A response to the consultation document 'A City for All Londoners'

Can recent trendbreaks become real tipping points for a new kind of growth?

Ying Jin
Department of Architecture, Cambridge University (yj242@cam.ac.uk)

I am writing on behalf of the Cities and Transport Research Group at Department of Architecture, Cambridge University, in the capacity of an academic working in the field of planning and design, and a resident in London's commuting catchment.

We strongly support the vision of Good Growth for its new, ground-breaking approach across so many policy areas. We also appreciate its prescient acknowledgement of the recent economic, social and cultural shifts that could well help build the momentum for implementing Good Growth.

However, we think that a successful implementation of the Good Growth policies would yet require a careful sift of the emerging evidence regarding the nature and magnitude of the trend shifts in order to be clear where the biggest gains are in momentum building, and where the toughest issues are likely to emerge in the process. This will help enhance the understanding of the benefits and costs of the new interventions, and make a coherent case for specific investments.

The challenges

Our understanding is that the ultimate aim of Good Growth is to steer the growth in jobs, housing and services across London in order to benefit all citizens. This is to challenging status quo, rebalance London's economy, relieve congestion and housing crises in areas that are under strains and pressures of growth, and create new, attractive places for living and working in wider London and - by implication - in the wider South East. This is a powerful idea that can lead to a paradigm change in how cities are reshaped for growth, liveability and sustainability. We are delighted that London is leading the way through advancing this vision.

The main challenge to policy implementation, as we see it, stems from the fact that the visionary objective is breaking into a new policy territory where the evidence base and political consensus are not strong. The policy aim to benefit all citizens is currently open to a diverse range of interpretations.

The London Plan is a natural instrument to lead the development and implementation of planning and transport policies through translating the vision into coherent alternatives regarding investment projects and regulatory measures. This is done through a clear demonstration on the need for growth and where such growth needs to happen, so that the stakeholders can make informed political choices.

Since its inception the London Plan has facilitated extraordinary breaks from the historic trends: London's overall population size has seen a dramatic reversal from decades of decline to rapid growth, now reaching well above the historic peak; by contrast, the overall road travel and traffic has also broken the year on year rises; in many areas car ownership has declined as the number of residents grow and their incomes rise; rail, Tube, buses, cycling and walking have regained their prominence in supporting and improving access to jobs, services and leisure; in a number of built-up areas, not least the CAZ, environmental amenities have improved as population and building densities increase. Those trendbreaks are the living proof of the effectiveness in the jointed up thinking in the London Plan, and they form the foundation for pursuing Good Growth.

Nevertheless, to achieve the objectives of Good Growth would imply confronting many hard issues that hitherto do not have set solutions, and few examples to copy from – the development initiatives will need to invent business models that do not yet exist. It is in such ventures that a more structured understanding of the recent successes in trendbreaking that would come useful, in gauging potential demand for infrastructure, as well as impacts upon the local economy, communities and the environment.

For instance, the London Plan has been very successful in promoting the integration of transport, land use and urban design, particularly through promoting growth in areas of good public transport activity. This approach has been very effective in regenerating central and inner London sites on the radial routes, and in the process reshaping the demand for travel and manage road traffic. The new Good Growth vision calls for spreading economic growth and jobs as well as securing affordable housing beyond the current high growth areas. To what extent will the approach to transport/land use integration and travel demand management measures work in areas of lower growth (e.g. much of south London), or along orbital corridors? How could the patterns of sustainable travel be extended from the successful areas into the more challenging opportunity areas in the outskirts?

A second aspect where more in-depth understanding is needed is the underlying background trends of work and lifestyles. For example, the high end of services is expanding and it can better exploit opportunities of tele-working. What impacts could there be if a growing percentage of commuters for CAZ can work in the outer London Boroughs one or two days a week upon the local economy, high streets, town centres, transport infrastructure, etc?

The most important aspect is of course to understand the knock-on effects and the trade-offs people will make. For example, on the demand side, how do people trade off their access to the buzz of business clusters (which would imply central sites and long commuting journeys) against flexible leases and cheap floorspace rents (which are vital to start-ups but disappearing fast in high growth areas); on the supply side, how do development proposals on parts of the Green Belt with easy access to fast public transport compare with intensification within existing suburban neighbourhoods?

For land use and transport analysts and modellers like us in the research group, it would seem that there are two major gaps in our current understanding in this context:

(1) How to translate the past London Plan successes in trendbreaking into an in-depth and quantified understanding of how integrated transport, land use and urban design interventions influence the choices of citizens and businesses, and apply this understanding to the specification and assessment of investment proposals in the many new situations in wider London where there are few best examples or set solutions.

(2) How to account for the complex knock-on effects and trade-offs that major transport or land use interventions will bring about. This is not just about unintended consequences – History of cities including that of London show that in the right circumstances the knock-on feedbacks (such as urban development responding to rail/Tube lines since the Victoria era) could engender societal benefits a magnitude higher than the direct returns anticipated by investors.

We see emerging opportunities to address Item (1) above through year-on-year monitoring using novel analytical methods and Item (2) through a new, city-region-wide simulation model. Since we are a leader in the respective fields we will briefly comment on our potential contributions to supporting the new London Plan.

Understanding and monitoring the trendbreaks

Trendbreaks by definition are hard to capture, understand and predict. Existing analyses of trend-breaks tend to focus on changes in overall patterns of e.g. population, travel demand, jobs, etc. This means that when such trend-breaks are identified in aggregate, the particular segments and areas where changes occur may have already undergone fundamental shifts. The timelag in detection would delay policy responses and narrow the time-window for effective intervention. Besides, aggregate analyses are often unable to predict forthcoming changes. More crucially, any incipient, new ‘trends' are often not a foregone inevitability, but contingent upon dynamics among societal changes, policies and regulation. It is therefore important not only to detect the signs of change, but also to monitor them over time. The London Plan analyses have already carried out analyses on population and here I provide two recent examples of our analyses on travel demand and jobs.

<u>Trendbreaks in car ownership and travel demand</u>

We have been working to detect and interpret early signs of trend-breaks in car ownership and travel demand first through a PhD project (see papers [a], [b], and [c]) and currently through a DfT Transport Research Innovation Grant at the UK level. For this analysis we develop a novel extension to the structural equation model (see Appendix A) which makes it possible for the first time to fully exploit the richness of recorded data in the UK National Travel Survey (NTS) regarding land use, built form, personal and household characteristics, car ownership, etc, and tease out the inter-dependencies among the personal and households profiles, their residential and job locations and the associated built environment characteristics, and their travel choices for all purposes and all travel outcomes.

This new approach can rigorously quantify the influences upon car ownership and travel choices of self-selection and spatial sorting, i.e. individuals and households being pre-disposed towards specific built-forms or car ownership status. So far our studies are carried out for the UK as a whole, but some of our findings highlight the potential to monitor London more closely, particularly the on-going transformation in its built environment. For instance, Paper [a] shows in dense urban areas, the built form characteristics are an increasing influence on car ownership and travel demand over the 2000s and early 2010s. Paper [b] reveals remarkably different built form influences upon the travel choices of the resident population living in dense, suburban and rural environments. Paper [c] the built form characteristics could potentially account for up to 50% of the influence upon travel distance and travel time once the self-selection and spatial sorting effects are accounted for.

In particular, the papers show that car ownership is neither a simple dependent nor independent variable concerning residential or travel choices - instead it is an endogenous variable that is interdependent with personal and households profiles, consumer choices and the built environment. This highlights the need to reconsider how car ownership is predicted, given the profound influence of car ownership on road traffic. By contrast, the state of the art in car ownership modelling in the UK (including though not limited to the most recent version of NATCOP, used in the National Transport Model) largely ignores such interdependencies.

Since the on-going London Travel Demand Survey (LTDS) data series has a similar data structure to the UK NTS, there is considerable potential for London Plan to consider adopting a similar SEM approach to detecting and monitoring the influence of resident socioeconomic profiles, land use and the built environment on car ownership and travel demand. The LTDS can be available at a finer geographic level than the NTS, thus providing the possibility of categorising the local areas in much greater detail, and thus detecting the evolution of land use and built form transformation across London. The SEM analysis can identify trendbreaks in travel in different areas and population groups, target help for economically and socially disadvantaged travellers, pinpoint new investments, and design effective measures to manage car use, especially in the outer areas and orbital corridors where the London Plan seeks to transform in the coming years.

Trendbreaks in the distribution of jobs and employment

We have been leading an analysis on detecting new trends in the distribution of jobs and employment through an on-going consultancy project for the Greater Cambridge-Greater Peterborough (GCGP) LEP, in collaboration with Dr Andy Cosh of the Judge Institute and Prof Peter Tyler of the Department of Land Economy at Cambridge, and LDA-Design Ltd. Compared with London, the GCGP LEP area has a similar, if not more pronounced, differentiation in the distribution of jobs and employment. Here we highlight the potential to exploit emerging data sources to carry out analyses at three different levels.

Our approach to the mapping of jobs and employment has often been integrated with advanced spatial equilibrium modelling, e.g. for interpreting and predicting commuting patterns (see next section). However, the emergence of new data sources, e.g. those associated with the ONS Real Time Census project and the VAT database have broadened the prospects and granularity when mapping the overall patterns of jobs and employment, particularly the interconnections among businesses.

In contrast with population data, jobs and employment are much harder to map. This is particularly because many of the new growth sectors are not well defined by the SIC codes. However, to achieve the Good Growth objectives it would be vital to have continuous monitoring of the evolving patterns of jobs and employment, particularly in the most dynamic sectors of industry.

The analytics and modelling for medium to long term predictions

The London Plan has been exemplary in joining up decisions regarding jobs, employment, population, housing, commercial floorspace, travel, the natural environment and the milieu for culture and innovation. Given the new challenges, particularly with respect to the policy objectives to steer economic and job growth and spread prosperity, we recommend that a greater emphasis be placed on the analytics and modelling that support this jointed up thinking, particularly in providing the evidence for the business case of new investment decisions.

We have a long tradition of building and using precision land use, built form and transport forecasting models for infrastructure investment and urban development that are based on rigorous measurements of prices, rents, wages and consumer utilities. Their spatial economic foundation provides the basis for its prominent role in developing the business cases in investment decisions in high, mid and low income countries[1]. In the London area, we built and run a forecasting model for DfT and ODPM in their Wider South East Regional Study (2003-2005), which has predicted the near doubling of rail travel over the 1997-2016 period. The Cambridge Futures studies[2] used a sister model which has made specific assessments of seven scenarios (minimum growth, densification, new towns, transport links, greenbelt swap, necklace of villages, virtual highway) and a convincing case for policy change, which has since reshaped the patterns of development in the Cambridge subregion.

Our latest model suite is LUISA (see paper [d], [e] and [f]). It is designed to investigate how changes in economic and social policies, land use planning and infrastructure service operations affect one another against background trends in global trade, production technology, demographics and consumer behaviour. In particular the model predictions are location and year specific at a suitable spatial and temporal granularity to provide insights for informing the impacts and phasing of investment, regulation, pricing and community action plans. Given explicit assumptions of the background trends and policy designs, it aims to answer questions like

- how changes in economic and social policies affect the location of jobs
- how changes in job location affect local services, housing demand, commuting, business travel, movements of goods

[1] For examples of such work, see UK Research Excellence Framework 2014 case study http://impact.ref.ac.uk/CaseStudies/CaseStudy.aspx?Id=23292
[2] http://www.cambridgefutures.org/

- how the location of jobs and residents affects the demand for land and infrastructure services (e.g. multimodal transport services, energy, water, waste), and how the supply of land and infrastructure shapes future location patterns.
- where the future policy hotspots are, in terms of congestion and bottlenecks in growth areas, and deficiencies and social exclusion in declining areas.

Accordingly LUISA has two components: an **inter-temporal module** for updating the background trends and the supply of land, business premises, housing and infrastructure services over a period (typically 5-10 years, in line with the Mayoral election cycles) and a **spatial activity and infrastructure demand module** for modelling how employers and residents adapt their choices to the available supply of land, business premises, housing and infrastructure services at a given year. A typical model application first involves a calibration of both modules for one or more base years (for instance in the model for London and the wider south east the base years are 1991, 2001 and 2011), where the model is tested to see how well it can replicate the observed patterns of location and infrastructure demand both at a base year and over time. For a future year test, the inter-temporal module is first run (for e.g. 2021), which is followed by the spatial activity and infrastructure demand module for that year. This can then be repeated for further policy horizons (2031, 2041, etc).

The **main inputs** for the modules are:

- Background trends including GDP and foreign trade projections, demographic and immigration assumptions including national level population, employment, residents socio-economic profiles, technical progress as embedded in baseline inter-industry input-output coefficients, base year consumer price elasticities and household utility functions
- Base year stock of land use by use category, stock of business premises and housing by category. Floorspace data is often incomplete in past studies and it can be estimated in LUISA through observed workplace and residents data from workplace and population censuses
- Future land use policies including restraints and allocations by location
- Travel costs and times including any congestion and local access between all locations in the study area for both base and future years, for all main travel options. This would include mobility as service and ride-sharing options for future scenarios.
- Unit costs of other infrastructure services, e.g. energy, water, waste, etc.

Alternative employment growth patterns, demographic projections etc are treated as variations in modelled policy/technology scenarios.

The **main model outputs** for each modelled year are

- Changes in urban land use, business premises and housing stock by type in each model zone
- Jobs and residents by category by model zone
- Production output and prices by zone
- Annual rents for housing and business premises by zone
- Wages and disposable incomes by zone

- Accessibility to jobs and services by zone
- Resident utility as a measure of economic well-being for each socio-economic group
- Commuting, education, business and other journeys between zones, using all means of travel including walking and cycling
- Travel demand with respect to transport corridor capacities – this may require running the LUISA model in tandem with a multimodal transport model
- Demand for other infrastructure services.

The model outputs facilitate pair-wise comparisons which lead to economic assessment through levels of output, jobs, prices, rents, and producer and consumer surplus by zone and model year. The segmentation of workers and residents according socio-economic profiles enables analyses of fairness in distribution and social inclusion. The forecast demand for land, floorspace, travel and other infrastructure services may be input into environmental assessment models for resource efficiency and sustainability analyses.

