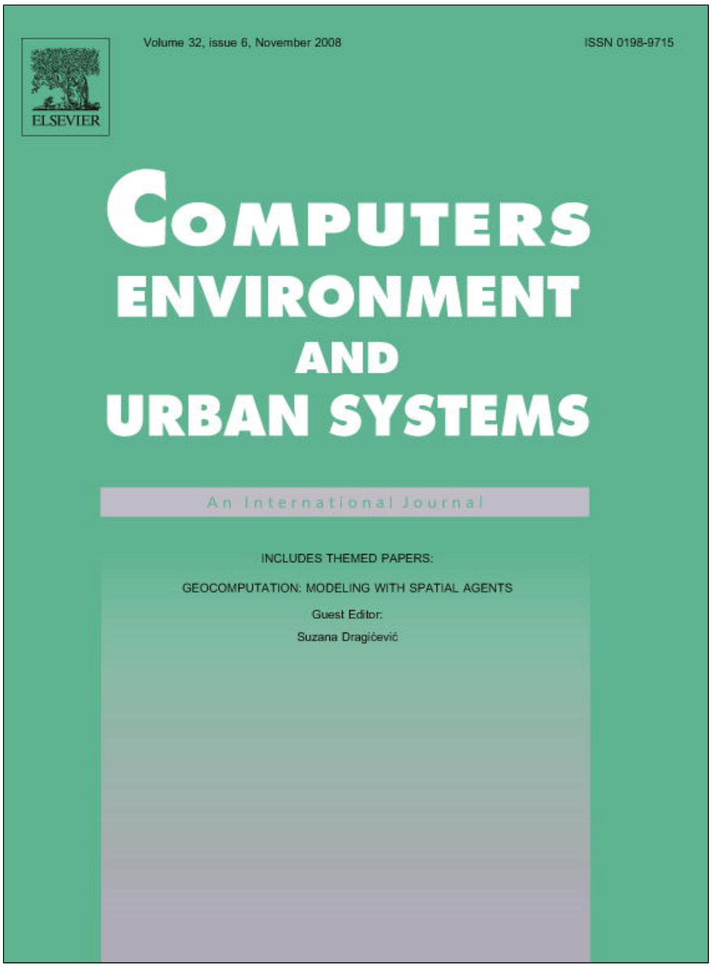

Volume 32, issue 6, November 2008
ISSN 0198-9715
ELSEVIER
COMPUTERS
ENVIRONMENT
AND
URBAN SYSTEMS
An International Journal
INCLUDES THEMED PAPERS:
GEOCOMPUTATION: MODELING WITH SPATIAL AGENTS
Guest Editor:
Suzana Dragićević

Computers, Environment and Urban Systems 32 (2008) 440–453

Contents lists available at ScienceDirect

Computers, Environment and Urban Systems

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

A spatial microsimulation model with student agents

B.M. Wu *, M.H. Birkin, P.H. Rees

School of Geography, University of Leeds, Woodhouse Lane, Leeds LS2 9JT, UK

ARTICLE INFO

Keywords:
Spatial microsimulation
Agent-based model
Spatial interaction model
Migration
Urban planning
Policy scenario

ABSTRACT

In this paper, we present a dynamic simulation model which projects the future population of the city of Leeds as a basis for policy analysis and scenario planning. We argue that microsimulation modelling is not entirely effective in the representation of student populations. Alternative approaches using both spatial interaction models and student agents are presented and evaluated. The results from the agent-based model are found to be particularly encouraging. We suggest that agent-based modelling and microsimulation are powerful as complementary technologies for individual-based modelling.

1. Introduction

This paper reports on research from a project which aims to produce a simulation model of the UK population, as it now is and as it can be expected to develop over a twenty-five year time horizon. The project is called Moses, and is additionally concerned with the potential effects of demographic change on services and infrastructure, particularly those relating to housing, health care and transportation (Birkin et al., 2005). If planners were equipped with the means (through simulation) to understand social and demographic changes in response to shifts in policy, such a device would have valuable practical applications as both a 'decision support system', and as a pedagogic tool for understanding how cities work. As an academic and intellectual challenge, the ability to reproduce and predict the behaviour of real city systems must demonstrate a deep understanding of such systems.

The technique of microsimulation has been adopted at the heart of our approach. The benefits of microsimulation (in contrast to macroscopic modelling approaches with similar objectives) within a demographic modelling context have been argued persuasively and eloquently by van Imhoff and Post (1998). In particular, these authors demonstrate the richness of microsimulation as a device for the representation of both relationships between members of a population, and of the transitions between states within a population. Microsimulation has a 50-year history as a technique within economic analysis (Orcutt, 1957), while more recent applications of spatial microsimulation have embraced problems such as transportation, healthcare and housing (Brown & Harding, 2002; Magne, Rabut, & Gabard, 2000; Wilson & Pownall, 1976). Although simulation modelling has been rather unfashionable amongst geographers in recent years, we follow authors such as Hamnett (2003), Johnston et al. (2003) and Unwin (2005) in arguing for a reappraisal of modelling methods on the basis of policy relevance, but also because there has been such a marked change in the background conditions of applied geographic research. To an extent, these changes are encapsulated within the notion of e-(social)-science in the UK and Europe, or cyberinfrastructure in North America (Foster, 2003).

The literature on spatial modelling reflects an increasing interest in models which represent the behaviour of individual entities or "agents" and has generated policy-relevant applications to problems of disease control (Barrett, Eubank, & Smith, 2005; Eubank et al., 2004; Ferguson et al., 2005), transportation (Transims Travelogue, 1996) and urban energy markets (Mozumder & Marathe, 2004). Whilst there are also limited examples of the incorporation of agent-based models (ABM) into demographic models (Billari, Ongaro, & Prskawetz, 2002; Espindola, Silveira, & Penna, 2006; Loibl & Toetzer, 2003; Makowsky, Tavares, Makany, & Meier, 2006) this paper represents an innovative attempt to combine ABM with microsimulation.

In Section 2 of the paper, we present a microsimulation model in which migration between small geographic areas is driven by transition rates. The difficulty with transition rates is that they are dependent on the background conditions at the time they were captured. For this reason, a spatial interaction model (SIM) of migration flows is introduced in Section 3.

In the SIM, flows are modelled in relation to the characteristics of both the origin and the destination areas, as well as the distance or travel cost between them. However the second approach, of spatial microsimulation model (MSM) + SIM, is also shown to be inadequate. The most particular weakness of the model is in relation to areas with high student populations.

* Corresponding author. Tel.: +44 (0)113 34 36838; fax: +44 (0)113 34 33308.
E-mail addresses: B.Wu@Leeds.ac.uk (B.M. Wu), M.H.Birkin@Leeds.ac.uk (M.H. Birkin), P.H.Rees@Leeds.ac.uk (P.H. Rees).

0198-9715/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compenvurbsys.2008.09.013

For these reasons, it may be more appropriate to model student migration using a set of behavioural rules, and the ideal framework for this is an agent-based model. In the fourth section of the paper, we therefore propose a model in which student migration is represented using a population of agents.

In the discussion which concludes the paper, we appraise the importance of this combination of approaches – MSM (microsimulation model) + SIM (spatial interaction model) + ABM (agent-based model) – for demographic simulation.

2. Model specification

Van Imhoff and Post (1998) pointed out that "macrosimulation" and "microsimulation" models (MSM) are alternative ways of representing the same demographic processes. Moses uses a spatial microsimulation model (MSM) for modelling the population and its dynamics, but the model structure parallels the macro multi-state cohort-component projection model. Populations are projected at the ward level to enable us to study the characteristics of the population within a local context. A synthetic population with rich attributes is simulated at the individual level to provide detailed information for decision making.

The model comprises two components: a baseline population reconstruction model (PRM) and a dynamic simulation model. The difficulties in constructing dynamic MSM at a national level have been noted by Harding (2007), although models such as SVERIGE (Rephann & Holm, 2004) and its antecedent CORSIM (Caldwell, Clarke, & Keister, 1998) have provided a template for modelling at the regional scale. Intra-regional or local applications have tended to concentrate on comparative static reweighting of individual populations to future projections or targets (Murphy, 2000; O'Hare, 2000). However such projections are not available for small areas in the UK. Our intention in building the dynamic model is therefore to create small area forecasts which are valuable for scenario planning and spatial policy analysis.

