levels can be explained by household socioeconomic profiles. However, the between built form variation in car ownership has increased by 36% post 2007, and this increased influence on car ownership arises principally from the transformations among some UK built up areas, particularly the dense cities. Furthermore, after controlling for the household self-selection and spatial sorting among the built-up areas, respectively 54%, 43% and 53% of the travel distance variations in commuting, shopping, and other travel can be attributed to the intrinsic built form characteristics in the UK. The percentages of travel time variations that can be attributed to built form characteristics are respectively 75%, 43% and 77%. Our new quantifications are built on more robust statistical theories and thus imply significant implications on how to interpret the impact of recent transformations in UK's inner cities. The results are to an extent corroborated by the Census and traffic data in the UK, which show that the UK inner cities have gained the momentum in population growth, whilst seeing car ownership and car traffic levels drop in most instances. Strictly speaking, it may still be too early to predict how such trends will play out next, but by working with an on-going travel survey like the NTS, this new SEM-Random Intercept method can be used to produce a regular and timely update on any shifts in the influences on travel, and provide feedbacks for land use planning and integrated built form/transport interventions.

REFERENCES

[1] Charlene, R., and J. Fox. Evidence review of car traffic levels in Britain: A rapid evidence assessment.In, RAND Corporation, Santa Monica, CA, 2014.

[2] Boarnet, M. G. The Built Environment and Physical Activity: Empirical Methods and Data Resource. Presented at Transport Research Board, 2004.

[3] Cao, X., P. L. Mokhtarian, and S. L. Handy. Cross-sectional and quasi-panel explorations of the connection between the built environment and auto ownership. *Environment and Planning A,* Vol. 39, No. 4, 2007, pp. 830-847.

[4] Silva, J. d. A. e., C. Morency, and K. G. Gouliasc. 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,* Vol. 46, No. 8, 2012, pp. 1252–1264.

[5] Handy, S., X. Y. Cao, and P. Mokhtarian. Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D-Transport and Environment,* Vol. 10, No. 6, 2005, pp. 427-444.

[6] Giuliano, G., and J. Dargay. Car ownership, travel and land use: a comparison of the US and Great Britain. *Transportation Research Part A: Policy and Practice,* Vol. 40, No. 2, 2006, pp. 106–124.

[7] Silva, J. d. A. e., T. F. Golob, and K. G. Goulias. Effects of Land Use Characteristics on Residence and Employment Location and Travel Behavior of Urban Adult Workers. *Transportation Research Record: Journal of the Transportation Research Board,* Vol. 1977, 2007, pp. 121-131.

[8] Van Acker, V., F. Witlox, and B. Van Wee. The effects of the land use system on travel behavior: A structural equation modeling approach. *Transportation Planning and Technology,* Vol. 30, No. 4, 2007, pp. 331-353.

[9] Van Acker, V., P. L. Mokhtarian, and F. Witlox. Car availability explained by the structural relationships between lifestyles, residential location, and underlying residential and travel attitudes. *Transport Policy,* Vol. 35, 2014, pp. 88–99.

[10] Mokhtarian, P. L., and X. Cao. Examining the impacts of residential self-selection on travel behavior: A focus on methodologies. *Transportation Research Part B-Methodological,* Vol. 42, No. 3, 2008, pp. 204-228.

[11] Cao, X., P. L. Mokhtarian, and S. L. Handy. Do changes in neighborhood characteristics lead to changes in travel behavior ? A structural equations modeling approach. *Transportation,* Vol. 34, No. 5, 2007, pp. 535-556.

[12] ---. The relationship between the built environment and nonwork travel: A case study of Northern California. *Transportation Research Part a-Policy and Practice,* Vol. 43, No. 5, 2009, pp. 548-559.

[13] Gao, S., P. L. Mokhtarian, and R. A. Johnston. Exploring the connections among job accessibility, employment, income, and auto ownership using structural equation modeling. *Annals of Regional Science,* Vol. 42, No. 2, 2008, pp. 341-356.