LUISA can fill a gap between the macro regional economic models and micro-economic models on the one hand and the conventional land use/transport interaction and road traffic model on the other. Because it is based on rigorous economic formulations, it can interface with the macro and micro models, and because it is a geographically detailed model, it can interface with the land use/transport interaction and road traffic models. Its main role is to predict and assess the financial, economic, social and environmental costs and benefits for specific infrastructure and regulatory interventions.

Summary

We strongly support the vision of Good Growth and would wish to contribute to the development of a firmer evidence base, particularly in view of the new challenges.

Since its inception the London Plan has facilitated extraordinary breaks from the historic trends: London's overall population size has seen a dramatic reversal from decades of decline; the overall road travel and traffic has also broken the year-on-year rises; in many areas car ownership has declined as the number of residents grow and their incomes rise; sustainable travel has regained its prominence in supporting and improving access to jobs, services and leisure; in a number of built-up areas, not least the CAZ, many environmental amenities have improved as population and building densities increase. Those trendbreaks are the living proof of the effectiveness in the jointed up thinking in the London Plan, and they form the foundation for pursuing Good Growth.

Can the recent trendbreaks become real tipping points for a new kind of growth? We think that they can, if there is an in-depth understanding and careful monitoring of the new trends. In other words, a successful implementation of the Good Growth policies would greatly benefit from a careful sift of the emerging evidence regarding the nature and magnitude of the trend shifts in order to be clear where the biggest gains are in momentum building, and where the toughest issues are likely to emerge in the process.

In line with our specialisms we would propose to contribute in

(1) the identification and monitoring of the trendbreaks (especially in terms of the influence of land use and the built environment on travel demand across London, and the evolution in the interconnections of London's businesses), and

(2) advanced land use, built form and travel demand modelling, with which we examine the business case for the major investment and regulatory projects.

Bibliography

[a] Jahanshahi, Jin & Williams (2015) Direct and Indirect Influences on Employed Adults' Travel in the UK. Transportation Research Part A. http://www.sciencedirect.com/science/article/pii/S096585641500227X;

[b] Jahanshahi & Jin (2016). The built environment typology in the UK and their influences on travel behaviour. Transportation planning and technology. http://www.tandfonline.com/doi/full/10.1080/03081060.2015.1108083;

[c] Jahanshahi & Jin (2016). Trendbreaking influences of built form upon travel in UK cities. Transportation Research Record. https://trid.trb.org/view.aspx?id=1393764#

[d] Echenique, MH, V Grinevich, AJ Hargreaves and V Zachariadis (2013). LUISA: A Land-Use Interaction with Social Accounting Model; Presentation and Enhanced Calibration Method. Environment and Planning B: Planning and Design. (http://epb.sagepub.com/content/40/6/1003.abstract)

[e] Jin Y, M Echenique and A Hargreaves (2013). A recursive spatial equilibrium model for planning large-scale urban change. Environment and Planning B: Planning and Design (http://epb.sagepub.com/content/40/6/1027.abstract)

[f] Wan, L and Y Jin (forthcoming). Historical Validation of the Recursive Spatial Equilibrium Model – An Empirical Application for Beijing, Tianjin and Hebei City Region. Paper prepared for Journal of Regional Science.

Appendix A: an example of SEM analyses – this is Paper [b]

Appendix B: an outline design of the dynamic spatial equilibrium model – this is paper [e]

Transportation Planning and Technology

ISSN: 0308-1060 (Print) 1029-0354 (Online) Journal homepage: http://www.tandfonline.com/loi/gtpt20

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

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

Published online: 25 Nov 2015.

TRANSPORTATION PLANNING AND TECHNOLOGY, 2016
VOL. 39, NO. 1, 59–77
http://dx.doi.org/10.1080/03081060.2015.1108083

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 inventions and guidance for land use planning and urban design. The statistical analysis is

CONTACT Kaveh Jahanshahi kj299@cam.ac.uk

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 interdependencies[1] 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 traverllers' 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.

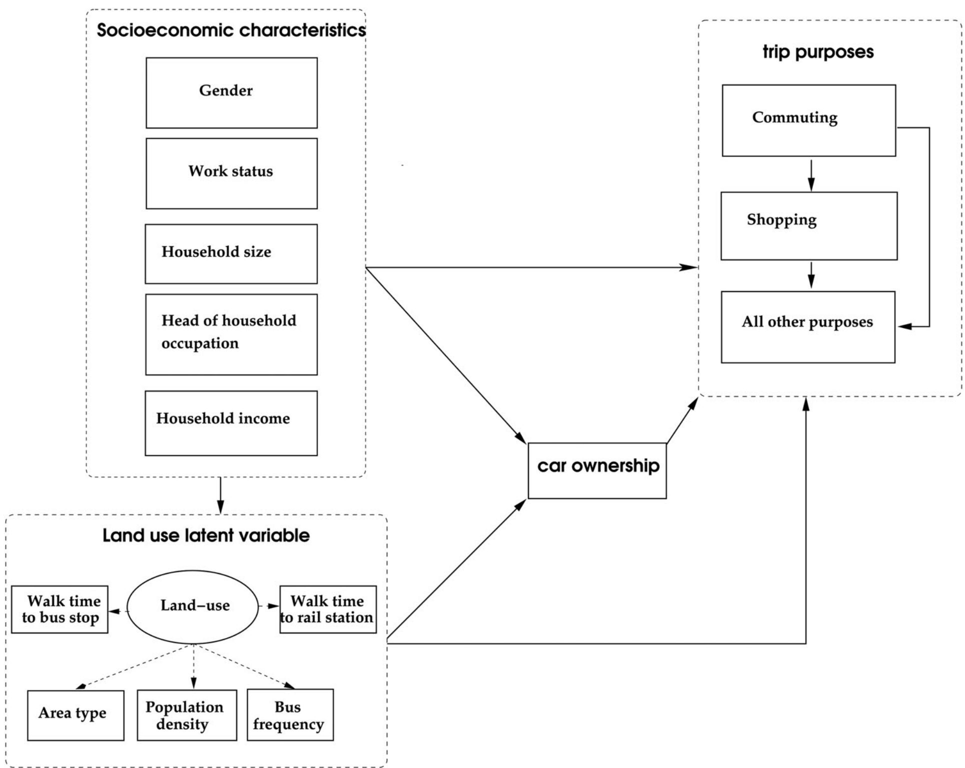

Figure 1. The conceptual structural equation model (SEM) for influences on travel.

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 intercorrelations 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 jth 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

$$Y_{ij} = s|C_i = c \Leftrightarrow \tau_{cj,s} < Y^*_{ij} < \tau_{cj,s+1} \tag{1}$$

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

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

$$Y^{*}{}_{ij}|_{C_i=k,x_i} = \nu_{kj} + K_{kj}X_i + \varepsilon_{ij} \tag{2}$$

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:

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

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

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

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

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

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

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

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

$$Y_{ij} = \nu_j + \mathrm{K}_j \mathrm{X}_{ij} + \epsilon_{ij} \tag{7}$$

where Y_{ij} refers to the ith respondent and jth 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 jth 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

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

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

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

$$\sum_{ij} \log\left(f_{ij}(Y_{ij})\right) \tag{9}$$

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 ≥99% confidence interval (i.e. the respective p-values are ≤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	Individual related variables – gender [male, female], work status [full time (FT), part time (PT)]
Journey	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.

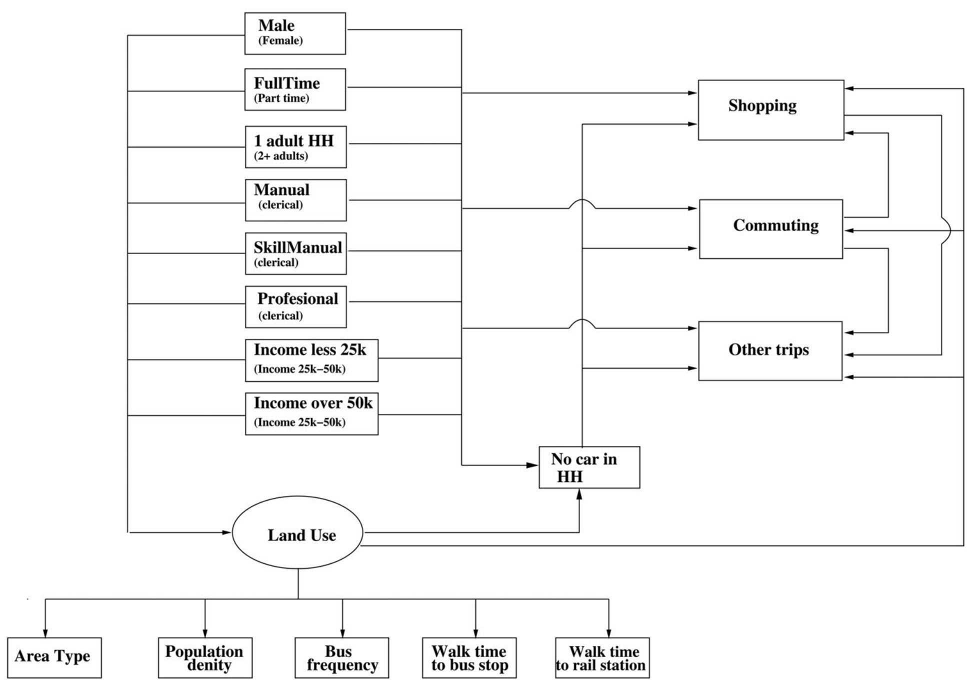

Figure 2. The SEM structure for testing the NTS data.

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 ($N = 20301$)
Panel 4a: Unconditional probabilities			
	0.18	0.54	0.27
Panel 4b: Conditional probabilities			
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.

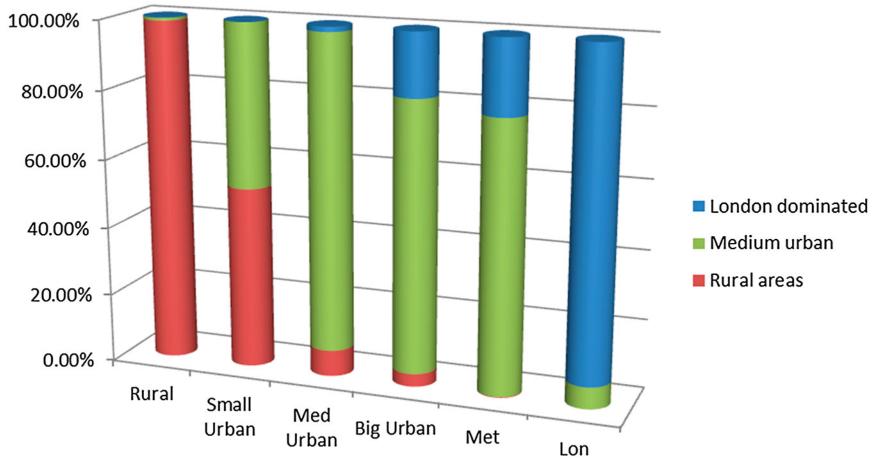

Figure 3. (Colour online) Composition of built environment latent classes by NTS land use area type.

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	007	−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 unfailing 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.

Environment and Planning B: Planning and Design 2013, volume 40, pages 1027 – 1050

doi:10.1068/b39134

A recursive spatial equilibrium model for planning large-scale urban change

Ying Jin, Marcial Echenique, Anthony Hargreaves
Martin Centre for Architectural and Urban Studies, Department of Architecture, University of Cambridge, 1–5 Scroope Terrace, Cambridge CB2 1PX, England;
e-mail: yj242@cam.ac.uk; me15@cam.ac.uk; ajh91@cam.ac.uk
Received 3 August 2012; in revised form 25 April 2013; published online 19 October 2013

Abstract. This paper presents a recursive spatial equilibrium model for urban activity location and travel choices in large city regions that anticipate major development or restructuring. In the model, producer and consumer choices that adjust quickly to stimuli reach temporary equilibria subject to recursively updated activity churn, background trends, estate development, and transport supply. The city region's performance at each time horizon affects the recursive variables for the next. The model builds on field leaders of urban general equilibrium, spatial interaction, and nonequilibrium dynamic models, and offers theoretical and practical improvements in order to fill an important gap in long-range urban forecasting. Linking the equilibrium and nonequilibrium models enables the simulation of path dependence in urban evolution trajectories that neither could produce in isolation. At the same time the model provides quantification of impacts of different policy interventions on a consistent basis for a given time horizon. The model is tested on the main archetypal urban development strategies for large-scale development and restructuring.

Keywords: land-use and transport model, infrastructure investment, travel demand forecasting, spatial equilibrium, recursive dynamics, urban restructuring, urban futures

1 Introduction

The 21st century as an 'urban century' has started to witness urban development and restructuring that are unprecedented in nature and scale. Over the next thirty to forty years accelerated urbanisation and lifestyle changes in the emerging economies are expected to lead to city building of a magnitude hitherto unseen in human history (UN Habitat, 2008); in countries that are already urbanised, some cities are still growing strongly. Numerous existing cities face challenges of restructuring and retrofit to tackle productivity growth, urban poverty, energy inefficiency, high per capita resource use, environmental degradation, and aging of citizens (Batty, 2010; Wegener, 1982; 2011). Bolder interventions have been called for (Fiorello et al, 2006; Wegener, 2011). Large-scale urban change may result from major new growth or restructuring.

Evolution in the governance of cities has cast a new light upon growth and restructuring. In addition to existing powers of land-use planning and regulation, municipal governments are often offered responsibilities for infrastructure investment, major transport and urban service operations, and ultimately attracting inward investment. For instance, such powers have been gradually decentralised to the municipal level in China since the 1980s (Lin and Liu, 2000); the on-going implementation of the 2010 election pledge (The Conservative Party, 2010) in the UK is a prominent example among the developed countries.

Since the 2008 financial crisis, productivity growth has jumped to the top of the policy agenda across the world's municipalities. Under tight public finance, productivity growth holds the key to social and environmental policies since the large investments required ultimately have to come from increased per capita output.

This modifies the context for computer modelling that supports municipal decision making. It is instructive to review the experience in London, a UK city which has over the years seen a fair share of active development and use of computer models for major policy decisions on both new development and restructuring.

1.1 Policy concerns versus unmet modelling needs

In any city the most relevant policy concern is the viability to fund (electoral) commitments to local constituencies. This was true even before the financial crisis. Volterra and CBP (2007, page 47) provide an insight into the unmet modelling needs in London, particularly concerning "the links between productivity, wages and rents and the full implications of these for output growth". They go on to list the unanswered questions as: What are the behavioural responses to overcrowding and to new transport availability? What are the effects of co-location and clustering of different firms, and do these vary among industries? What are the trade patterns and how do they change? How can we test that the models we use reflect the world in which we operate?