In this section of the paper, a microsimulation model of the population and its dynamics is described. The application of the model to the urban area of Leeds is introduced. The urban area of Leeds, a city of 730,000 people in the north of England, is used for illustrative purposes throughout this paper. Some key features in the geography of the city and which are significant in the spatial analysis later in the paper are shown in Fig. 1. The baseline model which we describe below is available for the whole of the UK, while the dynamic model uses parameters which have to date only been computed for the city of Leeds. The experiments with student agents which are described later (see Section 4 below) have currently been undertaken for the Leeds area only.

The baseline model (PRM) creates a complete synthetic representation of 24 million households and 60 million individuals in the UK. The data is reweighted from the UK census sample of anonymised records (SAR) which provides microdata for 1% of UK households from which the spatial codes have been aggregated into regions to protect individual confidentiality. The SARs are reweighted in accordance with known distributions from the census small area statistics – a procedure originally introduced by Williamson, Birkin, and Rees (1998) (see also Ballas et al., 2005). The PRM (population reconstruction model) is described in more detail in Birkin, Turner, and Wu (2006).

Six important demographic processes have been modelled in the dynamic model: ageing, mortality, fertility, health change,

Fig. 1. Map of Leeds area.

marriage and migration (Fig. 2). Transition probabilities for each of these events are applied at discrete one year intervals. Many different orderings are possible for this sequence of events. Rephann and Holm (2004) consider a model which also includes education and employment, in which fertility precedes marriage, and mortality follows migration. We consider that it is more logical to evaluate fertility following the formation of marriages and partnerships. Mortality is considered early in the process for practical reasons, since if an individual lifecourse is terminated subsequent processes can then be ignored. The operation of these processes also leads to the formation and dissolution of households within the simulation. Due to the nature of the demographic events, some processes are more complicated than the others and need to be modelled in multiple stages, for example, marriage and migration processes may often be connected, as a change in marital status will frequently occur alongside the move to a new home. A more detailed discussion of the dynamic model is provided by Birkin and Wu (2008).

Population is annually projected at the ward level currently by gender and single year of age. The population at the end of the year is equal to the population at the beginning of the year (starting population) plus births and in-migration, minus deaths and out-migration. The model projects each component of population change separately, but each component of change affects the others. The rate of change for each of the components depends on both observed historical trends in the region and on forecast national trends. The model will provide simulation results of the UK population for 30 years (2001–2031). However, all probability calculations are based on the most up-to-date census data in 2001. The most up-to-date BHPS data (2000–2004) is also used in order to capture the variations in migration behaviour across 10 key demographic attributes – age, ethnicity, gender, headship, health, household size, housing type, marital status, occupation, tenure.

Moses dynamic model v1.0 has been released and the initial analysis of 30 years' simulation results has been carried out. The

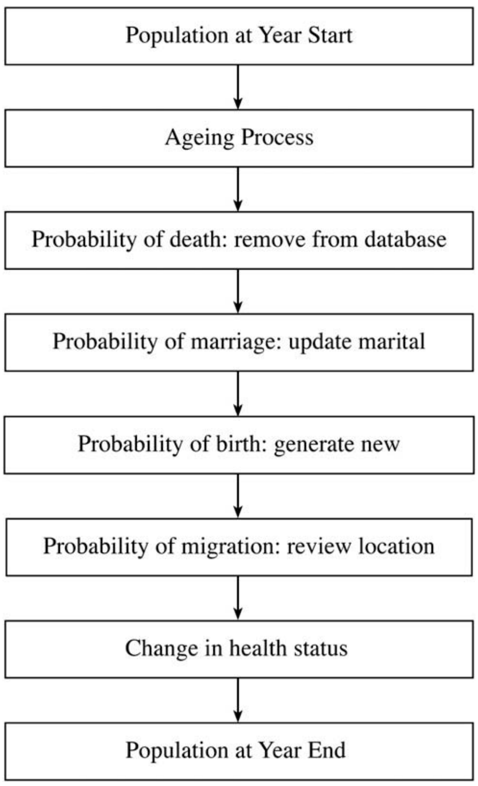


Fig. 2. Process of the population simulation.

population changes over time such as the ageing pattern can be clearly observed from population pyramids generated by the MSM (microsimulation model) to facilitate high level decision making. At this stage, we have not attempted to reconcile the model projections with other forecasts, such as the office for national statistics projection for 2029 (ONS, 2007), as shown in Fig. 3. There are three important sets of differences between the projections. The Moses MSM assumes that age-specific mortality rates will stay constant over the next two decades, whereas ONS expects a steady improvement in life expectancy. Hence the elderly populations are somewhat larger in the ONS projection. Moses MSM also assumes constant fertility rates, whereas ONS expects these rates to rise. Hence the number of young people is also higher in the ONS projection. Finally, student migration into Leeds, mostly in the 16–25 age range, is substantially greater in the ONS projections because these projections allow for acceleration in student numbers following the 2001 census.

The mortality and fertility assumptions are not of crucial importance in the present paper. As far as student migration is concerned, the patterns are of great interest, but mostly in relation to the distribution of migrants within the city of Leeds. Because the emphasis in the current paper is on spatial distributions, we did not consider it necessary to revise the student population upwards. Alternative scenarios, including those which match government projections much more closely, are considered elsewhere (Birkin & Wu, 2008).

The migration process within Leeds is modelled using the 2001 census special migration statistics (Level 2) data, which provides us the ward-based migration flows from one ward to another. The 33 local authority wards in Leeds are as shown in Fig. 1. Having incorporated the migration impact in MSM, the small area demographic change is then assessed using location quotients as a measure of the concentration of a particular group within each geographical area at a point in time. The population structure is analysed in five year age bands. Where the location quotient is greater than 100, this indicates an over-representation of the population group in area, and vice versa. Some examples are shown in Fig. 4.

Aireborough and Cookridge are both established suburban areas in the north of the city. In these areas, the changing concentration of demographic groups over time looks reasonable. For example, in Cookridge there is initially an expansion in the very elderly population, but later the quotient for the elderly falls relative to the rest of Leeds as other areas begin to experience a similar growth in the older age groups. However in wards where student migration has a great impact like University and Headingley, the MSM failed to reproduce the student population renewal. The peaks of young people aged 20–25 disappears after 10, 20 and 30 years simulation. The distinctive population structure in such areas has been lost in the small area projections.

This indicates that the subtlety of the local migration patterns has not been captured successfully in our MSM (microsimulation model). The pure spatial MSM used here cannot differentiate students from other migrants in the migration process. The MSM is probability driven. Migration probabilities at small area scale have been generated using the 2001 census special migration statistics (Level 2) data. The data provides us the ward-based migration flows from one ward to another for migrants, but we cannot determine the number of the university students within each flow. Therefore student migrants are modelled exactly the same as the rest of the migrants in the pure MSM using the generalised probabilities. Another factor affects the results could be that the migrants are not disaggregated by age. However, even if we refine the model with disaggregated ages, it still cannot capture the distinctive migration pattern of student migrants. In reality students are highly mobile during their study in the universities and mostly

Fig. 3. The age–sex structure of the Leeds population in 2001 and 2029 generated by the MSM compared with the 2029 population estimate of national statistics. Source: 2001 Census: standard area statistics (England and Wales).

only move around the area close to the universities where they study, not in the suburban areas. More importantly, most of them will leave the city once they finish their study, instead of settling down and growing old in the area. Due to the replenishment of the student population each year, the population of the wards in which university student stay tends to remain younger than that in other wards.

Given the above reasons, a hybrid modelling approach is adopted to better model the local migration patterns. In the next sections of the paper, we explore a hybrid approach combining SIM and ABM (agent-based model) techniques.

3. Spatial interaction models (SIM) for migration

3.1. Background

The outline simulation model at Section 2 incorporates migration flows for individuals and households entering or leaving an area which are disaggregated by age and sex, but which does not recognise either variations in age–sex-specific rates of migration between areas, or the spatial pattern of migration between small areas. For example, we cannot incorporate lifecycle effects with high migration of young people into the centres of cities, and the outward movement of families towards the leafy suburbs. In this section of the paper, we describe the use of a spatial interaction model to represent variations in the pattern of migration flows between small geographical areas. After a general review, the modelling is described as a two stage process: first, migrant populations are generated within small areas (Section 3.2); then the migrants are distributed between receiving areas (Section 3.3). The use of the spatial interaction modelling probabilities is described in Section 3.4.

The classic 'family of spatial interaction models' (Wilson, 1967; Wilson, 1974) considers possible alternative models in relation to possible constraints on the migration process. Fotheringham et al. (2004) considered an origin-constrained spatial interaction model as the most appropriate representation of migration flows between local government regions in the United Kingdom and explore a wide range of pull factors which can affect the attractiveness of different destinations, including house prices, environmental quality, and so on. It then becomes possible to link specific policy drivers to the attractiveness of destinations – for example, if more housing is made available then the area becomes more attractive and generates extra migrants.