[14] Bohte, W., K. Maat, and B. van Wee. Measuring Attitudes in Research on Residential Self-Selection and Travel Behaviour: A Review of Theories and Empirical Research. *Transport Reviews,* Vol. 29, No. 3, 2009, pp. 325-357.

[15] Cervero, R., and J. Murakami. Effects of built environments on vehicle miles traveled: evidence from 370 US urbanized areas. *Environment and Planning A,* Vol. 42, No. 2, 2010, pp. 400-418.

[16] Handy, S., X. Y. Cao, and P. L. Mokhtarian. Self-selection in the relationship between the built environment and walking - Empirical evidence from northern California. *Journal of the American Planning Association,* Vol. 72, No. 1, 2006, pp. 55-74.

[17] Jahanshahi, K., Y. Jin, and I. Williams. Direct and indirect influences on employed adults' travel in the UK: new insights from the National Travel Survey data 2002-2010. *Working paper at latest review stage - Transportation Research Part A: Policy and Practice*, 2015.

[18] Bhat, C. R., and R. Gossen. A mixed multinomial logit model analysis of weekend recreational episode type choice. *Transportation Research Part B: Methodological*, No. 9, 2004.

[19] Spissu, E., A. R. Pinjari, C. R. Bhat, R. M. Pendyala, and K. W. Axhausen. An analysis of weekly out-of-home discretionary activity participation and time-use behavior. *Transportation,* Vol. 36, No. 5, 2009, pp. 483-510.

[20] Cho, S.-J., K. J. Preacher, and B. A. Bottge. Detecting Intervention Effects in a Cluster-Randomized Design Using Multilevel Structural Equation Modeling for Binary Responses. *Applied Psychological Measurement*, 2015.

[21] Dunn, E. C., K. E. Masyn, S. M. Jones, S. V. Subramanian, and K. C. Koenen. Measuring psychosocial environments using individual responses: an application of multilevel factor analysis to examining students in schools. *Prev Sci,* Vol. 16, No. 5, 2015, pp. 718-733.

[22] Marsh, H. W., O. Lüdtke, A. Robitzsch, U. Trautwein, T. Asparouhov, B. Muthén, and B. Nagengast. Doubly-Latent Models of School Contextual Effects: Integrating Multilevel and Structural Equation Approaches to Control Measurement and Sampling Error. *Multivariate Behavioral Research,* Vol. 44, No. 6, 2009, pp. 764-802.

[23] Marsh, H. W., A. J. Morin, B. Nagengast, and L. F. Scalas. Doubly latent multilevel analyses of classroom climate : an illustration. *Journal of Experimental Education,* Vol. 82, No. 2, 2015, p. 143.

[24] Jahanshahi, K., I. Williams, and X. Hao. Understanding Travel Behaviour and Factors Affecting Trip Rates.In *European Transport Conference*, 2009.

[25] Jahanshahi, K., Y. Jin, and I. N. Williams. Analyzing Sources of Variability in Travel Time Use in a Combined Framework Using Extended Structural Equation Models and UK National Travel Survey Data. Presented at Transport Research Board Annual Meeting, Washington DC, 2013.

[26] Jahanshahi, K. a. J. Y. 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,* Vol. 39, No. 1, 2015.

[27] Jin, Y., I. Williams, and M. Shahkarami. A new land use and transport interaction model for London and its surrounding regions.In *European Transport Forum*, Cambridge, 2002.

[28] WSP. Research into Changing Trip Rates over Time and Implications for the National Trip End Model: Final Report.In, 2009.

[29] Muthén, B. A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators. *Psychometrika,* Vol. 49, No. 1, 1984, pp. 115-132.

[30] Muthén, B., and T. Asparouhov. Growth mixture analysis: Models with non-Gaussian random effects.In *Advances in Longitudinal Data Analysis*, 2007. pp. 143-165.

[31] Morris, S., A. Humphrey, K. Pickering, S. Tipping, I. Templeton, and J. Hurn. National Travel Survey 2013 technical report.In, UK Department for Transport, 2014.