Decentralised decision making strengthens the above concerns. Local authorities are focused on the 'business case' of any intervention and feasibility under financial and fiscal constraints. Since any assessment of a large development proposal will be subject to debate, the models must be transparent and empirically robust (Rosewell, 2011). The criticism is that methods for assessment (eg, of transport investments in the UK) are "unconnected to the real economy" (Wenban-Smith, 2011). Similarly unmet policy needs are apparent across the OECD (OECD, 2012). In the developing countries our experience shows that the policy concerns are similar, but the modelling tools remain unavailable in most cities.

It seems that it was not technical complexity of models per se that deterred policy applications. For instance, the aspiration for identifying 'the full implications' of productivity, wages, and rents shows that there is a genuine appetite for general equilibrium modelling. However, large urban models are seen as 'black boxes' by critics (Lee, 1973) as well as modellers (Eliasson and Mattsson, 2001), and users often avoid the large models, even if that means reduced form rather than general equilibrium modelling (DfT, 2006; Volterra and CBP, 2007). Short mayoral election cycles and the need to face the public call for quick turnaround and transparency.

The world after the financial crisis does not seem to have fundamentally altered the key modelling questions. Rather, the need to understand drivers to productivity and offer practical insights to policy making are highlighted. This means that the models need to operate in the world of markets, prices, finance, budget constraints, physical and institutional inertia, individual behaviour, and their combined impacts.

1.2 Existing modelling methods

It is useful to contrast user needs with what is already available for policy modelling. Such models sprang from many different fields and disciplines, and they are far from paradigm convergence (Batty, 2009).

Given the traditional emphasis on land-use and transport planning, the main urban models in policy use since Lowry (1964) are built on spatial interaction models (Batty, 1976; Wilson, 1967). Effective and practical models have been created for assessing property development and transport options at detailed geographic scales through a close integration of the spatial interaction model with random utility theory (McFadden, 1974), national/regional input–output tables (Leontief, 1986), land-use and floorspace stock market models (Echenique, 2004; Echenique et al, 1969), transport demand forecasting (Ben-Akiva and Lerman, 1985; Daly and Zachary, 1978; Domencich and McFadden, 1975), road traffic assignment

(Sheffi, 1985), GIS and big data analyses (Batty, 2010; Batty et al, 2013). Their strengths lie in the explicit incorporation of planning and infrastructure constraints and the incorporation of policy inputs over explicit time horizons. However, those models rarely address endogenous productivity growth or urban dynamics.

A second strand of models investigates general equilibrium of the spatial economy. The relationships between the economy, activity location, and transport costs have been a focus of new economic geography (Fujita, 1989; Fujita et al, 1999; Krugman, 1991; Venables, 1996) and of spatial general equilibrium models (Anas and Kim, 1996; Anas and Liu, 2007; Bröcker, 1998; Ivanova and Tavasszy, 2007; Oosterhaven et al, 2001). Those models are focused on the effects of spatial costs on producers and consumers whilst giving a fuller representation of product varieties and economies of scale. Some models account for urban agglomeration and related productivity effects. Significant progress has been made in empirical model estimation (Redding, 2010). Production, trade, transport demand, and location are endogenously and mutually determined at spatial general equilibrium. Although, like the spatial interaction models, they can be used for discrete time horizons, existing spatial equilibrium models in their published form tend to focus on the end state rather than on the trajectories leading to the equilibrated state. Anas and Liu (2007) have introduced a dynamic property development sector within a general equilibrium model with exogenously determined total size of the city and of development. Dynamic general equilibrium models that represent intergeneration linkages and forward-looking behaviour have been at an exploratory stage (see Bröcker and Korzhenevych, 2011) or on the longer term research agenda (Anas, 2013).

A third strand of models is focused on urban dynamics, which are either represented in the aggregate (Allen, 1997; Forrester, 1969; Simmonds, 2001; Wegener, 2001; Wilson, 2000; Zondag and de Jong, 2011) or at a microlevel through cellular automata, agent-based models, and other forms of microsimulation (Batty, 2005; Chapin and Weiss, 1968; Clarke, 1996; Ingram et al, 1972). Microlevel dynamic models have been developed for land-use activities (UrbanSim, 2011; Waddell, 2002) and traffic flows (Nagel et al, 1999). They offer insights into microscopic interactions among agents, particularly in property development and traffic management. They also introduce physical inertia explicitly. However, they are predominantly used for investigating mechanisms and system-level emergence of microscopic interactions rather than for policy analysis (Batty, 2009), with a few exceptions such as those models developed by Wegener (2001), Simmonds (2001), Zondag and de Jong (2011), and UrbanSim (2011) which have been used for policy studies. A prominent feature of the applied models is their disregard for market equilibrium (Simmonds et al, 2013). It is clear that the needs of policy analysis will be better served if the model features could be applied across paradigms.[(1)] In particular, policy making requires not only insights into interdependencies at any point in time but also into how cities evolve.

In summary, cities facing major growth and restructuring would require planning models that can examine (1) implications of planned intervention on productivity, wages, and rents over policy horizons that relate to tenure lengths of mayoral offices; (2) effects of planning, building, and infrastructure constraints which are dominated by inertia and take decades to reach any equilibrated state if ever; and (3) dynamics of people and investment in response to prices, productivity, and citizens' well-being. In addition, such models should (4) be built upon technical data that most cities already have, such as censuses, input–output tables, urban traffic models, travel behaviour surveys, and any emerging big data. So far as we are aware,

[(1)] Where such progress has been made, the results are promising: for example, in linking spatial interaction and general equilibrium modelling (de la Barra, 1989; Echenique, 2004), cellular automata with input–output modelling (eg, White et al, 2000), or incorporating principles of microsimulation within aggregate urban land-use activity and stock modelling (Simmonds et al, 2013; Wegener, 2001).

no models currently meet the above four requirements simultaneously, least of all in those emerging economies that matter the most to world poverty alleviation and sustainability.

1.3 Aims of this paper

The aims of this paper are (1) to present the design of a new, generic model that starts to incorporate the above four requirements simultaneously for practical urban applications; and (2) to test it on a wide range of archetypal development scenarios for insights into fundamental model assumptions, roles of key parameters, and the added value of the new method. The tests help to set a prioritised research agenda for empirical implementation for assessing individual projects and policy initiatives.

The paper provides a summary of the model and tests for model users whilst addressing the key concerns of specialist modellers. More specialist material on equations, data, algorithm, and tests are presented in a supplementary working paper (Jin et al, 2013).

2 Model design

We consider each model component in turn before linking them together. Key concepts are reviewed where the context requires but space does not allow a literature survey—for such surveys see Wegener (2005; 2011), Hunt et al (2005), Iacono et al (2008), and Batty (2009).

As it is a spatial model, locations are defined as discrete and contiguous zones; the model divides the world into two categories of zones: 'internal' ones that represent areas within a city region;[(2)] and 'external' ones for the city region to trade with and to exchange migrants, supercommuters, and investment funds with.

2.1 Components for a new model

We follow a widely shared convention between spatial interaction and general equilibrium models and classify the economy into *producers* which include private, public, and voluntary businesses; and *final consumers* which include households, governments, collectives, investors, and exports. We further follow that convention and consider *trade in labour, goods, and services* between locations which is determined simultaneously with prices at market equilibrium, subject to idiosyncratic circumstances. We follow nonequilibrium dynamic models and define the stock of existing urban activities, buildings, transport infrastructure, and land as *stock constraints* which may be updated periodically subject to background trends, inertia, investment, and planning regulations. Finally, we consider how *boundary conditions*—such as business relocation and household migration between internal and external areas and cross-boundary investment—occur subject to prices, physical constraints, citizens' well-being, and idiosyncratic circumstances.

For simplicity, when the model components are discussed for one period only the time period subscripts t, $t + 1$, etc, are omitted; to account for flows of money (eg, production, consumption) and effort (eg, hours of labour, utility gains) all such quantities are defined in annual units unless noted otherwise.

2.1.1 *Producers*

The producers are represented by a set of production functions that define how they use capital, labour, properties, raw materials, and services, particularly how their input choices and productivity change with prices and externalities. A nested Cobb–Douglas–constant elasticity of substitution (CD-CES) function has been broadly accepted as a standard for this purpose in spatial general equilibrium analyses since Krugman (1991) and Fujita et al (1999). We follow Anas and Liu (2007), who developed a leading urban general equilibrium model, and define

[(2)] This is usually a reasonably self-contained area for daily commutes.

the production function as a variant of their CD-CES specification:

$$X_j^n = E_j^n A_j^n (K_j^n)^{v^n} \left(\sum_w (L_j^{wn})^{\theta^n}\right)^{\frac{\delta^n}{\theta^n}} \left(\sum_k (B_j^{kn})^{\zeta^n}\right)^{\frac{\mu^n}{\zeta^n}} \prod_m (Y_j^{mn})^{\gamma^{mn}} , \qquad (1)$$

where X_j^n is the output of industry n in zone j. The main inputs to production are capital K, labour L, buildings B, and intermediate inputs Y; and the function implies constant internal returns to scale of production through specifying the sum of cost share parameters for the respective input groups, $v^n + \delta^n + \mu^n + \sum_m \gamma^{mn} = 1$. For w varieties of labour and k varieties of buildings, a CES function is used to represent the substitution effects within each input, the elasticities of substitution being governed by parameters θ^n and ζ^n . A_j^n is a function of the economic mass for producer n in zone j that represents Hicksian-neutral total factor productivity effects resulting from learning and transfer of tacit knowledge (Graham and Kim, 2008; Rice et al, 2006), which are an important component of urban agglomeration effects. We define $A_j^n = \underline{A}_j^n (M_j^n / \underline{M}_j^n)^{\pi}$, where $\underline{A}_j^n$ is a constant representing baseline economic mass effects; M_j^n is a function of the economic mass accessible by producer n from zone j; $\underline{M}_j^n$ is a constant representing the baseline economic mass for product n in zone j; and π is a parameter to be calibrated. Following Graham et al (2009) we define $M_j^n = \sum_{w \in n} \sum_i L_i^w (d_{ij})^{-\chi}$, where M_j^n is a measure of the accumulated economic mass for industry n in location j; L_i^w is the total size of employment of type w that is relevant to industry n in zone i; d_{ij} is the economic distance from location i to location j; and $\chi > 0$ is a distance-decay parameter. Finally, E_j^n is a constant scalar representing any additional zonal effects on total factor productivity, which is to be calibrated empirically.

The production function (1) differs from that of Anas and Liu (2007) in two ways. First, an economic mass function A_j^n is introduced to represent increasing external return to scale in production: that is, those urban agglomeration effects that arise from land-use and transport changes.(3) Secondly, labour and intermediate inputs enter the production as quantities by zone rather than by zone pair. This makes it easier to calibrate the models empirically, because zonal observations are much more easily found; also the production function is more readily interfaced with existing social accounting matrices (Echenique et al, 2013) and four-step transport models for commuting and for goods transport (see subsection 2.1.3).

Each type of labour and of intermediate inputs consists of commuters and goods/services, respectively, supplied from all available model zones i (including $i = j$); the sourcing of those inputs among zones is modelled through spatial interaction. Each type of building stock in zone j is fixed for the period and updated in the following period as a result of obsolescence, renovation, new construction, etc represented in a recursive model (see subsection 2.1.4).

We follow standard assumptions that producers are cost minimisers under budget and input supply constraints, and operate with zero economic rent and constant internal returns to scale. The price of goods or service n produced in zone j can then be derived as an average and marginal cost. In turn, given X_j^n, the demands for inputs of capital, labour, buildings, and intermediate inputs can be derived from equation (1).(4) Imports into the city region are included as external production.

(3) When such agglomeration effects are strong the model could produce multiple equilibria (Anas and Kim, 1996). Here we expect the parameter π for most cities to be generally below 0.1 (Graham and Kim, 2008; Rice et al, 2006; Rosenthal and Strange, 2004). Zhu (2012) has tested parameter π in the range 0.0–0.2 for primary and secondary industries, and 0.0–0.4 for tertiary industries with a model calibrated for southern England and found that a single quilibrium exists from a reasonable range of alternative input values. The higher the π value, the more is required in calibration to check for possible multiple equilibria.

(4) For further equations and discussions, see Jin et al (2013). This split between summary and detail also applies to the rest of this paper.

2.1.2 *Final consumers*

For the final consumers, we model household choices and leave government budgets, other collective spending, investment decisions, and exports as scenario inputs. The reconciliation between production output (subsection 2.1.1), budget, spending, and investment is a policy decision that should be made explicit as model input. On the other hand, inward investment and export levels may be recursively updated to reflect productivity and prices in the city region.

Household choices here refer to how households source goods and services, choose where to live, and, in the case of working households, determine how to divide time between work and leisure on the basis of utility, prices, and externalities. Households are assumed to maximise utility under constraints of income and time. We follow Anas and Rhee (2006) in including households' consumption of leisure time as well as goods, services, and housing:

$$V_i^H = \sum_m (\alpha^{mH} \ln z_i^{mH}) + \beta^H \ln \left[\sum_k (b_i^{kH})^{\zeta^H} \right]^{\frac{1}{\zeta^H}} + \gamma^H \ln l_i^H \ , \tag{2}$$

where V_i^H defines the economic well-being for household type H which is derived from consumption and leisure in residential zone i; z_i^{mH} is the demand per household H for goods/services of type m in zone i; similarly b_i^{kH} is the housing demand; l_i^H is the leisure time in hours for household of type H during working days of the year; α^{mH}, β^H, and γ^H, where $\left(\sum_m \alpha^{mH}\right) + \beta^H + \gamma^H = 1$, are parameters for consumption in goods/services, housing, and leisure time, respectively; and ζ^H is a parameter for the nested CES function for choosing among housing varieties. Household consumption utility increases not only through consumption, but also through a rise in the number of varieties of housing available for better matching with needs. Households may also trade off consumption against leisure time.