In our model, in-migrants and out-migrants are explicitly linked. The migrants who leave an area automatically create housing vacancies for new migrants seeking to enter that area. A number of simplifying assumptions are adopted. Firstly, all flows into or out of a city region are ignored. Second, that there is no vacant housing stock. Third, that all moves involve complete households. Fourth, that the probability of migration is driven by two factors – proximity to the existing residence, and the availability of suitable housing at a destination. The combined effect of these assumptions is that the process can be represented as a doubly-constrained spatial interaction model (DCSIM) in which vacancies are simultaneously created and filled by mover households. The interaction terms take the following form:

$$\boldsymbol{T}_{ij}^{km} = \boldsymbol{A}_i^k \boldsymbol{O}_i^k \boldsymbol{B}_j^m \boldsymbol{H}_j^m \mathrm{e}^{-\beta d_{ij} + \lambda^{km}} \tag{1}$$

in which the flow of migrants ($\boldsymbol{T}_{ij}^{km}$) is positively related to both population ($\boldsymbol{O}_i^k$) and housing ($\boldsymbol{H}_j^m$), and inversely related to the distance between areas ($\boldsymbol{d}_{ij}$). The full specification of the model and the means for its calibration is discussed in Appendix A.

3.2. Migrant generation model

In order to calibrate the variation in rates of migration between different demographic groups, we used data from the British household panel survey (BHPS). The BHPS in a longitudinal survey which has been tracing a demographic cohort of about 16,000 individuals in annual waves since 1991 includes a very wide array of demographic and household characteristics (University of Essex, 2005). Migrants are identified from a specific question within the BHPS, which "indicates whether sample members have moved since (the) last wave" (Taylor, Brice, Buck, & Prentice-Lane, 2005). When successive waves are connected from the survey, it is possible to determine whether an entire household has moved, whether a single individual has moved from one household to another, or

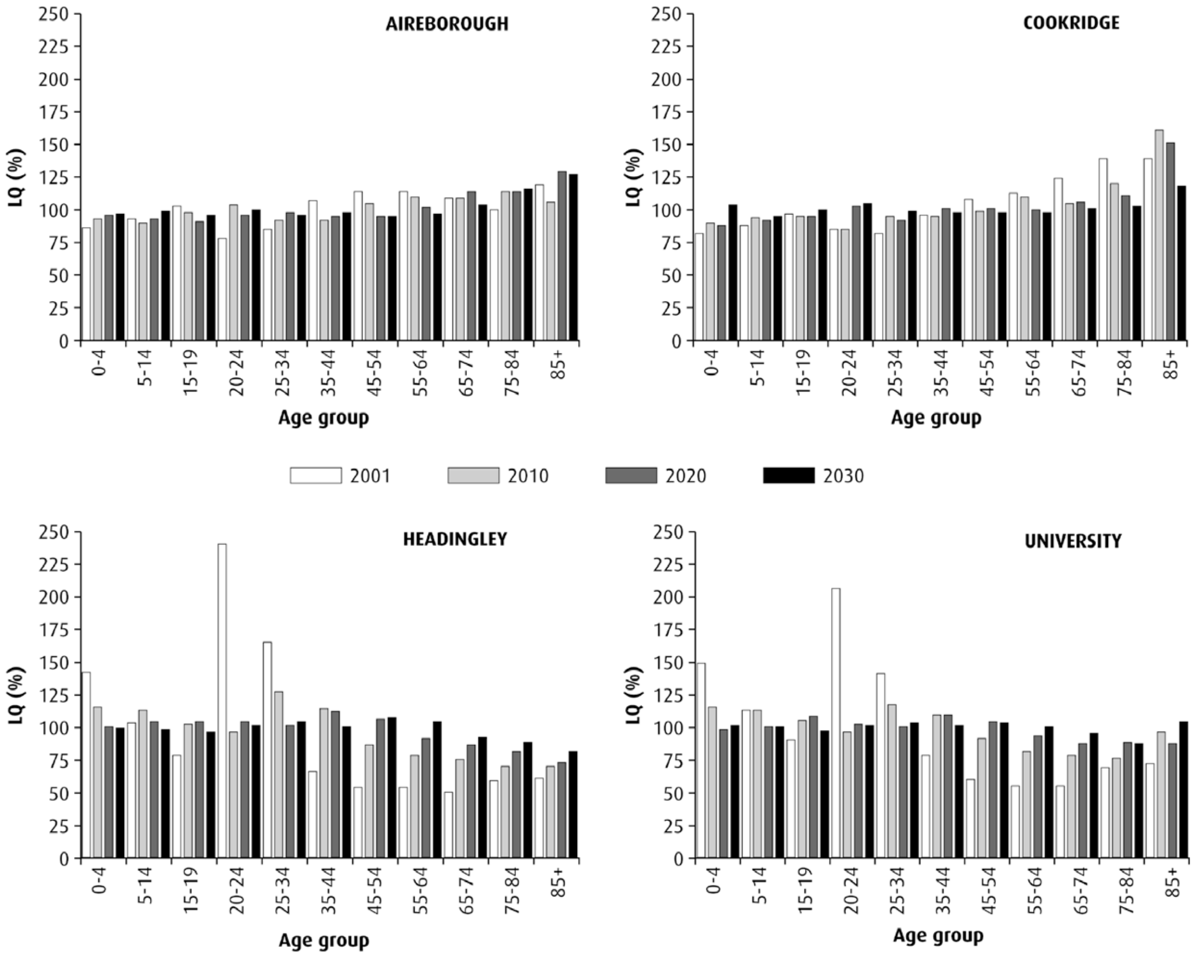


Fig. 4. Location quotient analyses of ward level projections. Source: 2001 Census: standard area statistics (England and Wales).

whether a household has fragmented into one or more constituent parts.

Data from five waves of the BHPS (2000–2004) were analysed for variations in migration behaviour across ten key demographic attributes – age, ethnicity, gender, headship, health, household size, housing type, marital status, occupation, tenure. Other attributes were either unavailable or considered to be inappropriate or unnecessary. For example, income is not available in the UK census and therefore not represented in the population database. Education is proxied by occupation which includes 'full-time students' as a category. It was quickly established that partial household moves are unimportant relative to individual moves and whole household moves. In BHPS wave 13 (2004) for instance, only 15 out of 993 total moves were of this type. The analysis therefore concentrated on individual and household moves.

A stepwise chi-squared procedure was implemented to rank the demographic variables in order of their importance as predictors of migration. If one variable ranks lower than another as a potential predictor, then the effects of the lower ranking variable are only assessed after allowing for the effect of the higher ranking variable. The estimator is similar to CHAID as used within SPSS and elsewhere (Hill & Lewicki, 2006, chap. 7). Stepwise estimators are widely employed in bioinformatics applications (Ott & Rao, 1985). The computation is straightforward and so a bespoke computational procedure was deployed. The results for both individual and household moves are summarised in Table 1, from which it can be seen that a small number of characteristics are highly significant. For household moves, then the key factors are age of head, house type and household size. A number of other variables are less important, although still statistically significant with 99% confidence – tenure, occupation, marital status and health. Other variables found to be insignificant include ethnicity and gender. The equivalent interpretation for the drivers of individual moves is evident from Table 1.

As a basis for the model, it was decided to limit the drivers to those three characteristics which were found to be overwhelmingly significant in each case. This gives a model which is sparse and robust, in which there are no potential doubts about the sign

Table 1
Chi-square significance measures for demographic predictors of migration.

	Mover type			
	Individual	Sig.	Household	Sig.
Age	1265	–**	1036	–**
Ethnicity	22	–*	6	
Gender	0		5	
Headship	2		0	
Health	32	–*	22	–*
Household size	290	–**	80	–**
Housing type	27	–*	184	–**
Marital status	114	–**	25	–*
Occupation	19	–*	28	–*
Tenure	2		28	–*

Source: 2001 Census: standard area statistics (England and Wales).
* Indicates significant with 99% confidence.
** Indicates 'overwhelming' significance: chance occurrence of less than one in one thousand.

or significance of individual relationships. Students are characterised here by their age and mover type – they are typically aged between 18 and 25 and either leaving a family home or moving as single person households. However in the SIM they are not differentiated explicitly from other young people with these characteristics.

3.3. Migrant distribution model

The model has a single distance parameter (β) and a vector of household choice parameters for people type k living in houses type m (λ^{km}). The distance parameter is calibrated using data from the 2001 census special migration statistics (SMS), and the household choice parameters are calibrated to match observed housing preferences within the BHPS. This calibration is achieved using a maximum likelihood estimation procedure, in which parameter estimates are successively rescaled to reflect the divergence in model distributions from the known data. The procedure generates stable and convergent parameter estimates within a relatively short space of time: for example, with a convergence threshold of 0.001 (0.1%) then stable estimates are achieved after 13 iterations for each parameter.

The average distance travelled by inter-ward migrants is 3.7 km, with a parameter estimate of $\beta = 0.35$. (More than 40% of moves are short distance moves within an existing ward). Stable values for the household choice parameters (λ) are illustrated in Table 2. These estimates show an intuitively reasonable pattern in which small households with young heads show the highest propensity for flats and terraces. Indeed single person households prefer flats in all age groups, although this preference becomes progressively weaker through the lifecourse. Larger households tend to show a preference for detached and semi-detached housing, and this pattern is most pronounced in the middle age ranges.

3.4. Combining the two models

Having generated aggregate flows within the DCSIM, we now need to assign individual destinations within the microsimulation model. The obvious approach to this problem is to take a set of probabilities from the spatial interaction model, and to sample individual preferences using the normal Monte Carlo procedures. The objective is to take a migrant with known demographics, and to assign probabilities for household type and destination. These probabilities will look like this

$$p(j,m|i,k) = \frac{\boldsymbol{T}_{ij}^{km}}{\sum_{jm}\boldsymbol{T}_{ij}^{km}} \tag{2}$$

This methodology has been adopted recently in the context of simulation recent trips by Nakaya, Fotheringham, Clarke, and Ballas (2007) and originally suggested by Birkin and Clarke (1987).

The fit of the model can be assessed through a regression of observed against predicted interactions. This generates a satisfactory r^2 value of 0.6446, so that the model explains just under two-thirds of the variance in local migration patterns. However this overall result masks significant spatial variations in the correlation between flows for individual wards, which are shown in Fig. 5. There is a general pattern for the poorest fits to be in the more central areas, including but not limited to those wards which are most heavily influenced by student migration (such as Weetwood – see Fig. 1). For example, Seacroft has a very high level of intra-ward moves, but because this is quite a small zone in proximity to several others, the model tends to predict quite a lot of cross-boundary movement. The model is most robust in the free-standing peripheral towns, e.g. Wetherby, Pudsey, Otley and Morley.

4. An agent-based model of student migration

4.1. The need for a hybrid approach combining MSM and ABM

Microsimulation models (MSM) have become one of the principal techniques in demographic modelling. Being a widely discussed and applied instrument in studying and predicting the evolution of population using an individual-based model (IBM), MSM is as important to the analysis of event histories as macrosimulation is to traditional aggregated demographic analysis. It is now generally accepted that the explanation of demographic behaviour requires a microscopic approach, however, there is a gap between the microdemographic theory and demographic techniques. In demography, there is a poor level of precision in theoretical constructions and of particular importance it lacks sufficient theory at the basis of the applications of statistical models and data collection. Sometimes there is also insufficient accounting for observability of important quantities used in the theory (Billari et al., 2002).

The combined demographic MSM and spatial interaction modelling approach has been successful to a degree in providing a rich representation of population change in a city region. However certain problems, particularly in relation to student housing areas, have proved to be rather intransigent to both of these methods.

Table 2
Variations in the affinity parameter (lambda).

Age group	Household size					
	One person housing type			Two people housing type		
	Detached	Terraced	Flats	Detached	Terraced	Flats
16–24	0.16	0.72	2.10	0.19	1.06	1.75
25–34	0.21	0.90	1.86	0.72	0.88	1.38
35–44	0.47	0.73	0.74	0.07	0.91	0.03
45–54	0.81	0.54	0.62	0.60	0.97	0.49
55–64	0.80	0.82	1.36	1.65	0.68	0.73
65–74	1.02	0.61	1.37	1.71	0.71	0.65
75+	1.08	0.53	1.41	1.44	0.69	0.92
	Three people housing type			Four or more people housing type		
16–24	0.31	1.34	1.36	0.86	1.36	0.81
25–34	0.95	1.09	0.98	1.36	1.32	0.39
35–44	1.51	1.25	0.33	1.77	1.22	0.17
45–54	1.87	1.15	0.16	2.17	1.38	0.02
55–64	2.00	1.31	0.06	2.00	1.17	0.10
65–74	2.00	1.41	0.04	1.67	0.62	0.77
75+	2.27	0.98	0.06	2.85	2.48	0.00

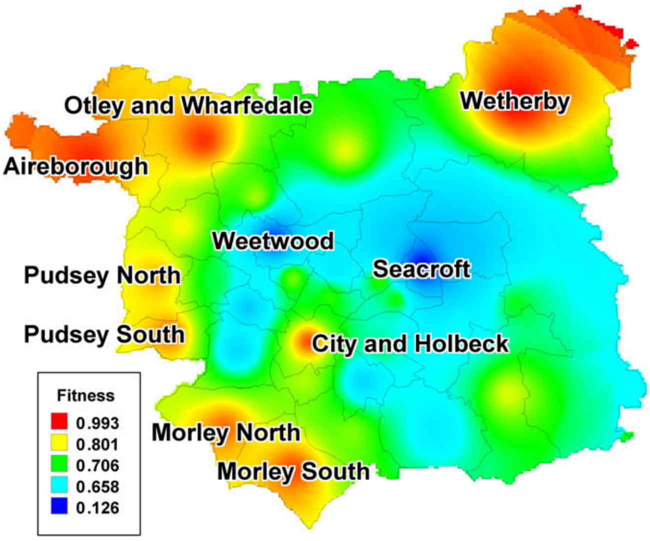


Fig. 5. Goodness of fit for the SIM, by area. Source: 2001 Census: standard area statistics (England and Wales); Ward Boundaries. Crown copyright 2003.

For instance, in wards where student migration has a great impact, the microsimulation failed to reproduce the renewal of the student population and they grow old in these areas as other people do. The problem of modelling student migration is quite well-known in small area demographic modelling (see for example Baryla & Dotterweich, 2001; Fotheringham et al., 2004; Koser & Salt, 1997; Rees, 1994). We know however for a fact that university students tend to only stay in such areas during the period of their study and then leave while other new students move in. Due to the replenishment of the student population each year, the population in such wards stays younger than that in other wards.

One possible refinement to the SIM (spatial interaction model) would be to include further disaggregation of the model parameters by age. This can soon make the models unwieldy. Agent-based modelling (ABM) is an alternative approach that can model individual behaviours through multiple agents. In an ABM, each agent follows their built-in rules and makes decisions and takes actions on the basis of the rules and the knowledge gained through interactions with each other and the environment they live in. Through such interactions, simple and predictable local interactions can generate familiar but unpredictable global patterns. Therefore ABM provides theoretical leverage where the global patterns of interest are more than the aggregation of individual attributes.

MSM and ABM have been closely allied, but with significant difference in terms of

1. Programming approach: MSM uses a more traditional programming approach that uses a central control to manipulate data flow, i.e.: A central manipulation mechanism in the system processes, the input data and outputs of the results. Whilst ABM adopts object oriented programming (OOP) that models through objects with their own attributes and methods. All objects process the input data and produce results according to their built-in methods (Russell & Norvig, 1995).
2. Computing method: MSM (and SIM) is more statistical/probability driven simulation where each event occurrence is based on relevant probabilities. Whilst ABM is more of a rule-based (built-in intelligence) simulation process. These rules can, but do not always have to incorporate probabilities or equations (Billari et al., 2002).
3. Interaction and behaviour modelling: Due to its modelling approach, MSM lacks useful means for interaction and behaviour modelling. Whilst each agent in ABM makes their decisions and takes actions on the basis of the interactions with other agents and the environment they live in, according to their own built-in rules. This provides a very flexible way to model the heterogeneous behaviours of different agents (Conte & Gilbert, 1995).

Another approach in need of mention is CA (cellular automata). CA is more similar to ABM in the sense that they are both object-oriented and uses agents. However ABM provides agents far more capable of mobility, interactions and behaviours, as CA agents are confined within cells and the simulation is driven by the changes of the states of each cell.