Households' demands for consumption and leisure time are derived through the household budget and the level of incomes, prices, and rents. The households may be segmented by socioeconomic profile, life-cycle, size, etc,

2.1.3 *Location choices and trade patterns*

In many cities, commuting, shopping, and goods delivery patterns and residential location choice have already been modelled by spatial interaction models that are embedded in transport models, often with a richness in market segmentation and behavioural calibration that is worth building upon. The zonal production and consumption functions defined above facilitate a relatively easy interface with spatial interaction models. Following the random utility interpretation of such models (McFadden, 1974), if the location utility for obtaining input m in zone i for user n in zone j is $\nu_{ij}^m = \nu_i^m - d_{ij}^m - \psi_{ij}^m - \Psi_i^m + \varepsilon^m$ (where ν_i^m is the utility of input m from zone i; d_{ij}^m is a generalised transport cost function including travel and logistical costs to transport a unit of input m from zone i to zone j; ψ_{ij}^m and Ψ_i^m are observable nonmonetary barriers for trading from zone i to zone j and for production in zone i, respectively; and ε_{ij}^m is a constant representing unobservable idiosyncratic variations in utility that follow an independent and identical distribution of the Gumbel type) then the trade volume Y_{ij}^m can be expressed generally as:

$$Y_{ij}^m = Y_j^m \left\{ \frac{S_i^m \exp[\lambda^m(\nu_i^m - d_{ij}^m - \psi_{ij}^m - \Psi_i^m)]}{\sum_i S_i^m \exp[\lambda^m(\nu_i^m - d_{ij}^m - \psi_{ij}^m - \Psi_i^m)]} \right\} , \tag{3}$$

where Y_j^{mn} is the total demand for input m by user n in zone j; S_i^m is a size term that corrects for the bias introduced by the uneven sizes of zones in the model (see Ben-Akiva and Lerman, 1985); and λ^m is a scale parameter that measures the concentration of trade among alternative sources which is empirically calibrated along with parameters ψ_{ij}^m and Ψ_i^m.

More specifically, for sourcing of goods and services, $\nu_i^m = -p_i^m$, where p_i^m is the factory-gate price for goods m[5]. This applies to both intermediate and consumer goods/services, including the special cases where the services are travel for leisure and personal business. Commuter households choose where to live based on $\nu_i^m = V_i^H$ and on d_{ij}^m, a generalised cost function for commuting. For noncommuter households, equation (3) is relevant only in cases where their residential locations are determined by previous commuting choices.

An important aspect of spatial choice that has been overlooked in both urban general equilibrium models and land-use and transport interaction models at the city-region scale is the formulation of the d_{ij}^m function. City regions with a reasonably self-contained commuting catchment today tend to have a radius of 50 km or more. At this metropolitan scale, extensive analyses of travel choices data show that a d_{ij}^m function that is linear to travel costs and times will have great difficulties in representing realistic demand elasticities throughout; a nonlinear, Box–Cox transformation of utilities is required (Gaudry and Laferrière, 1989). Fox et al (2009) put forward a log-linear transformation that is a close equivalent to the Box–Cox function whilst being easier to calibrate. This function should fit, in the form:

$$d_{ij}^{mn} = a^m \Big(\sum_k \eta_k^{mn} \chi_{ijk}^m\Big) + (1 - a^m)\ln\Big(\sum_k \eta_k^{mn} \chi_{ijk}^m\Big) - a^m , \tag{4}$$

where the χ_{ijk}^{mn} are the attributes of travel, such as cost or time, and the η_k^{mn} and a^m are parameters.

2.1.4 *Stock constraints*

We define stock constraints in line with Wegener (2001) to cover not only land, buildings, and transport infrastructure but also existing urban activities such as job and home locations which may evolve or 'churn' slowly. For instance, there may be a lag of many years between a utility change and household relocation. For each period, only a proportion of the existing households will be ready to move. Whilst the commuter households make their choices according to equation (3), the moving noncommuter households face the utility level $U_j^H = V_j^H - \Psi_j^H - \eta^H d_{ij}^H + \varepsilon_j^H$, where Ψ_j^H is a nonmonetary barrier for locating in zone j, and d_{ij}^H is a measure of perceived distance from zone i to zone j. We thus obtain a discrete choice model:

$$X'^H_{ij} = X'^H_i \left\{ \frac{S_j^H \exp[\lambda^H (V_j^H - \Psi_j^H - d_{ij}^H)]}{\sum_j S_j^H \exp[\lambda^H (V_j^H - \Psi_j^H - d_{ij}^H)]} \right\} . \tag{5}$$

At spatial equilibrium, the demand for all types of buildings stock, B_i^k and b_i^k, must be equal to available supply, $\hat{B}_j^k$ and $\hat{b}_j^k$. $\hat{B}_j^k$ and $\hat{b}_j^k$, as well as the associated land supply, respond to demand through development/restructuring but subject to regulation, planning, speculation, procurement, construction/renovation, commission and decommission, and inertia. It is thus more appropriate for a model user to specify detailed estate development plans, subject to expected rental revenue and costs. The model can then account for the asymmetry between growth and decline—for example, in the case of business buildings:

$$\hat{B}_j^{k(t+1)} = (1 - \overline{\Upsilon}_j^k)\hat{B}_j^{k(t)} + \vec{B}_j^{k(t+1)}, \qquad \text{if } B_j^k(t+1) \geqslant B_j^k(t) , \tag{6}$$

$$\hat{B}_j^{k(+1)} = (1 - \underline{\Upsilon}_j^k)\hat{B}_j^{k(t)}, \qquad \text{otherwise.} \tag{7}$$

[5] Here we present a simplified model by assuming that input m is shipped straight from zone i to zone j. The logistical channels may be added through a supply-chains model consisting of a series of random utility models for intermediate logistical stages; for an application to the UK, see WSP UK Ltd (2005).

In equation (6), as the total building demand for type k in zone j increases, the existing building stock is depleted by $\overline{\Upsilon}_j^k (0 \leqslant \overline{\Upsilon}_j^k \leqslant 1)$ through demolition and conversion, and the user-specified building stock increment at period $t+1, \vec{B}_j^{k(t+1)}$, is added for period $t+1$. In equation (7), as the total demand falls, the user-specified building increment does not materialise, and the existing building stock is depleted by $\underline{\Upsilon}_j^k (0 \leqslant \underline{\Upsilon}_j^k \leqslant 1)$. In other words, when building demand increases, the user-specified plan is adopted; if demand falls, the existing stock will reduce through depletion, and the user-specified plan is left unimplemented. Similar equations may apply to housing or urban land. The equations reflect the indivisibility of user's development plans (ie, all or nothing for the new stock increment) and can be further refined as proposed by Glaeser and Gyourko (2005).

Similarly, transport infrastructure and services respond to demand subject to regulation, planning, procurement, construction/renovation, commission and decommission, and thus respond to demand slowly and indivisibly. Like land and buildings, user-defined transport supply scenarios are likely to be the most appropriate subject to transport revenues and costs; the growth/decline asymmetry can be applied: that is, new projects are implemented only in the test if the related demand grows.

2.1.5 *Boundary conditions*

External shocks cover decisions that at least partly depend upon factors outside the city region. Business investment and household migration across the city region boundary are such examples. Naturally, external shocks are case dependent. Traditionally external shocks are exogenous, scenario inputs. Nevertheless, policy makers are interested in how changes within a city region may trigger certain shocks under prevailing external conditions.

For such decisions we continue to follow the notion of utility: $U_\mathrm{I} = V_\mathrm{I} + \varepsilon_\mathrm{I}$, where V_I is the measurable average utility for the city region as a whole, and ε_I is a Gumbel-distributed error term. This leads to a discrete choice model

$$X_\mathrm{I}^{(t+1)} = X_\mathrm{All}^{(t+1)} \left\{ \frac{S_\mathrm{I}^{(t)} [\exp(\lambda_\text{I-E} V_\mathrm{I}^{(t)})]}{S_\mathrm{I}^{(t)} [\exp(\lambda_\text{I-E} V_\mathrm{I}^{(t)})] + S_\mathrm{E}^{(t)} [\exp(\lambda_\text{I-E} V_\mathrm{E}^{(t)})]} \right\}, \tag{8}$$

where X_I predicts the activity (eg, migrants or investment) that chooses the city region, and S_I is a size term. All terms with an E subscript denote corresponding values assumed for the external area. Under this random utility framework which accounts for idiosyncratic circumstances through parameter $\lambda_{\mathrm{I-E}}$, we may define a migration function $V_\mathrm{I}^{(t)} = V_\mathrm{I}^{M(t)} = \bar{V}_\mathrm{I}^{\mathrm{H}(t)} - d_{\mathrm{I-E}}^{\mathrm{M}(t)}$ for migration choices subject to average household utility $\bar{V}_\mathrm{I}^\mathrm{H}$ and migratory distance $d_{\mathrm{I-E}}^\mathrm{M}$ at period t, and a business floorspace investment function $V_t^{(t)} = V_\mathrm{I}^{\mathrm{B}(t)} = \ln(\bar{R}_\mathrm{I}^{\mathrm{B}(t)}) - \ln(\bar{p}_\mathrm{I}^{(t)}) + \pi \ln(\bar{M}_\mathrm{I}^{(t)})$ that consists of expected rentals $\bar{R}_\mathrm{I}^\mathrm{B}$, average production cost $\bar{p}_\mathrm{I}$, and productivity effects from the economic mass $\pi \ln(\bar{M}_\mathrm{I})$ at period t.

2.2 Model assembly

Central to model assembly is the fact that urban change processes vary over time scales (Simmonds et al, 2013; Wegener et al, 1986). Some processes adapt quickly to constraints and are thus amenable to equilibrium modelling, such as producer and household relocation and transport choices; others are more inertia prone, lumpy, and indivisible, such as estate development, transport supply, and life-cycle churns of producers and households.

Existing spatial interaction and general equilibrium models, to a varied extent, all adopt a strategy to solve for equilibrium quantities and prices subject to exogenous constraints; the equilibrium condition provides a consistent platform for comparing alternative policy interventions at each time horizon, but such models rely on exogenous scenarios to articulate trajectories between time horizons. Nonequilibrium dynamic models offer insights into the effects of life-cycles, churns, and inertia on temporal trajectories, but have to rely on an interface with other models with an equilibrium mechanism (most often a transport model) to assess any costs and benefits. Cities facing large urban change require both cross-sectional assessment of prices, rents, and wages and the cumulative effects of urban evolution. This calls for a more radical interface between equilibrium and nonequilibrium models. An appropriate articulation of the model components has to be considered for model calibration, validation, and forecasting.

Calibration of a recursive model requires not only a representation of the city region at a base year t, but also at least one transitional period to the next horizon $t + 1$, preferably more. For calibration at base year t, all boundary conditions and constraints including the activity stocks are needed as inputs, as are the quantities and prices of goods and services, labour, buildings, land, and trade patterns. The model estimates the demand for goods and services, labour, buildings, land, travel, traffic flows, and all associated prices based on input boundary conditions, stock constraints, and an initial set of model parameters that are derived through successive partial equilibrium model estimations (see the left-half of figure 1). The solution algorithm proceeds iteratively through each of the markets until all demand and prices reach equilibrium. The model predictions are compared with known zonal quantities and prices to refine the parameters. The model parameters are then retained for use for period $t + 1$.

For transition to period $t + 1$, the known changes in boundary conditions, stock constraints, and associated knowledge on policy interventions are used to establish recursive models that

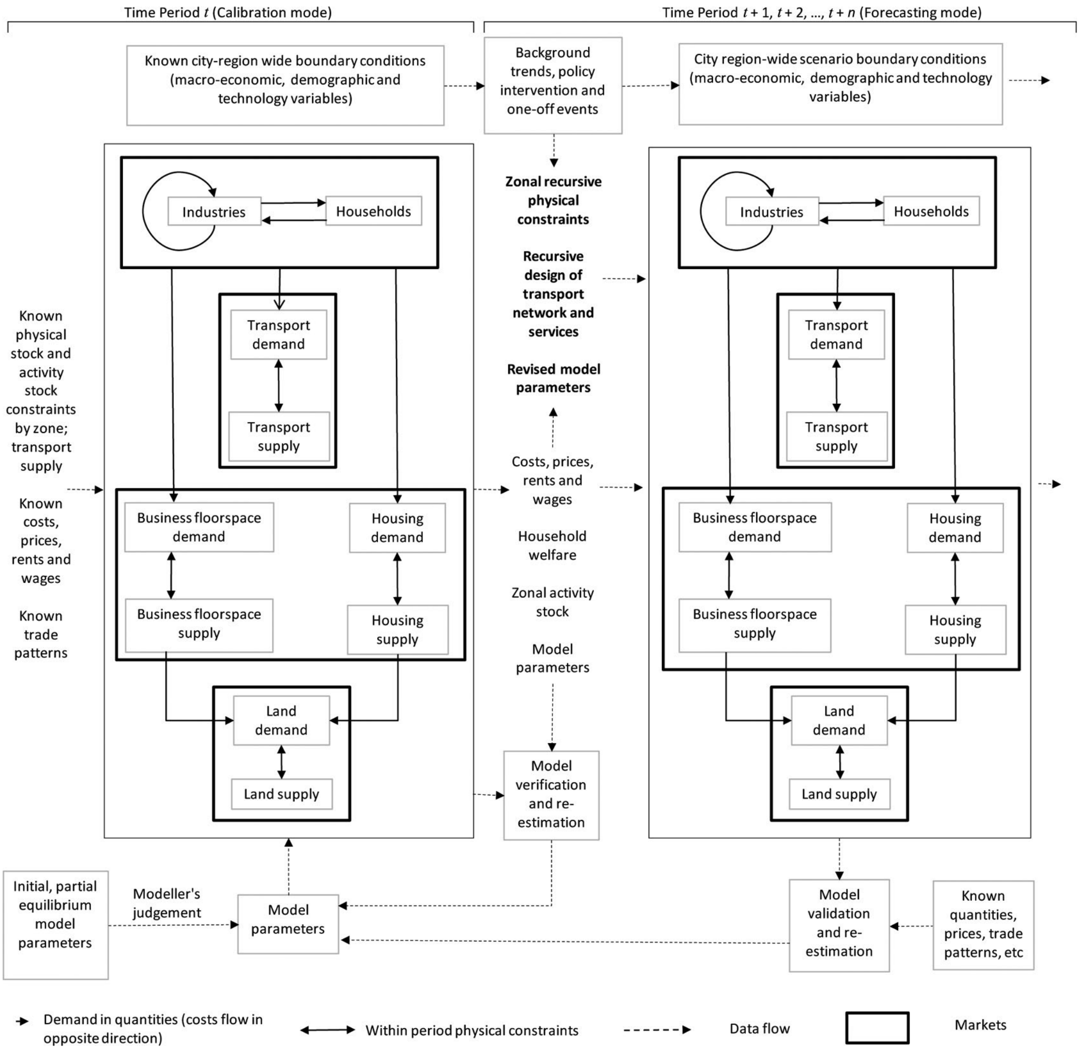

Figure 1. Main information flows within and between recursive spatial equilibria.

predict boundary conditions and stock constraints for period $t + 1$. The recursive models may include one-off events which enter as period-specific constants. The spatial equilibrium model is run in forecasting mode for period $t + 1$. The model predictions are validated through comparing with known zonal quantities and prices for period $t + 1$ which have not been used in model calibration. The recursive and spatial equilibrium model calibration may have to be repeated many times in a calibration–validation loop until a satisfactory goodness of fit has been achieved (see the right-hand half of figure 1).