Although there are limited examples of using ABM approach in demographic models, Billari et al. (2002) consider ABM as a promising approach to help improve our understanding of demographic behaviours. ABM can provide us an alternative means to study demographic processes as the outcome of interacting agents. It is also useful in explaining individual behaviours by taking into account both micro and macrofactors. For instance, agents in an artificial society can follow their built-in rules while subject to the macro behaviour rules (e.g. conventions, institutions). The explanation of behaviour is based on simple propositions about individual behaviour, but can produce complex situations and feedback at the macro level. By focusing on dynamics of the population instead of equilibria, ABM is better suited for modelling demographic processes. The question whether those dynamic changes are due to compositional changes or changing behaviour rules of the individual agents can also be studied within the framework of ABM.

Migration is a complex demographic process where interactions and behaviours play an important role. Using ABM, individual activities and diversity of migration decisions leading to the observed complex migration patterns can be simulated in detail. Some attempts have produced fruitful outcomes. Espindola (2006) analysed the rural–urban migration using ABM, where the migration of workers is modelled as a process of social learning by imitation. As emergent properties of the model, transitional dynamics are observed with continuous growth of the urban fraction of overall population towards equilibrium. While Loibl and Toetzer (2003) studied urban sprawl patterns through modelling suburban migration and residential area occupation. Distinctive migration behaviours of households with varying socio-economic status have been simulated in an ABM. Makowsky et al. (2006) build an ABM to simulate crisis-driven migration of agents within a multi-ethnic population. This study reveals that cultural networks temper an agent's security calculus, with strong social ties dampening the human security dilemma.

The above studies all demonstrate that ABM (agent-based model) is very useful in understanding the functioning of complex models and in modelling particular phenomena without the need of being limited to mathematical tractability. ABM to a degree provides a link to the theory and knowledge gap in demographic modelling. ABM is also a much more flexible way to model various behaviours and dynamics during the migration process, as different type of agents can have their own rules. Based on such rules and the information they gathered from the environment they live in, they can make different decisions and act on it accordingly. On the other hand, MSM (microsimulation model) is a widely applied and tested approach in demographic modelling as discussed previously. While ABM is useful in modelling features in the model where knowledge and theory is lacking, MSM can provide important statistical mechanisms that ensure the similarity between what it predicts and what is actually observed in the gathered data. For instance, the probabilities used in the MSM that are calculated using the empirical data can provide a guideline for the population evolution patterns. This study is one of the first attempts to bring the strength of the two approaches together to provide a better demographic model.

Leeds has been attracting students from all over the UK, due to its central geographical location in and its reputation for university education. However, most students will leave the area when they finish their study, for example to take up employment elsewhere. Therefore student migration is an important component of Leeds migration. Using the pure MSM, it is difficult to model the distinctive behaviours of individual student migrants. However, as a result of students joining the general migration process as described in previous sections, the MSM results in a considerable number of students being found in suburban areas and many students continue to stay in Leeds after their study period and grow old in the central area. This projection is not an accurate reflection of reality.

Therefore a hybrid approach combining MSM and ABM techniques is adopted to strengthen the modelling of the subtlety of the local migration patterns in our model and the behaviour modelling of the student migrants, as this is less well studied and lacks an appropriate theoretical basis in MSM. The main advantage of using ABM here is that it is relatively easy to introduce heterogeneous agents with distinctive behaviours that are not mathematically tractable, as ABM is most useful for problems where "writing down equations is not a useful activity" (Billari et al., 2002). It also provides us a way to experimentally test hypotheses on decision processes and behaviours at an individual level.

4.2. ABM construction

The student population has been selected from the baseline population to form a sub-group based on two criteria: age and full-time student status. The assumption we made for age is based on empirical data from the Higher education statistics agency (HESA): admission data 2004–2005. As most people started at age 17, 18 or 19 for their 1st degree study, most people will finish their PhD program before 30. Therefore the people aged 17–29 who are also "full-time students" have been selected as the baseline population of students for the initialisation.

We then recognise the following groups of students on the basis of age:

- First year undergraduates
- Other undergraduates (years 2 and 3)
- Master students and
- Doctoral students

The reason for such groupings is that the four types of students have distinctive behaviours during the process of migration. For instance, the first year students tend to stay in university accommodations. From the second year on, as students are more familiar with the area, they move out and find privately rented accommodation that are close to the university where they study, often with their fellow students in the same area. The transitional probability to a higher degree of study is generated by comparing counts of students in different degrees of studies in the HESA admission data for the current and the following year, for instance, during the period of 2004–2005 and 2005–2006.

The number of vacancies and their spatial distribution comes from three sources:

- Accommodation at University of Leeds: http://www.leeds.ac.uk/accommodation/overview.html
- Accommodation at Leeds Metropolitan University: http://www.leedsmet.ac.uk/visiting/accomm/index.htm
- Privately rented accommodations from the Unipol database[1] at http://www.unipol.org.uk/leeds/.

[1] Unipol is a charity working nationally in student housing and Unipol Leeds is the largest student housing provider in Leeds.

The ABM of student migration simulation is based on the following assumptions/hypotheses:

- Students prefer to stay with their fellow students – choose to stay the area where student population is.
- Students prefer to stay close to the university, subject to the availability of accommodation. Due to the simplicity of the model at this stage, price and other attributes such as housing quality are not modelled.

Such assumptions are based on the behaviour in housing choice that has been recognised in previous studies. Many studies including Schelling's (1971) famous model of housing segregation have suggested that people have the tendency to live where "similar" types of residents are. The criteria can be based on characteristics such as ethnic group and social economic class. Here we just focus on the different renting behaviour between the four types of university students.

Based on the above assumptions, we then apply the following general rules to the "student agents":

- Each group is allowed a set number of years to stay in an area (based on their study period in the universities).
- Students stay close to their university of study, subject to housing availability.
- They are excluded from the processes of marriage and fertility during their study (we recognised the simplification of this, but with limited research carried out on student migration, there are no suitable rates for student marriage and fertility specifically for university students).

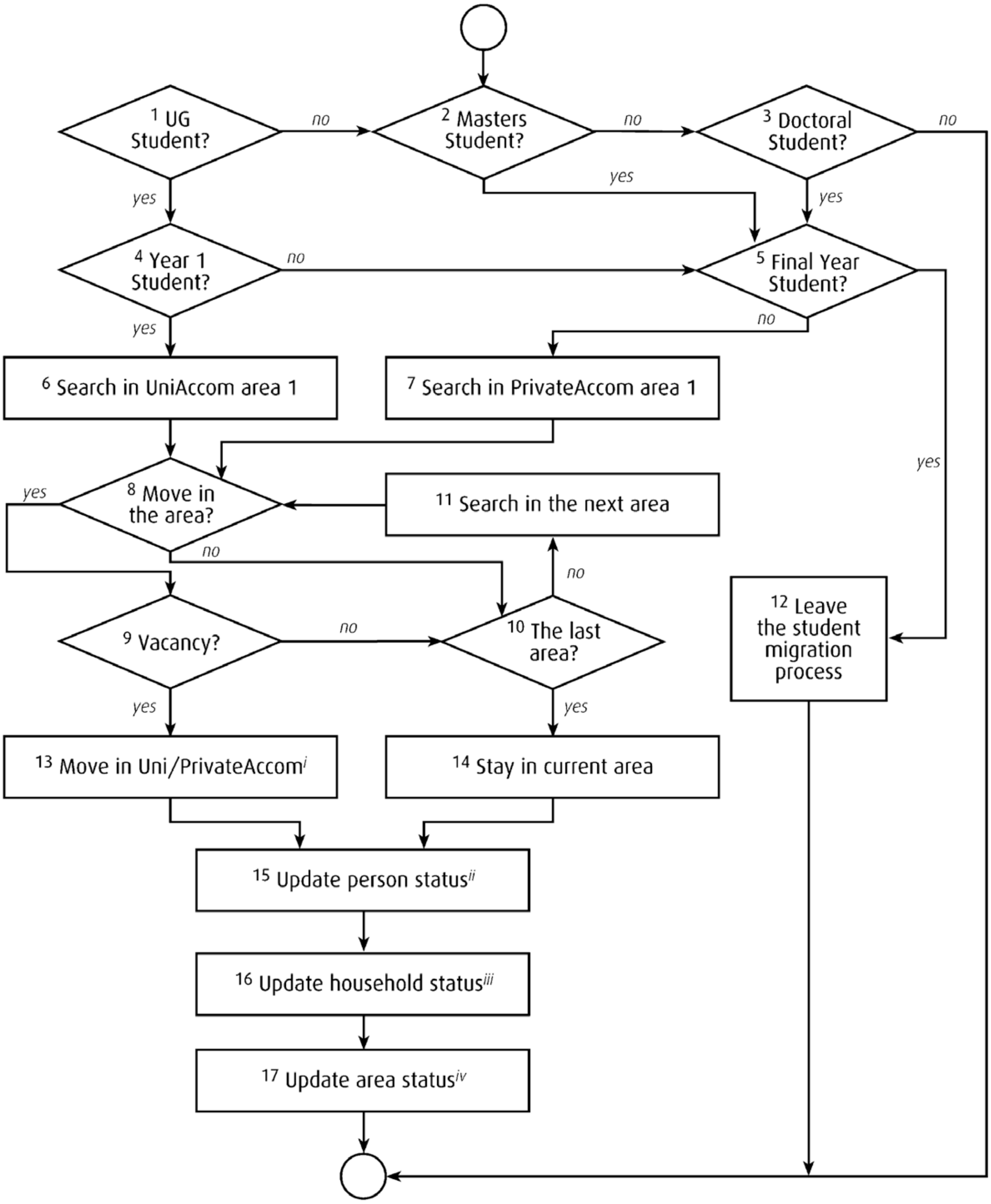
Depending on the type of the students, their rules vary slightly. For example, a year two undergraduate student can stay in the area for two years. He/she then can have the chance to continue studying towards a master's degree for one more year or leave. While a masters students can stay in the area for one year and then continue with a doctoral study for three more years or leave.