Ideally, more than one known transition period exists so that the recursive models for boundary conditions and stock constraints can be tested repeatedly, and the model builds up a validated track record. In practice, it is rarely feasible to trace back more than one historic period for data problems and modeller resources. An effective way to achieve multiple-period validation may be to retain existing models and extend them through time, and use the successive model development exercises to extend the series of recursive models. From period $t + n$ ($n \geqslant 2$) the model will be used in forecasting mode. The recursive and spatial equilibrium models share the same running procedure as for model validation at period $t + 1$.

Whilst the spatial equilibrium model for each horizon $t + n$ ($n \geqslant 0$) is a static equilibrium model, the recursive model representing the transition of boundary conditions and stock constraints are nonequilibrium in nature. Although the recursive models are perhaps the most uncertain to begin with, their outputs for transition between time horizons are nevertheless made plain to see by all model users. In fact, in the case of forecasting, the model users may wish to intervene and revise the projections, either at the city-region level or for specific zones, as a form of scenario design. Nevertheless, a gradual establishment of evidence-based recursive models is particularly useful for radical development and restructuring scenarios—however much they are interested in such scenarios, the far-sighted decision makers might not want to be seen specifying them for political reasons.

The number of years elapsed between two modelled time horizons is a local matter. The standard assumption of the recursive model is that an urban administration goes through a stereotypical cycle from new initiatives to policy implementation and ultimately to the lame-duck phase: in such cases, the majority of the stimuli to boundary condition and stock constraint changes would occur early in the period; producers and consumers then adapt before the next round of radical changes. However, development cycles are hardly universal, and the time horizons are heavily constrained by data availability (eg, the census years) and masterplan horizons. Locally specific considerations are thus crucial in determining the period length. In our experience, ten or more years may be required for development and restructuring effects to work through producer and consumer choices. This is true even during the recent fast growth in China since the late 1970s, where distinct policy cycles are generally around ten years (Zhang, 2010).

2.3 Model outputs for policy assessment

The model outputs are quantities (production, factor inputs, and consumer demand) and prices (of goods/services, wages, and rents) in each zone, and movements of people and goods/services between zones. A multimodal transport model or a collection of unimodal traffic models need to be incorporated to estimate travel demand, costs, operation characteristics, and congestion/overcrowding levels. The outputs provide the basis for assessing economic, social, and environmental benefits (Echenique et al, 2012).

In the model, two types of prices are accounted for in parallel under spatial interaction: the consumption price of inputs that come from different zones are calculated as an average of the delivered prices weighted by respective trade volumes; the average utilities of the inputs are calculated as a log-sum (Ben-Akiva and Lerman, 1985; Williams, 1977) of the delivered prices.

Household utility is not linear in income and the marginal utility of income varies between policies and among zone pairs of spatial interaction (Anas and Rhee, 2006). The overall consumer surplus, ΔC, in the city region as a household well-being measure may be defined as the change in average household utility divided by the average marginal utility of money:

$$\Delta C = \left(\overline{V}^{\text{Alternative}} - \overline{V}^{\text{Base}}\right) \Big/ \left[\frac{1}{2}\left(\frac{1}{\overline{\Omega}^{\text{Alternative}}} + \frac{1}{\overline{\Omega}^{\text{Base}}}\right)\right], \tag{9}$$

where $\overline{V}^{\text{Base}}$ and $\overline{V}^{\text{Alternative}}$ are the average household utilities, and $\overline{\Omega}^{\text{Base}}$ and $\overline{\Omega}^{\text{Alternative}}$ are the average household incomes for the Base and Alternative scenarios, respectively.

3 Model tests

Although the model components follow three well-established model traditions, the new model design still needs thorough in-lab testing. This is because, first, the interactions between the recursive and equilibrium components create many new mechanisms that do not exist in current models. Secondly, an understanding of the range and uncertainty of parameter values helps to develop a prioritised agenda for empirical model estimation. Thirdly, large-scale urban change may be a challenge for the spatial equilibrium model to converge.

We set up test model code in MatLab (The MathWorks Inc.) with a flexible zone dimension. When it is used for a one-zone model, all model results may be traced easily by hand. Here we use a model with twelve zones which retains the fundamental features of a city region and can represent archetypal urban development strategies, in order to pressure-test the model with easily manageable data tables. We present the key results here and further details are to be found in Jin et al (2013).

We specify a narrow peninsular city region with the following zones: (1) an older, denser city centre at the cape where businesses concentrates with limited housing; (2) a built-up inner city with both homes and jobs; (3) a contiguous outer urban area where housing dominates; (4) a greenbelt where development has been restricted; (5) a far suburb beyond the greenbelt with multiple commercial centres scattered among towns and villages; (6) a wider rural hinterland which is sparsely populated (figure 2). We further distinguish a free-standing city in the far suburb, and five small areas which are the main catchment of large rail stations—we code them as zones 10 and R1–R5 respectively. The spatial configuration of this model has made land-use patterns more explicit, but otherwise it follows the tradition of the 'long narrow city' of Solow and Vickrey (1971), applied, for example, by Anas and Kim (1996) and Eliasson and Mattsson (2001).

To make the data flows easy to trace in a complex model, we make a number of simplifications. We assume that the total population in our city region is 1 million at time t (say 2010). There are nine other city regions of the same size in the country (thus the total number of households in the country is 10 million), although there are none nearby.[6] Periodically, households in other city regions as well as this one compare their well-being and make decisions to migrate between them. The city regions altogether face a population growth of 2.5 million per decade, thus doubling at period $t+4$ (2050) to 20 million. The boundary conditions are migration subject to average household migration utilities $V_{\mathrm{I}}^{\mathrm{M}}$ and business floorspace investment subject to attractiveness function $V_{\mathrm{I}}^{\mathrm{B}}$ (see subsection 2.1.5). It is also subject to the reservation utility for the rest of the country, V_{E}. The solving algorithm of the spatial equilibrium model is shown in figure 3.

We define one type of household. Each household supplies one worker who fills one job. Trade across the city region boundary is zero; the workers produce a product that is entirely consumed by the households in the city region. The households also own the estate properties

[6] If there are, the internal modelled area shown in figure 2 may be expanded to include them.

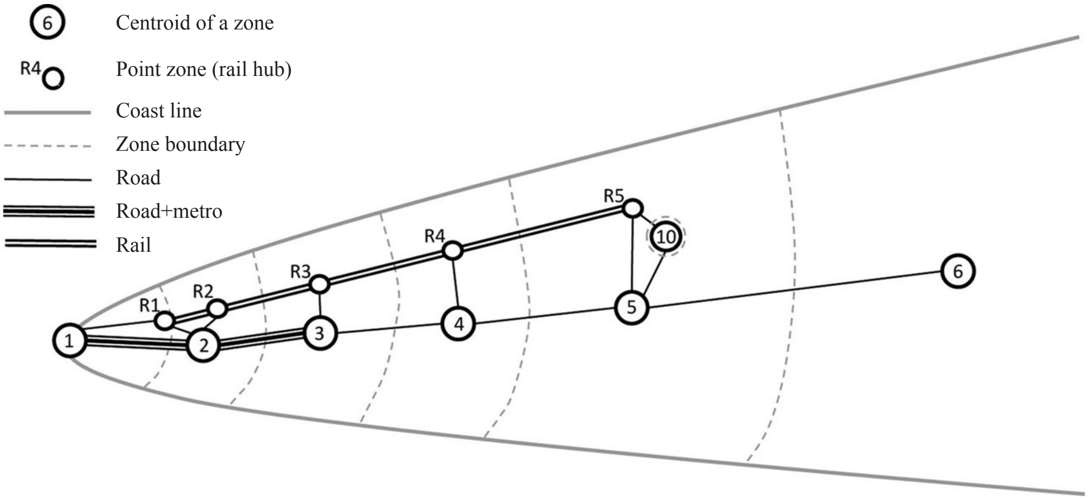

Figure 2. The model area, transport links, and zone numbers. Diagram not to scale; physical dimensions are specified by land-use and transport data.

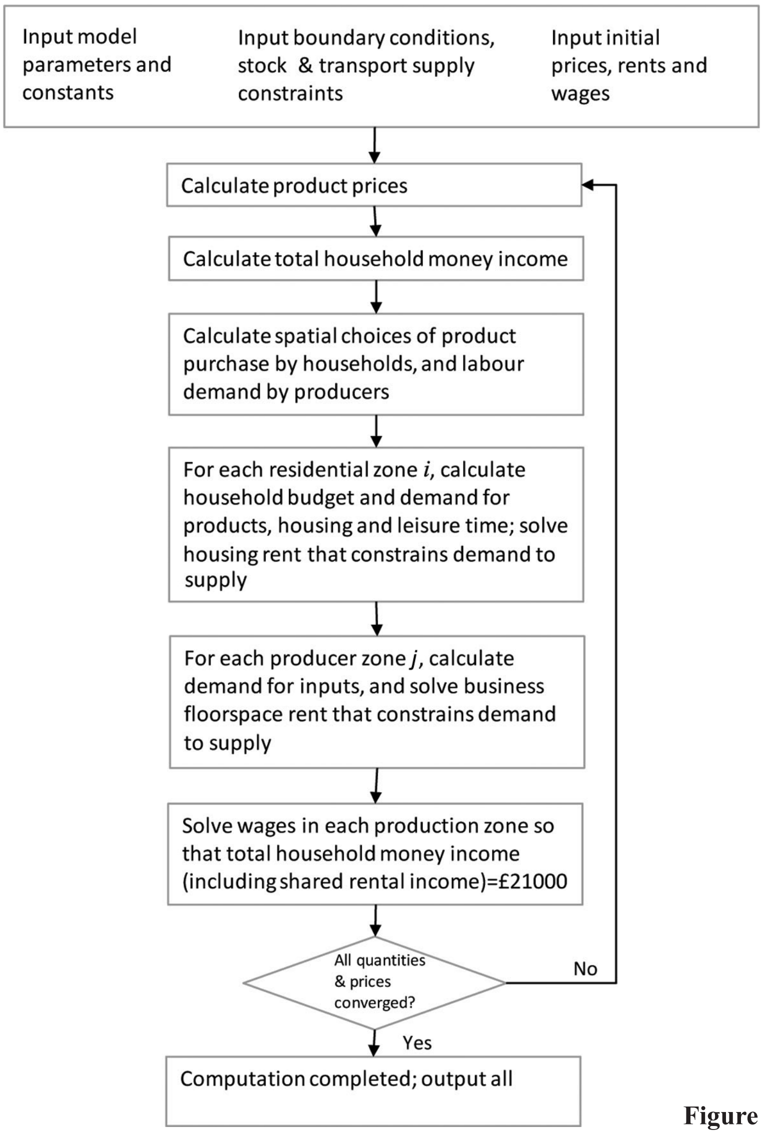

Figure 3. Model-solving algorithm.

collectively and share out the rental income equally. We assume an average household wage income of around £12 000 (or US $18 000) per year.[7] There are two types of housing (houses and apartments) and two types of business floorspace (bespoke and generic).

3.1 Model parameterisation

We take parameter values from established models. Where there are no commonly accepted parameters we carry out sensitivity tests in the model and adopt value ranges by judgment. Table 1 lists the model parameters that have been specified in the equations.

Table 1. Model parameters and their sources.

Model parameter	Value(s)	Sources
δ^n (labour cost share)	0.86	Anas and Rhee (2006)
μ^n (business floorspace cost share)	0.14	Anas and Rhee (2006)
ν^n (capital cost share)	0.00	Anas and Rhee (2006)
γ^{mn} (intermediate inputs cost share)	0.00	Anas and Rhee (2006)
ζ^n (business floorspace variety effects)	0.90	Own sensitivity tests
E_j^n (residual total factor productivity multiplier)	1	Anas and Rhee (2006)
π (economic mass effects on productivity)	0.05-0.10	DfT (2006); Graham and Kim (2008)
α (household utility parameter for goods/service)	0.36	Anas and Rhee (2006)
β (household utility parameter for housing)	0.15	Anas and Rhee (2006)
γ^H (household utility parameter for leisure time)	0.49	Anas and Rhee (2006)
ζ^H (housing variety effects)	0.90	Own sensitivity tests
λ^m (scale parameter for spatial interaction model)	1	Calibrated to reproduce an average commuting distance that is compatible with mid-income commuters in the London region in 1991 (Jin et al, 2002), in conjunction with a^m below
$\psi_{ij}^m, \Psi_i^m, \psi_{ij}^H, \Psi_j^H$ (zone-specific attractiveness)	0 for all i, j	The zones are featureless other than represented by land-use and transport data
a^m (log–linear travel cost function parameter)	0.0005	See above
η_k^m (log–linear travel cost function parameter)	500	A multiplier to converts travel costs and times of one trip to annual (2 trips a day, 250 days a year)
$\underline{\Upsilon}_j^k$ (building stock depletion)	0	Building stock depletion is not included here for simplicity
$\lambda_{\text{I–E}}^{\text{H}}$(scale parameter for household migration model)	1.0-4.0	Own sensitivity tests
$\lambda_{\text{I–E}}^{\text{B}}$ (scale parameter for business floorspace investment model)	1.0	Own sensitivity tests
Total number of working days a year	250	Anas and Rhee (2006)
Hours per day	24	Anas and Rhee (2006)
Cost for delivering a unit of local service as percentage of commuting trip cost	10%	Anas and Rhee (2006)

[7] This income is supplemented by shared rental income, implying an average household income of £21 000; this represents a reasonably affluent profile that the leading emerging economies are currently aiming towards.

3.2 Model runs

We present three types of run to highlight the key features of the model: (1) the base year t which represents 2010; (2) a set of static equilibrium runs for period $t + 4$ (2050) with given boundary conditions; (3) a set of recursive equilibrium runs from 2010 to 2050.

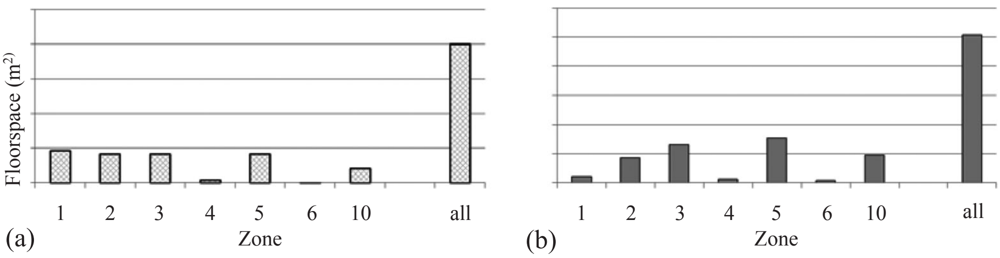

Figure 4. Floorspace constraints by zone in year t (2010): (a) business, (b) housing.

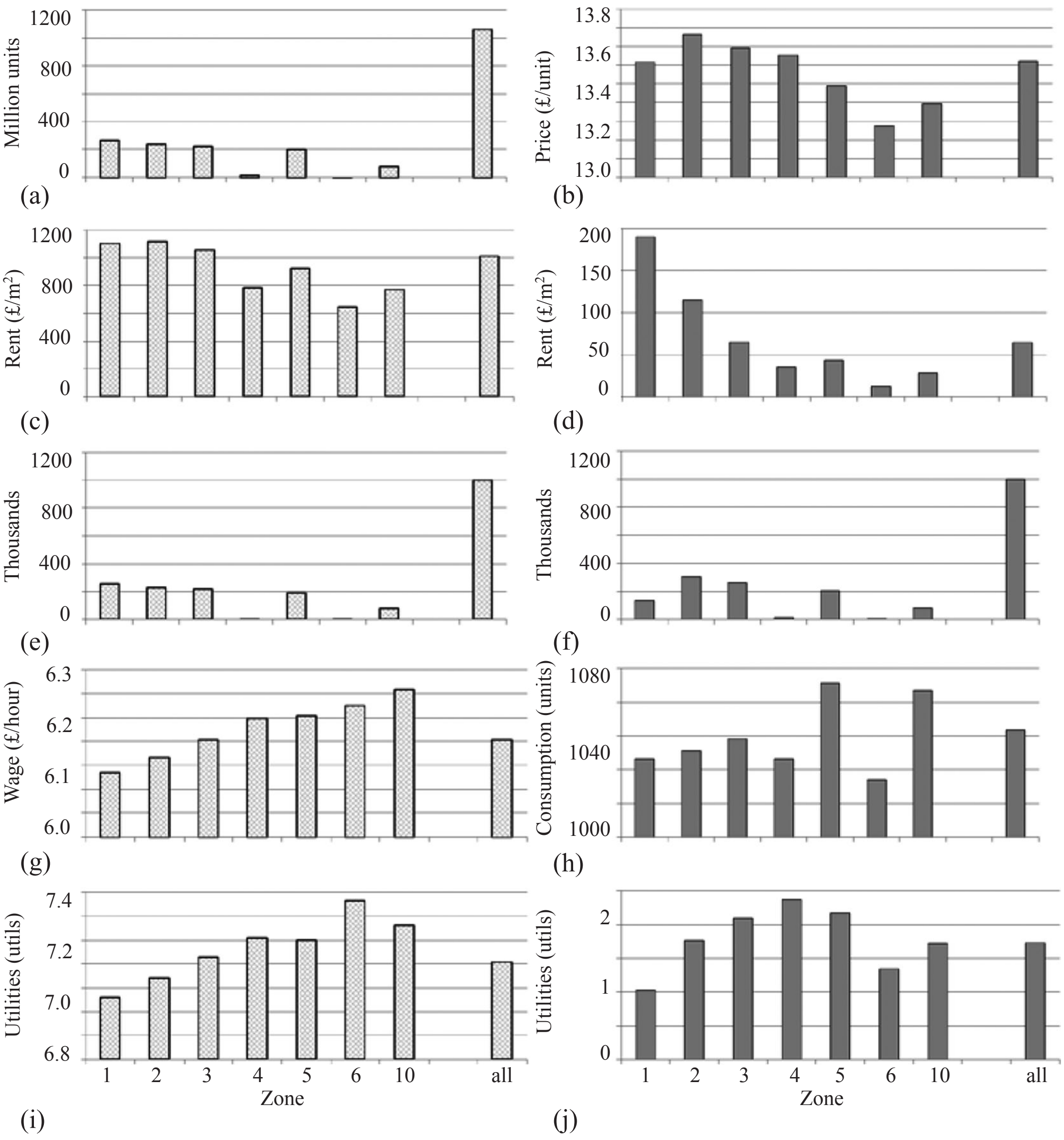

Figure 5. Model output quantities and prices by zone, in year t (2010): (a) production output; (b) product prices; (c) business rents; (d) housing rents; (e) number of jobs; (f) number of households; (g) wages (home location); (h) consumption per household; (i) consumption utilities; (j) commuter location utilities.

3.2.1 *Model run for* t *(2010)*

The model starts with inputs of transport supply, stock of housing and business buildings (figure 4), stock of households and jobs, and boundary condition (total households = 1 million) at time *t* (2010) for a static spatial equilibrium run. The output activity stock in this case equals the input; the model also outputs prices, rents, wages, and household utilities by zone (figure 5). Through the interface with the transport model, the travel distances, costs, and times incurred by labour and product flows are computed (which are summarised in figure 6). The model outputs depict a polycentric city region where the densely built-up areas have short average travel distances, long travel times, and high rents; the reverse is the case in the suburbs.

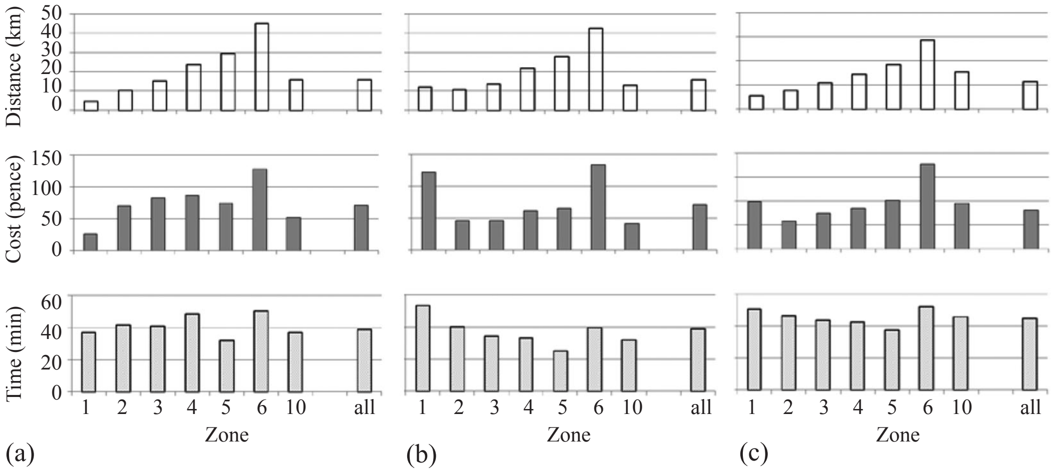

Figure 6. Model output average trip distances, travel costs, and travel times by purpose by zone, in year *t* (2010) for: (a) commuting—by home origin; (b) commuting—by workplace; (c) goods and services—by household location.

3.2.2 *Static spatial equilibria for 2050*

Before running the model recursively, we tested the spatial equilibrium component by static runs for four archetypal scenarios: (1) *trend growth* which targets development opportunities through inner-city regeneration and greenfield development beyond the greenbelt, (2) *compact* development of existing built-up areas without new greenfield land supply, (3) expansion of *garden suburbs* outside built-up areas at prevailing suburban densities, and (4) densification around urban *rail hub* locations which is an upscaled version of transit-oriented development. For these static runs we assume that the city region will grow at the country-average rate: that is, doubling the number of households to 2 million in 2050. Half of the expected floorspace construction will be natural growth which occurs pro rata to existing zonal stock, and the remainder is specified by the respective planning scenarios.

For each scenario (2)–(4), three variants are tested: (a) *maintaining the status quo:* average floorspace per household and per job, and average travel costs and times remain unchanged from 2010;[8] (b) *scale of floorspace construction following zonal profiles* per household and per job under each planning scenario: 30% less per household and per job in dense built-up zones and 30% more in suburban and rural zones; (c) *accompanying traffic speeds following zonal profiles* in addition to zonal floorspace profiles: in the case of compact development and garden suburbs, traffic congestion worsens—average intrazonal travel times increase by 5 minutes, and the access times to and from those zones increase by 10 minutes per trip; in

[8] This follows pragmatic policy targets used in many cities where infrastructure investment aims to keep network speeds on main transport corridors constant, through expanding network capacity and services, and peak time traffic management.

Table 2. A comparison of 2010 and 2050 static equilibrium tests.

	Base Year 2010	Trend growth 2050	Compact, 2050			Garden suburbs, 2050			Rail hubs, 2050		
			(a)	(b)	(c)	(a)	(b)	(c)	(a)	(b)	(c)
Inputs											
Total households (million)	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Total jobs (million)	1.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0	2.0
Total housing (million m^2)	101.4	202.8	202.8	187.6	187.6	202.8	228.1	228.1	202.8	187.6	187.6
Total business floorspace (million m^2)	2.0	40.0	40.0	37.0	37.0	40.0	45.0	45.0	40.0	37.0	37.0
Outputs											
Total production (million units)	1063.4	2135.1	2127.1	2105.1	2084.2	2145.7	2188.2	2185.6	2125.6	2102.6	2123.2
Average product prices (£/unit)	13.6	13.6	13.6	13.8	13.9	13.5	13.3	13.3	13.6	13.8	13.7
Average wages (£/hour)	6.2	6.2	6.2	6.2	6.3	6.2	6.1	6.1	6.2	6.2	6.1
Average housing rents (£/m^2)	64.2	64.2	64.2	69.4	69.7	64.1	56.9	57.0	64.4	69.5	69.3
Average business rents (£/m^2)	101.4	101.5	101.4	109.7	109.4	101.5	90.3	90.2	101.3	109.6	109.7
Average commuting time (min/trip)	38.6	37.2	39.1	39.0	42.9	34.9	33.3	32.7	36.2	37.1	32.5
Household consumption utility	7.208	7.214	7.209	7.193	7.193	7.223	7.250	7.237	7.218	7.200	7.192
Consumer surplus as percentage of household money income		0.0	−1.1	−4.4	−4.4	1.7	7.4	4.7	0.7	−2.9	−4.7
Average economic mass index	338.2	691.1	701.0	691.5	608.5	760.2	824.7	743.5	663.5	662.4	745.2
Effect of economic mass on productivity (elasticity = 0.05)			0.1	0.0	−0.6	0.5	0.9	0.4	−0.2	−0.2	0.4
Effect of economic mass on productivity (elasticity = 0.10)			0.1	0.0	−1.3	1.0	1.8	0.7	−0.4	−0.4	0.8

Note: the variants are (a) maintaining the status quo; (b) scale of construction following zonal profiles; (c) accompanying traffic speeds following zonal profiles. See text above for further explanations.

the case of rail hub developments, the intrazonal travel times remain unchanged, whilst the interzonal travel times to and from the rail hubs reduce by an average of 5 minutes thanks to a combination of improved headways of rail services and station access.

Using parameters from established models, the spatial equilibrium tests reveal stark differences among the scenarios and variants by working through the full implications of the supply constraints on prices, wages, rents, household utility, consumer surplus, and economic mass. Table 2 shows that floorspace and traffic congestion could reduce household welfare by an equivalent of 4.4% of average income whilst reducing per employee productivity by 0.6%–1.3% under the compact variant (c); better housing and business floorspace supply without worsening traffic congestion could raise household welfare by 7.4% of income whilst improving per employee productivity by 1.8% under the garden suburbs variant (b) (figures 7 and 8). The results are corroborated in nature by studies of real city regions

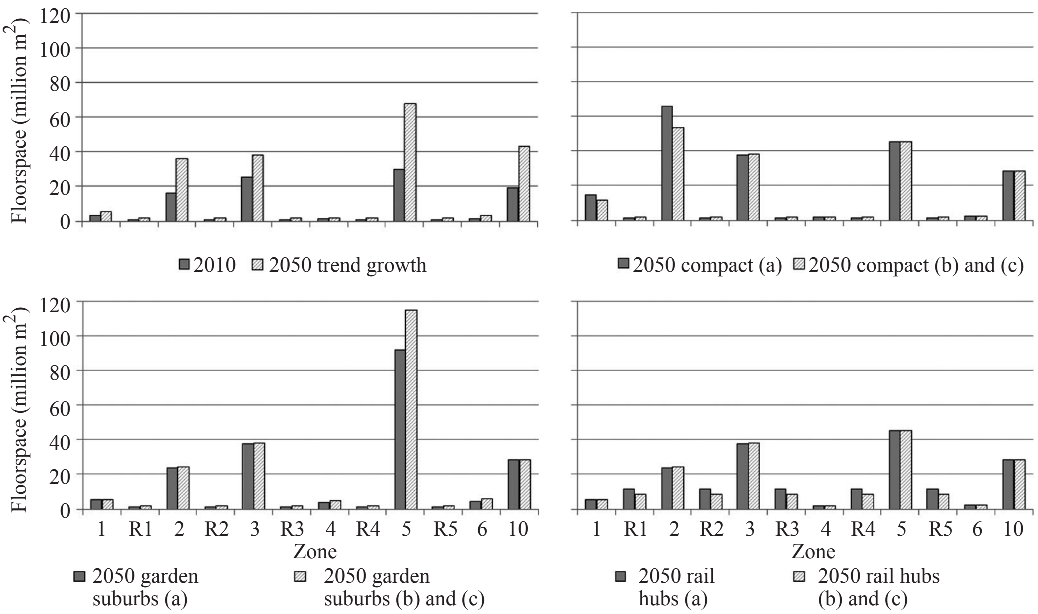

Figure 7. Housing floorspace inputs to 2010 and 2050 static equilibrium tests.