The typical interaction between the agents in this model would be finding the fellow students in order to move to the area they are and the interaction with the environment would be checking if there is a vacancy in that area. If the searching result is negative, the agent/student moves on to the next area. At the end of the search, the agent without success then stays in the area for one more year.

The typical student migration process that a student agent experiences each year during their stay in Leeds has been illustrated in Fig. 6. The student agents will carry out the similar process (depending on the updates of the student status) if they are going to stay in Leeds for at least one more year. At the end of their study, they will be marked as final year students and leave the student migration process and join the general migration process (Fig. 6).

4.3. Comparison of the simulation results using pure MSM and hybrid model

Using only the pure MSM, the model does not differentiate the student migrants from other migrants and they join the general migration process as described in previous sections. By applying different rules to individual "student agents", the hybrid model presents a better reflection of the observed student population in wards of Leeds. The maps below show the comparison between the currently observed student population (Fig. 7) and the simulation results two years later using a pure MSM approach (Fig. 8) and a hybrid MSM (microsimulation model) and ABM (agent-based model) approach (Fig. 9). In the maps, each red dot represents



i. Year 1 students in University accommodation and other students in private accommodation;
ii. type of students, year(s) stayed in Leeds and household ID etc.
iii. household size, household reference person, number of various household members etc.
iv. number of people in the area, number of household in the area, vacancies in area etc.

Fig. 6. Student migration process.

100 students. The decline of the density of the student population in an area is indicated by the intensity of the shading in the map.

From the maps we can observe that pure MSM has projected a relatively even scattering of students around the whole city. When we look at the levels of density of student population in that area (shown in the legend), there is no big gap between the highest and the lowest level of density, unlike the observed and the hybrid model results. Pure MSM simulation also results in some over-representation of student population in suburban areas. On the other hand, the hybrid model projection suggests that students tend to gather around the universities in the city centre and they normally do not live in suburban areas. Most of them will leave the area once they complete their studies. Comparing the three maps, we can see that the hybrid model results provide a much better reflection of the observed population in reality than the pure MSM simulation results.

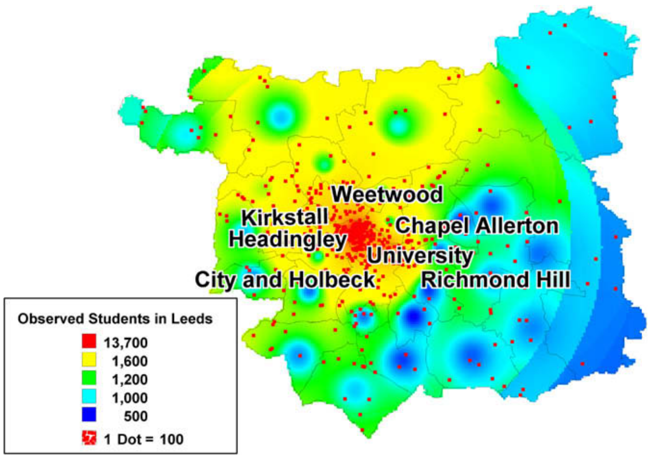


Fig. 7. Leeds students: observed. Source: 2001 Census: standard area statistics (England and Wales); Ward Boundaries. Crown copyright 2003.

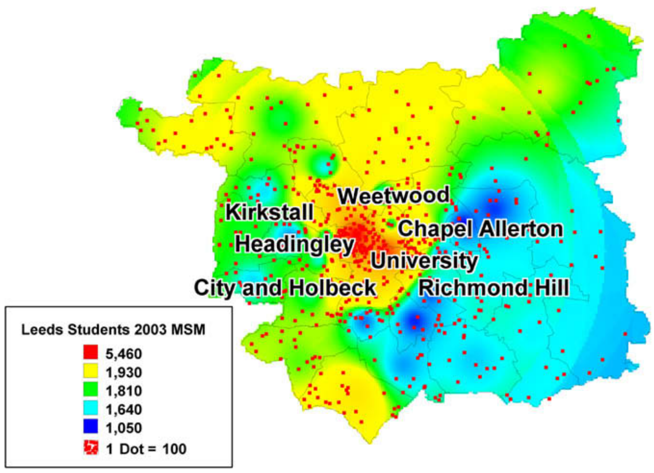


Fig. 8. Leeds students: pure MSM result. Source: 2001 Census: standard area statistics (England and Wales); Ward Boundaries. Crown copyright 2003.

A formal comparison of model results from the MSM and ABM is shown in Table 3. The level of similarity between four groups – the total Leeds population and the student population, as observed in the 2001 census data, the distribution of students from the MSM and the ABM. The comparison is presented as an index of dissimilarity (IoD). IoD has been used as a means for comparing the distribution of different social and ethnic groups in cities (Rees & Birkin, 1983) and reflects the average variation in concentration between areas. If all the students were in Headingley ward and the rest of the population evenly distributed elsewhere then the IoD would be 100. If the students are spread uniformly across the city then the IoD will be zero. The mathematical expression of this relationship is

$$IoD = 100\sum_i \left[\frac{\boldsymbol{X}_i}{\sum_i \boldsymbol{X}_i} - \frac{\boldsymbol{Y}_i}{\sum_i \boldsymbol{Y}_i}\right] \Big/ 2$$

where $\boldsymbol{X}_i$ is a count for the small area population of ward i, and $\boldsymbol{Y}_i$ is the count for students. Other distributions for $\boldsymbol{X}$ and $\boldsymbol{Y}$ are substituted as appropriate.

The results in Table 3 show that the student population is indeed highly concentrated against the general population of Leeds (IoD = 62). This is reflected well in the ABM (IoD = 62) and much less well in the MSM (IoD = 13). The same pattern is reflected if we compare the student population with the models: in this case we find a high level of dissimilarity to the MSM model (IoD = 50) and a much lower level of dissimilarity to the ABM (IoD = 12). Last, we find that the ABM distributions are rather different to the MSM (IoD = 51).

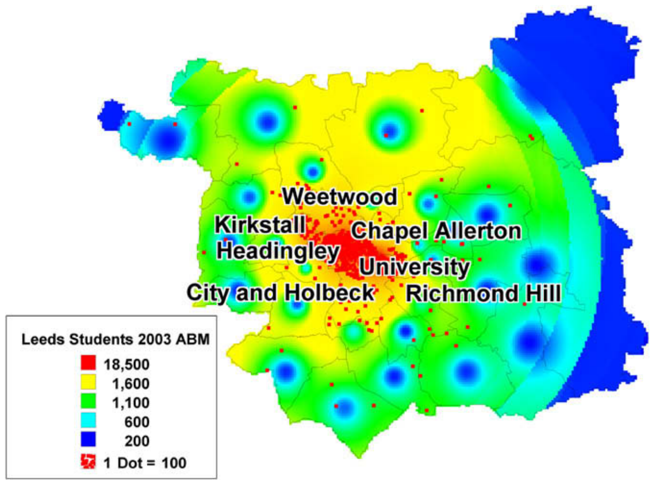


Fig. 9. Leeds students: hybrid model result. Source: 2001 Census: standard area statistics (England and Wales); Ward Boundaries. Crown copyright 2003.

Table 3
Index of dissimilarity for student populations.