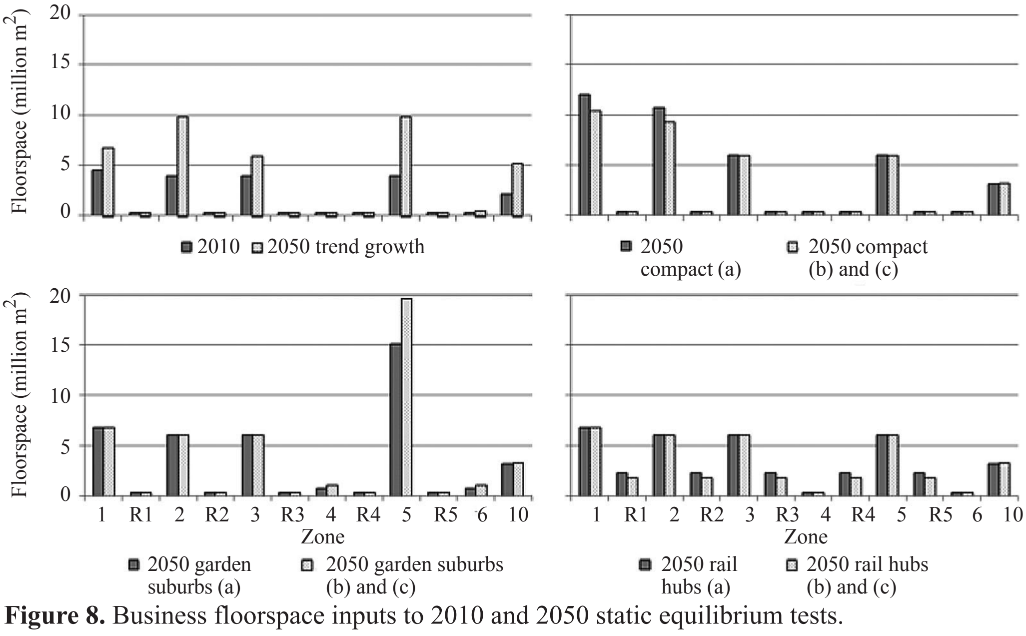

Figure 8. Business floorspace inputs to 2010 and 2050 static equilibrium tests.

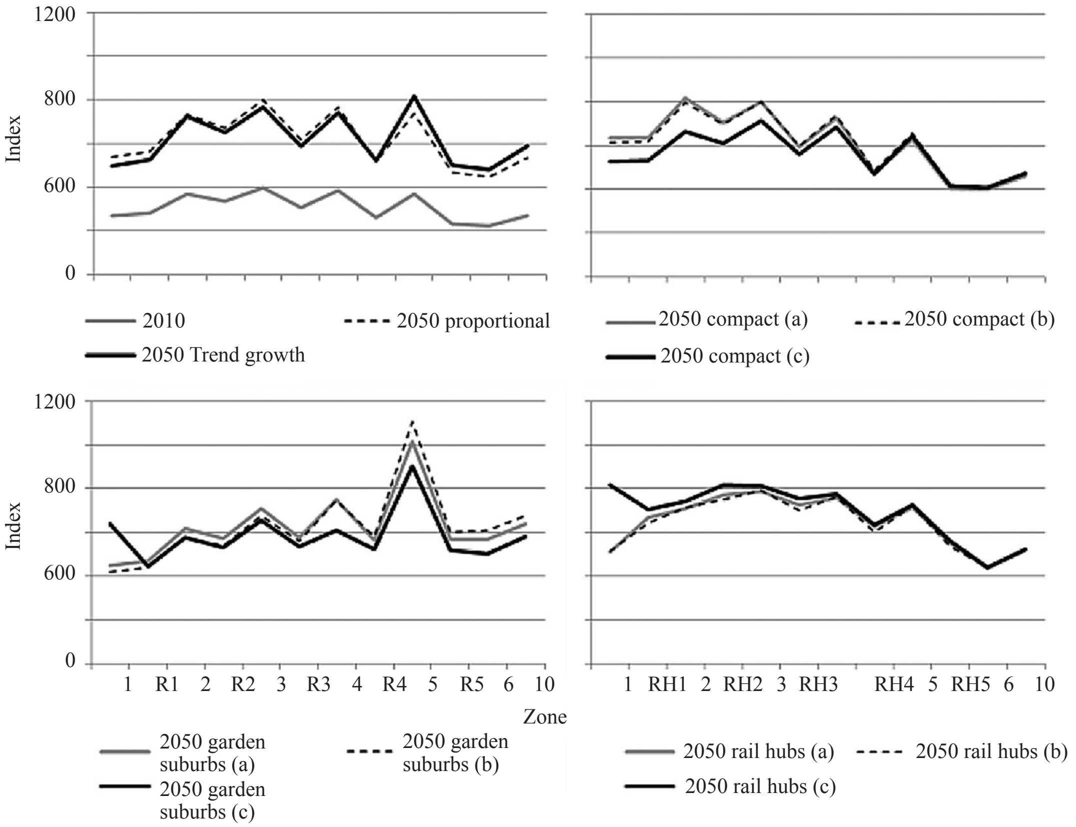

Figure 9. Zonal indices of economic mass: 2010 and 2050 static equilibrium tests.

(see Echenique et al, 2012). The most complex responses appear to be with the rail bub developments of which the overall impacts on welfare and productivity are very sensitive to detailed input specifications, with household welfare changes varying from 0.7% to −4.7% of average income, and −0.2% to 0.8% for productivity effects across variants (a) to (c). Figure 9 presents the implications of economic mass under the scenarios with different land-use and transport configurations.

The significant differences in household utility levels among the scenarios show that the assumption of a constant 2 million household size across scenarios may not be realistic. We now turn to this question by incorporating a recursive model for the boundary conditions.

3.2.3 *Recursive spatial equilibria (RSE): trend growth and rail hub tests: 2010–2050*

The RSE needs first to start with a baseline scenario, which we define as trend growth. The boundary conditions are total households in the city region and total new business floorspace investment. Without affecting generality, we assume that our city region leads the country by a decade: that is, the external reservation household location utility is equal to that for our region a decade earlier. As there are no consensus recursive model parameters, we present tests with household relocation parameter $\lambda^{\mathrm{H}}_{\mathrm{I\text{–}E}} = 1.0$ and 4.0 whilst keeping business floorspace investment parameter $\lambda^{\mathrm{B}}_{\mathrm{I\text{–}E}}$ constant at 1.0. Both boundary conditions are predicted through equation 8. New housing and business floorspace construction plans are then linked to household growth and business floorspace investment, respectively; zonal floorspace supply is subject to the asymmetric build-out (see subsection 2.1.4).

We then set up finer-grained variants for the rail hub scenario: (i) transport access to the five hubs is gradually improved with average access times shortened by 1, 2, 4, and 6 minutes, respectively, for each decade 2010–50; (ii) transport improvements delayed by a decade, so average access times are 1, 2, and 4 minutes shorter for respective decades 2020–50;

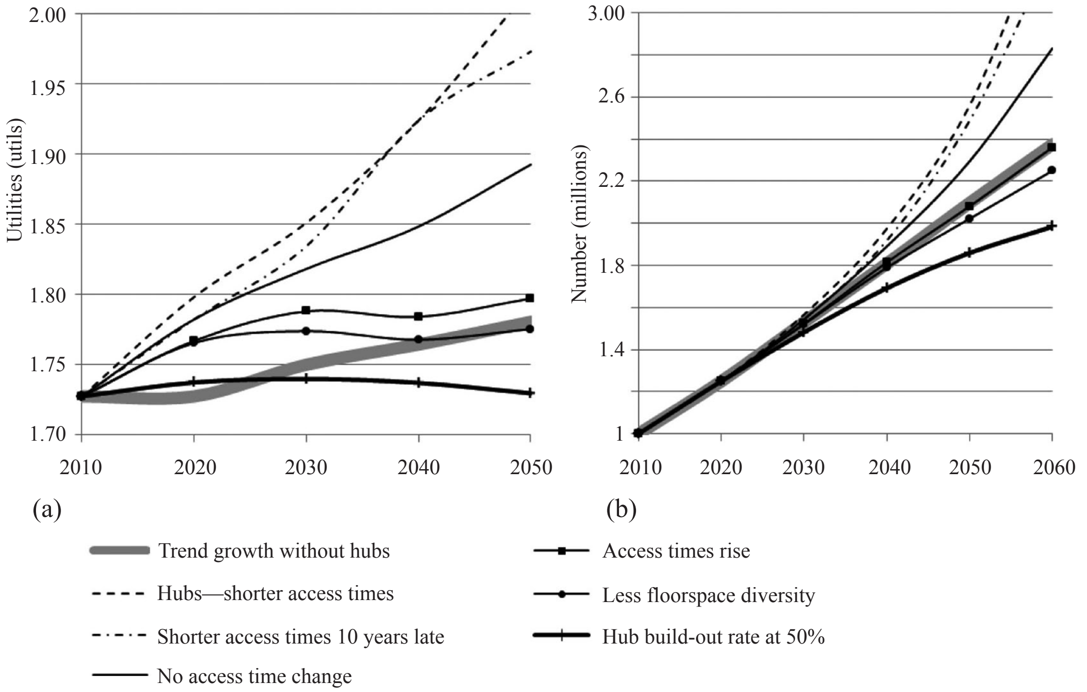

Figure 10. Summary of the household growth trajectories 2010–50 under recursive spatial equilibria (i): trend growth and rail hub scenarios ($\lambda^{\mathrm{H}}_{\mathrm{I-E}} = 1.0$ and $\lambda^{\mathrm{B}}_{\mathrm{I-E}} = 1.0$) for: (a) household location utility; (b) total number of households by year.

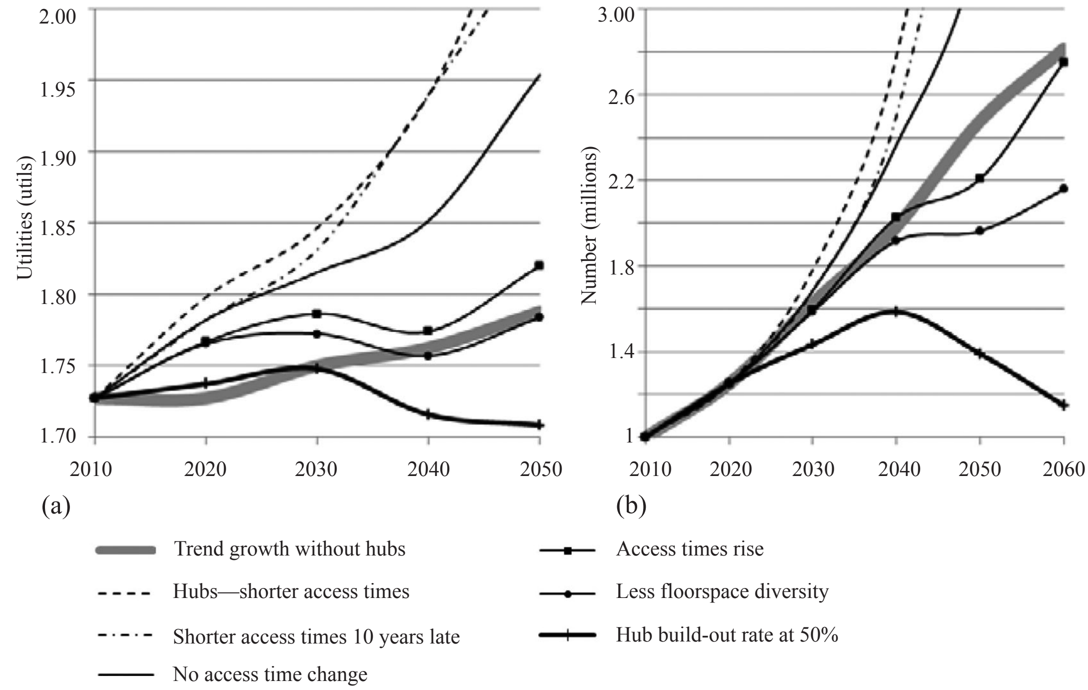

Figure 11. Summary of the household growth trajectories 2010–50 under recursive spatial equilibria (ii): trend growth and rail hub scenarios ($\lambda^{\mathrm{H}}_{\mathrm{I-E}} = 4.0$ and $\lambda^{\mathrm{B}}_{\mathrm{I-E}} = 1.0$) for: (a) household location utility; (b) total number of households by year.

(iii) access conditions remain as in 2010; (iv) gradually worsening access, the reverse of (i); (v) reduced business and housing floorspace diversity—instead of a 50:50 balance between the floorspace stock varieties, the balance is 90:10; this builds on test (iv); (vi) the floorspace completion rate in the hubs reduces by 50%, otherwise the inputs are the same as (v).

As one would expect from equation (8), the results show that when $\lambda^{\mathrm{H}}_{\mathrm{I-E}}$ is small the share of population in our city region follows more closely the historic household share, and the growth trajectories form monotonic trajectories around trend growth, the city region variously reaching 2 million to over 4 million households. Figure 10 shows both the evolution of average household location utility [figure 10(a)] and the resultant total household size change [figure 10(b)]. As in equation 8, the location utility of period t predicts the total household size of period $t+1$. As $\lambda^{\mathrm{H}}_{\mathrm{I-E}}$ increases to 4.0, the relocation decisions become more sensitive to household utility changes and the cumulative effects range from a dramatic growth in excess of 6 million by 2050 to a radical reversal of growth to under 1.2 million (figure 11). This does not only lead to changes in prices, wages, rents, household utility, consumer surplus, and economic mass at the zonal level in a way that cannot be predicted by static spatial equilibria; it also predicts qualitatively different city sizes (2 million to over 4 million) in terms of economic mass and productivity, even with relatively low values of $\lambda^{\mathrm{H}}_{\mathrm{I-E}} = 1.0$.

In the test model, all households can relocate in response to relocation utility levels subject to their idiosyncratic tastes. We have also carried out tests where the majority of households are subject to churns in their life-cycles and are not free to relocate in each time period. However, because our city region is experiencing 100% growth over the whole period, the conclusions reached above still hold if there is a reasonable activity churn rate.