	Leeds pop	Student pop	MSM	ABM
Leeds pop	–	61.6	12.6	62.3
Student pop	–	–	49.7	11.7
MSM	–	–	–	50.7

5. Potential applications of the model

In its current form, the Moses model can be used to investigate a number of policy scenarios for the UK or its constituent cities and regions. For example, we could look at Government proposals for the location of additional housing on brownfield sites, to understand whether opportunities are sufficiently widespread to meet latent demand for new housing, and in relation to alternative scenarios such as relaxation of greenbelt restrictions in peri-urban locations. On the other hand, we might explore the effect of continuing high immigration from new EU member states, perhaps including qualitative evidence or a growing body of empirical data about return migration.

Amongst many potential policy users, planners in the Leeds city council division of social services are particularly concerned with the impacts of an ageing population on provision of services within the city. A selection of questions which Moses seeks to answer is set out in Table 4. Some of the questions are reasonably straightforward, and could be addressed through standard analytical techniques. For example, the question of whether limiting long-term illness is correlated with morbidity depends on the acquisition of appropriate data on morbidity from one or more of the city health trusts, from which point it is a straightforward piece of analysis. Other questions play directly to the strengths of the modelling approach which we have proposed. For example, co-dependency of the elderly population within households and ethnic minority populations can both be projected directly from the IBM. This allows

Table 4
Elements of a model-based analysis of social care.

Target population	Prevalence/incidence data	Examples of data sources	Examples of some questions to consider
Older people (OP)	Limiting long-term illness	Census 2001 DoH – health survey for England 2004	Can self-reported limiting long-term illness be validated against other data sources? Does the profile for LLLI correlate with other data, e.g. morbidity data?
	Physical disability	Health survey for England 2000 (DOH)	Which wards/localities have the most OP with LLTI?
		Census 2001 (theme table 06)	Is there a correlation with the distribution of services, e.g. home care, equipment and adaptations, hospital admission data.
	Limiting long-term illness (LLTI)	Ageing: scientific aspects (house of Lords)	Can a small area analysis validate the city-wide picture?
	Sensory impairment	–	–
	Cardiovascular disease (e.g. stroke/heart attack)	Public health observatories	Is there any correlation between the current distribution of services and the proximity to District centres (e.g. for shopping) and health care facilities? Is there any significance in the age, geographical distribution) of people who attend day centres for older people?
	Ethnicity	Policy Research Institute on ageing and ethnicity	What is the likely impact for service delivery arising from the projections for growth in ethnic minority population, (both in numbers and the age profile of the projected increases in numbers)? What is the current and future likely level of co-dependency among older couples

flexible aggregation of outputs to allow the construction of indices of need or service uptake (physical disability, sensory impairment, etc) by combining estimates from survey data – health survey for England or BHPS, for example, or perhaps using data about existing social services users in the Leeds area. This is analogous to the process by which migrants are estimated within the current model. A third class of results necessitates the integration of behavioural models of service utilisation, which will only be captured through incorporation of third party data regarding both the provision and uptake of services. For example, where are the existing day centres and from where do they draw their users? Once such models have been estimated, however, the evaluation of alternative delivery options or future requirements becomes much more feasible.

One of the simpler examples is shown in Fig. 10. Here we have applied small area rates of limiting long-term illness disaggregated by age and sex to an evolving population over a 25 year time horizon. The results show a growth in the demand for services, which is most obvious in the northerly areas of the city. Some of these areas, e.g. Wetherby, are also the most sparsely populated and this could create special challenges for service delivery and access. An analysis of this data implies a slight shift of resources from the central areas to the more outlying areas. Of course, this scenario incorporates a rather crude assumption of no change in illness rates over

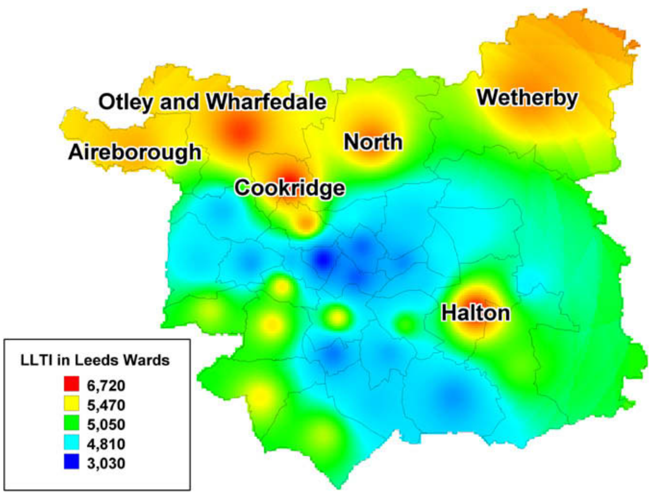


Fig. 10. Limiting long-term illness in Leeds, 2031. Source: 2001 Census: standard area statistics (England and Wales); Ward Boundaries. Crown copyright 2003.

quite a long period of time. We believe that this kind of assumption is appropriate as a starting point, from which the modelling can then move to explore further scenarios: what happens if there a 10% improvement at each age group? Can we extrapolate trends over the last 25 years? What assumptions about illness and service utilisation do we need to make so that provision is sustainable with current or expected resources?

6. Discussion

In this paper we have introduced a dynamic microsimulation model (MSM) to represent demographic change within an urban area. The model exploits the well-known strength of MSM for list processing based on well-defined rules. Ageing, fertility and mortality can all be simulated easily by this means. However once we start to introduce concepts of movement and interaction, then the standard MSM process begins to malfunction. We have considered two means for the extension of MSM in relation to the problem of household and student migration.

Spatial interaction modelling is a means to represent population flows at the meso-scale, thus the movement of people and households between (small) areas in a city, in which groups are typically disaggregated according to a small number of key attributes. This style of modelling is an ideal complement to MSM, and in a sense can be thought of as representing the influence of collective or 'market' processes on individual behaviour. For example, in the problem of residential choice then it is a straightforward matter to create a simulation in which houses are vacated on the basis of individual migration propensities, and then filled as the result of an individual search mechanism. But what happens when all the vacancies within an area are quickly used up? In real environments, what tends to happen is that price will go up to regulate supply with demand. This is essentially a meso-level process, and in our version of the SIM (spatial interaction model) is articulated through the balancing factors which operate as shadow prices on housing availability.

Agent-based modelling (ABM) is also complementary to MSM, although in a different way. It is generally recognised that agents are an effective way to represent individual entities which move around and interact with one another and with their environment. On the other hand, agents are usually represented using relatively sparse attribute sets and simple rules. The generation of complex or emergent systemic behaviour from simple rules is a touchstone for ABM. Thus mobile agents interacting with an environment (ABM) look like an ideal partner to static, self-contained individuals which are nevertheless richly specified (MSM). We have demonstrated one potential use for ABM within an individual-based model (IBM) of demographic change with respect to the management of student populations. At the moment, the agent-based model has been applied only to the city of Leeds. It has been shown that the agent model generates significantly more reliable estimates of local area population distributions than the equivalent microsimulation model. The ABM could be employed as part of the national Moses modelling framework, although this would require a national analysis of student accommodation in the manner described in Section 4.2.

There is obvious potential for the extension of both SIM and ABM within the broader project of simulating the development and behaviour of an urban population. The SIM is likely to be appropriate for the representation of behaviours such as journey-to-work, retail and leisure activities and utilisation of health services. Therefore we see the coupling to activity-based SIM as an ideal means to extend the attributes within demographic MSM, and to link to policy drivers such as the closure of a hospital, or development of a new retail mall. The incorporation of agent-style dynamics also has more general value in our opinion. Another simple example might be the process of marriage (or cohabitation) in which the search for a partner would lend it naturally to the representation of mobile and interacting individuals/agents. A more complex example, perhaps, might be associated with a process like household dissolution, where the movement of (elderly) residents into hospitals or other care environments might be mediated through local social and institutional networks. Just having a neighbour or relative to help with the shopping and other mundane domestic activities is crucial in many real situations. Thus the transition pathways could look very different for individuals in a neighbourhood with cohesive relationships and strong local institutions than for someone who is socially or geographically isolated.

A different kind of agent-based interaction could be between individuals and their environment. Schelling's seminal analysis of segregation posits a relationship between individual decisions to move and the (ethnic) composition of neighbourhoods (Schelling, 1971). This process could easily be represented within a migration model, and perhaps extended to embrace characteristics such as social class, income or lifestage. In other words, that individuals exhibit a preference to live close to other people, not just of the same ethnic group, but of a similar age and a shared social background. Therefore an agent-based mechanism might be used to maintain geodemographic discrimination within urban population projections.