4 Discussions

We return here to the questions posed by Volterra and CBP (2007). The analysts' aspiration to examine "the links between productivity, wages and rents and the full implications of these for output growth" could be met through a general equilibrium model; the difficulty lies with a spatially detailed application to answer their follow-up questions about behavioural responses, trade patterns, etc, Our proposed interface with detailed transport and traffic models brings spatial equilibrium models into play in assessing individual projects and initiatives.

"[Testing] that the models we use reflect the world in which we operate" links to growth trajectories. It is clear that the priority for model estimation has to be empirically robust models for recursive updating of the boundary conditions and stock constraints, which could generate qualitatively distinct urban futures which are of critical importance to major urban infrastructure and land-use decisions.

We acknowledge our enormous intellectual debt to three distinct modelling traditions. The proposed model has a fairly parsimonious structure and a relatively small number of parameters. Nevertheless, whether they are 'deep' parameters(9) will yet depend on model segmentation in empirical applications. There is already a wealth of literature regarding the likely values/ranges of some parameters. However, building a consensus on all the key parameters has far to go, particularly for the recursive models. Extensive 'in-lab' tests of the parameters would seem useful in guiding further work with the empirics.

5 Conclusions

The new RSE model combines two features that are required by policy makers: (1) it enables simulation of urban evolution trajectories that the existing equilibrium or nonequilibrium models cannot produce in isolation, and (2) it quantifies impacts of policy interventions on a consistent basis for a given time horizon. These two features cannot be simultaneously

(9) In the sense that parameters are invariant across the policy scenarios of interest (Lucas, 1976).

achieved by existing models. The proposed model has also incorporated new elements that enable the modelling of productivity effects of land-use and transport interventions, and a more precise handle on city-region-scale travel choice behaviour through a log–linear travel utility transformation. However, a recursive use of static spatial equilibrium models over successive policy horizons is but a very small and experimental step towards dynamic equilibrium modelling and much remains to be done.

Acknowledgements. The authors wish to acknowledge the funding support of the UK Engineering and Physical Sciences Research Council (EPSRC) through the EECi Project (for Marcial Echenique and Ying Jin) and the ReVISIONS Project (for Marcial Echenique and Tony Hargreaves). We wish to thank Michael Wegener, Ian Williams, and two anonymous referees for insightful comments, and Alex Anas for an unpublished model algorithm appendix to Anas and Rhee (2006). We thank UK Department for Transport for use of the LASER3.0 model which informed the choice of some parameters. Dr Jie Zhu assisted with MatLab programming. The authors alone are responsible for the views and any remaining errors.

References

Allen P M, 1997 *Cities and Regions as Self-organizing Systems: Models of Complexity* (Taylor and Francis, London)

Anas A, 2013, "A summary of the applications to date of RELU-TRAN, a microeconomic urban computable general equilibrium model" *Environment and Planning B: Planning and Design* **40** 959–970

Anas A, Kim I, 1996, "General equilibrium models of polycentric urban land use with endogenous congestion and job agglomeration" *Journal of Urban Economics* **40** 232–256

Anas A, Liu Y, 2007, "A regional economy, land use, and transportation model (RELU-TRAN©): formulation, algorithm design, and testing" *Journal of Regional Science* **47** 415–455

Anas A, Rhee H-J, 2006, "Curbing excess sprawl with congestion tolls and urban boudaries" *Regional Science and Urban Economics* **36** 510–541

Batty M, 1976 *Urban Modelling* (Cambridge University Press, Cambridge)

Batty M, 2005 *Cities and Complexity: Understanding Cities with Cellular Automata, Agent-based Models, and Fractals* (MIT Press, Cambridge, MA)

Batty M, 2009 "Urban modeling", in *International Encyclopaedia of Human Geography, Volume 12* Eds R Kitchin, N Thrift (Elsevier, Oxford) pp 51–58

Batty M, 2010, "Integrated models and grand challenges" *ArcNews Online* Winter 2010/11, http://www.esri.com/news/arcnews/arcnews.html

Batty M, Vargas C, Smith D, Serras J, Reades J, Johansson A, 2013, "SIMULACRA: fast land-use–transportation models for the rapid assessment of urban futures" *Environment and Planning B: Planning and Design* **40** 987–1002

Ben-Akiva M, Lerman S, 1985 *Discrete Choice Analysis* (MIT Press, Cambridge, MA)

Bröcker J, 1998, "Operational spatial computable general equilibrium modeling" *The Annals of Regional Science* **32** 367–387

Bröcker J, Korzhenevych A, 2011, "Forward looking dynamics in spatial CGE modelling", Kiel Working Papers 1731, http://www.ifw-members.ifw-kiel.de/publications/ forward-looking-dynamics-in-spatial-cge-modelling/forward-looking-dynamics-in-spatial-cge-modelling.pdf

Chapin F S, Weiss S F, 1968, "A probabilistic model for residential growth" *Transportation Research* **2** 375–390

Clarke G (Ed.), 1996 *Microsimulation for Urban and Regional Policy Analysis* (Pion, London)

Daly A, Zachary S, 1978, "Improved multiple choice models", in *Determinants of Travel Choice* Eds D Hensher, M Dalvi (Saxon House, Farnborough, Hants) pp 325–357

De la Barra T, 1989 *Integrated Land Use and Transport Modelling* (Cambridge University Press, Cambridge)

DfT, 2006 *Transport, Wider Economic Benefits, and Impacts on GDP* Department for Transport, London, http://www.dft.gov.uk/pgr/economics/rdg/webia/transportwidereconomicbenefi3137

Domencich T, McFadden D, 1975 *Urban Travel Demand: A Behavioural Analysis* (North-Holland, Amsterdam)
Echenique M, 2004, "Econometric models of land use and transportation", in *Transport Geography and Spatial Systems* Eds D A Hensher, K J Button (Pergamon, Kidlington, Oxon), pp 185–202
Echenique M, Crowther D, Lindsay W, 1969, "A spatial model for urban stock and activity" *Regional Studies* **3** 281–312
Echenique M H, Hargreaves A J, Mitchell G, Namdeo A, 2012, "Growing cities sustainably: does urban form reallymatter? *Journal of the American Planning Association* **78** 121–137
Echenique M H, Grinevich V, Hargreaves A J, Zachariadis V, 2013, "LUISA: a land-use interaction with social accounting model; presentation and enhanced calibration method" *Environment and Planning B: Planning and Design* **40** 1003–1026
Eliasson J, Mattsson L-G, 2001, "Transport and location effects of road pricing: a simulation approach" *Journal of Transport Economics and Policy* **35** 417–456
Fiorello D, Huismans G, López E, Marques C, Steenberghen T, Wegener M, Zografos G, 2006 *Transport Strategies under the Scarcity of Energy Supply* STEPs Final Report, edited by A Monzon, A Nuijten, Buck Consultants International, The Hague, http://www.steps-eu.com/reports.htm
Forrester J W, 1969 *Urban Dynamics* (MIT Press, Cambridge, MA
Fox J, Daly A, Patruni B 2009, "Improving the treatment of cost in large-scale models", paper presented to European Transport Conference, Noordwijkerhout; copy available from the authors
Fujita M, 1989 *Urban Economic Theory: Land Use and City Size* (Cambridge University Press, Cambridge)
Fujita M, Krugman P, Venables A J, 1999 *The Spatial Economy: Cities, Regions and International Trade* (MIT Press, Cambridge, MA)
Gaudry M, Laferrière R, 1989, "The Box–Cox transformation: power invariance and a new interpretation" *Economics Letters* **30** 27–29
Glaeser E, Gyourko J, 2005, "Urban decline and durable housing" *Journal of Political Economy* **113** 345–375
Graham D J, Kim H Y, 2008, "An empirical analytical framework for agglomeration economies" *Annals of Regional Science* **42** 267–289
Graham D J, Gibbons S, Martin R, 2009, "Transport investment and the distance decay of agglomeration benefits", mimeo, Centre for Transport Studies, Imperial College, London
Hunt J D, Kriger D S, Miller E J, 2005, "Current operational urban land-use–transport modelling frameworks: a review" *Transport Reviews* **25** 329–376
Iacono M, Levinson D, El-Geneidy A, 2008, "Models of transportation and land use change: a guide to the territory" *Journal of Planning Literature* **22** 323–340
Ingram G K, Kain J F, Ginn J R, 1972, *The Detroit Prototype of the NBER Urban Simulation Model* (Columbia University Press, New York)
Ivanova O, Tavasszy L, 2007 *RAEM: Version 3.0* Transport and Mobility Leuven, Leuven
Jin Y, Echenique M, Hargreaves A J, 2013, "A recursive spatial equilibrium model for planning large scale urban change: supplementary working paper", Energy Efficient Cities Project working paper, available on request
Krugman P, 1991 *Geography and Trade* (MIT Press, Cambridge, MA)
Lee D B, 1973, "Requiem for large scale models" *Journal of the American Institute of Planners* **39** 163–178
Leontief W, 1986 *Input–Output Economics* 2nd edition (Oxford University Press, New York)
Lin J Y, Liu Z, 2000, "Fiscal decentralization and economic growth in China" *Economic Development and Cultural Change* **49** 1–21
Lowry I, 1964 *Model of Metropolis* RM-4035-RC (Rand Corporation, Santa Monica, CA)
Lucas R, 1976, "Econometric policy evaluation: a critique", in *The Phillips Curve and Labor Markets* Eds K Brunner, A Meltzer (American Elsevier, New York) pp 19–46
McFadden D, 1974, "Conditional logit analysis of qualitative choice behavior" in *Frontiers in Econometrics* Ed. P Zarembka (Academic Press, New York) pp 105–142

Nagel K, Beckman R J, Barrett, C L, 1999, "TRANSIMS for urban planning", LA-R 984389, Los Alamos National Laboratory, Los Alamos, NM
OECD, 2012 *Compact City Policies: A Comparative Assessment* (OECD, Paris)
Oosterhaven J, Knaap T, Ruijgrok C, Tavasszy L, 2001, "On the development of RAEM: the Dutch spatial general equilibrium model and its first application to a new railway link", paper presented to the 41st Congress of the European Regional Science Association, Zagreb, 29 August – 1 September; copy available from the authors
Redding S J, 2010, "The empirics of new economic geography" *Journal of Regional Science* **50** 297–311
Rice P G, Venables A J, Patacchini E, 2006, "Spatial determinants of productivity: analysis for the regions of Great Britain" *Regional Science and Urban Economics* **36** 727–752
Rosenthal S, Strange W C, 2004, "Evidence on the nature and sources of agglomeration" *Review of Economics and Statistics* **85** 377–393
Rosewell B, 2011, "Planning curses: how to deliver long term infrastructure", Policy Exchange, London
Sheffi Y, 1985 *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods* (Prentice-Hall, Englewood Cliffs, NJ)
Simmonds D, 2001, "The objectives and design of a new land-use modelling package: DELTA", in *Regional Science in Business* Eds G Clarke, M Madden (Springer, Berlin) pp 159–188
Simmonds D, Waddell P, Wegener M, 2013, "Equilibrium v. dynamics in urban modelling" *Environment and Planning B: Planning and Design* **40** forthcoming
Solow R M, Vickrey W, 1971, "Land use in a long narrow city" *Journal of Economic Theory* **3** 340–447
The Conservative Party, 2010, "Open source planning: green paper", The Conservative Party, London, http://www.conservatives.com/~/media/Files/Green%20Papers/ planning-green-paper.ashx
UN Habitat, 2008 *State of the World's Cities* (Earthscan, London)
UrbanSim, 2011 *UrbanSim* http://urbansim.org
Venables A J, 1996, "Equilibrium locations of vertically linked industries" *International Economic Review* **37** 341–359
Volterra ,CBP, 2007, "Modelling transport and the economy in London: framework and literature review, Volterra Consulting Ltd and Colin Buchanan and Partners Ltd, London
Waddell P, 2002, "UrbanSim: modelling urban development for land use, transportation and environmental planning" *Journal of the American Planning Association* **68** 297–314
Wegener M, 1982, "Modelling urban decline—a multilevel economic–demographic model for the Dortmund region" *International Regional Science Review* **7** 217–241
Wegener M, 2001, "The IRPUD model", Institute of Spatial Planning (IRPUD), Dortmund
Wegener M, 2005, "Urban land use transportation models", in *GIS, Spatial Analysis, and Modeling* Eds D J Maguire, M F Goodchild, M. Batty (ESRI Press, Redlands, CA) pp 203–220
Wegener M, 2011, "Transport in spatial models of economic development", in *Handbook in Transport Economics* Eds A de Palma, R Lindsey, É Quinet, R Vickerman (Edward Elgar, Cheltenham, Glos) pp 46–66
Wegener M, Gnad F, Vannahme M, 1986, "The time scale of urban change", in *Advances in Urban Systems Modelling* Eds B Hutchinson, M Batty (North-Holland, Amsterdam) pp 145–197
Wenban-Smith A, 2011, "Are current appraisal methods adequate?" *The Transport Economist* **38** 2–10
White R, Engelen G, 2000, "High resolution integrated modelling of the spatial dynamics of urban and regional systems" *Computers, Environment, and Urban Systems* **24** 383–4000
Williams H C W L, 1977, "On the formation of travel demand models and economic evaluation measures of user benefits" *Environment and Planning A* **9** 285–344
Williams I N, 1994, "A model of London and the South East" *Environment and Planning B: Planning and Design* **21** 535–553
Wilson A G, 1967, "A statistical theory of spatial distribution models" *Transportation Research* **1** 253–269

Wilson A G, 2000 *Complex Spatial Systems: The Modelling Foundations of Urban and Regional Analysis* (Pearson Education, Harlow, Essex)
WSP UK Ltd, 2005, "The EUNET2.0 freight and logistics model—final report", Regional Pilot for Economic/Logistic Methods, DfT Contract PPAD 9/134/24, Department for Transport, London, www.dft.gov.uk/pgr/economics/rdg/ coll_theeunet20freightandlogisti/ theeunet20freightandlogistic3127
Zhang P (Ed.), 2010 *Strategic Research for the Twelveth Five Year Plan* (People's Publishing House, Beijing)
Zhu J, 2012 *A Spatial Computable General Equilibrium Model for London and Surrounding Regions* PhD dissertation, Martin Centre for Architectural and Urban Studies, University of Cambridge
Zondag B, de Jong G, 2011, "The development of the TIGRIS XL model: a bottom-up approach to transport, land-use and the economy" *Research in Transportation Economics* **31** 55–62