At the same time, we should recognise that the importance of MSM is hardly exhausted by the representation of natural change processes within a demographic model. For example, MSM will be an ideal basis for the determination of health status, need for services, and lifestyle characteristics such as car ownership, in much the same way as we used the MSM in the generation of migrants within the current application. So we believe that the framework of an individual-based model which combines MSM, SIM and ABM is a potent and necessary cocktail for the representation of urban social systems and associated policy modelling (Fig. 11).

As a concluding remark, we would wish to assert one caveat in relation to the terminology of this paper and the concluding discussion. We have characterised a set of approaches – MSM, SIM, ABM – and then argued that it makes sense to combine these approaches within a single modelling framework. A reasonable objection to this argument would be that actually the three approaches aren't that different in the first place. For example, the equivalence between spatial interaction models and discrete choice models, say in relation to housing market search, is well-known (Williams, 1977), and in principle our SIM could be represented equally well

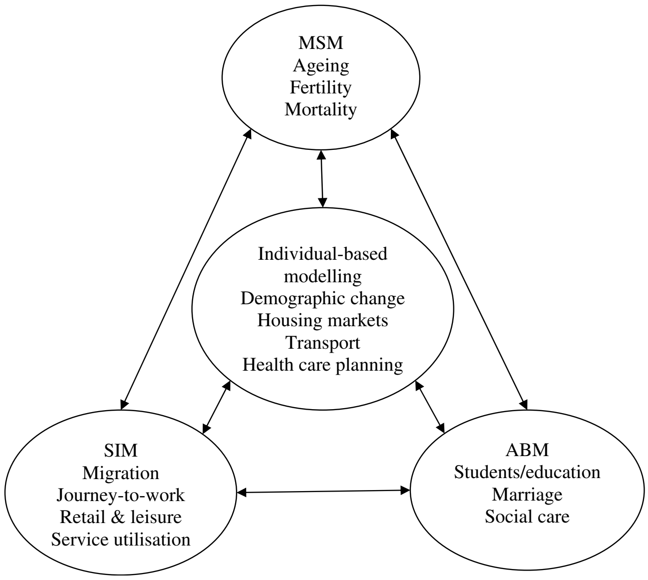


Fig. 11. Framework for an individual-based model (IBM).

by a discrete choice process within the MSM. If the fundamental characteristic of ABM is that 'the agent society and its spatial environment are coupled' (Epstein & Axtell, 1996) then the MSM, with either a SIM or discrete choice model, already exhibits these characteristics. We also believe that different schools are often separated by a disparity in purpose as much as method. Thus while MSM is heavily focused towards applied problems, ABM tends to favour simplicity of rules, with an emphasis on understanding and explanation. Again, we suggest that this distinction is largely artificial. We feel that it is important for advocates of microsimulation to consider richer rule sets and behaviours within their models, and for agent enthusiasts to pay more attention to the calibration and realism of their models. The achievement of this kind of balance is an alternative interpretation of our adoption of a plurality of modelling styles within an individual-based modelling framework.

Acknowledgements

This research has been funded by ESRC through the National Centre for e-Social Science (Award Reference RES-149-25-0034).

Census output is Crown copyright and is reproduced with the permission of the Controller of HMSO and the Queen's Printer for Scotland.

British Household Panel Survey is copyright to the Institute for Social and Economic Research, UK Data Archive, University of Essex. The UK Data Archive bears no responsibility for the processing and analysis of these data.

We are grateful for the comments of three anonymous referees on an earlier draft of this paper.

Appendix A. Spatial interaction model for the distribution of migrants

We assume that migration processes are driven by

(1) The distance between the origin zone and the destination zone (d_{ij}).
(2) The demographic characteristics of the migrant households (age and household size) ($\boldsymbol{O}_i^k$).

(3) The housing characteristics of the destination area (detached/semi-detached/terraced/flats/other) ($\boldsymbol{H}_j^m$).

Migration distances are regulated through a distance deterrence parameter, β, and a set of housing preferences for each demographic group, λ^{km}. This gives us a model with the following structure:

$$\boldsymbol{T}_{ij}^{km} = \boldsymbol{A}_i^k \boldsymbol{O}_i^k \boldsymbol{B}_j^m \boldsymbol{H}_j^m \, \mathrm{e}^{-\beta d_{ij} + \lambda^{km}}$$

$$\boldsymbol{A}_i^k = \frac{1}{\sum_{jm} \boldsymbol{B}_j^m \boldsymbol{H}_j^m \, \mathrm{e}^{-\beta d_{ij} + \lambda^{km}}}$$

$$\boldsymbol{B}_j^m = \frac{1}{\sum_{ik} \boldsymbol{A}_i^k \boldsymbol{O}_i^k \, \mathrm{e}^{-\beta d_{ij} + \lambda^{km}}}$$

The input vectors $\boldsymbol{O}_i^k$ and $\boldsymbol{H}_j^m$ can be obtained from the migra-production model:

$$\boldsymbol{O}_i^k = \sum_{mw} \boldsymbol{M}_i^{kmw}$$

$$\boldsymbol{H}_j^m = \sum_{kw} \boldsymbol{M}_j^{kmw}$$

where $\boldsymbol{M}_i^{kmw}$ is the number of household migrants by age of head and household size (k), housing type m, and other characteristics w.

The calibration of the model is undertaken in relation to the distance patterns and household preferences of migrants, so that: β calibration

$$\sum_{ijkm} \boldsymbol{T}_{ij}^{km} d_{ij} = \sum_{ij} \boldsymbol{T}_{ij}^{obs} d_{ij}$$

where $\boldsymbol{T}_{ij}^{obs}$ is the observed flow of household migrants from ward i to ward j (from the 2001 census special migration statistics). λ calibration

$$\sum_{ij} \boldsymbol{T}_{ij}^{km} = \sum_{iw} \boldsymbol{M}_i^{kmw}$$

Given the derivation of the migrant households, this effectively means that the preference for housing by demographic group is calibrated against the British household panel survey, adjusted to local market conditions.

A.1. Calibration method

Maximum likelihood estimates for the parameters are obtained from an iterative adjustment procedure which is slightly messy in view of the need to also incorporate recursive estimates of the model balancing factors. The procedure therefore looks like this

1. Establish random starting values for β and λ.
2. Set initial values for the destination balancing factors, $\boldsymbol{B}_j^m = 1$ for all j, m.
3. Calculate values for the origin balancing factors, $\boldsymbol{A}_i^k$.
4. Calculate new estimates for $\boldsymbol{B}_j^m$.
5. Test for convergence of the balancing factors. If convergence achieved, continue to step 6. Otherwise return to step 3.
6. Calculate new values of beta such that β (new) = β (old) $\times$ predicted distance/observed distance.
7. Calculate new values of lambda such that λ (new) = λ (old) - observed housing preference/predicted housing preference.
8. Test for convergence in the parameter estimates. If convergence achieved then stop, otherwise return to step 2.

References

Ballas, D., Clarke, G., Dorling, D., Eyre, H., Thomas, B., & Rossiter, D. (2005). SimBritain: A spatial microsimulation approach to population dynamics. *Population, Space and Place, 11*(1), 13–34.

Barrett, S., Eubank, S., & Smith, J. (2005). If smallpox strikes Portland. *Scientific American, 292*(3), 54–62.

Baryla, E. A., & Dotterweich, D. (2001). Student migration: Do significant factors vary by region? *Education Economics, 9*(3), 269–280.

Billari, F., Ongaro, F., & Prskawetz, A. (2002). Introduction: Agent-based computational demography. In F. Billari & A. Prskawetz (Eds.), *Agent-based computational demography: Using simulation to improve our understanding of demographic behaviour* (pp. 1–18). London/Heidelberg: Springer/Physica.

Birkin, M., & Wu, B. (2008). Dynamic social simulation models enabled by e-research. In *Proceedings of the fourth international conference on e-social science*. Manchester: National Centre for E-Social Science. <http://www.ncess.ac.uk/conference/programme>. Accessed 9.09.2008.

Birkin, M., & Clarke, M. (1987). Comprehensive models & efficient accounting frameworks for urban & regional systems. In D. Griffith & R. Haining (Eds.), *Transformations through space and time* (pp. 169–195). The Hague: Martinus Nijhoff.

Birkin, M., Clarke, M., Chen, H., Dew, P., Keen, J., Rees, P., et al. (2005). MoSeS: Modelling and simulation for e-social science. In *Proceedings of the first international conference on e-social science*. Manchester: National Centre for E-Social Science. <http://www.allhands.org.uk/2005/proceedings/papers/341.pdf>. Accessed 18.04.2008.

Birkin, M., Turner, A., & Wu, B. (2006). A synthetic demographic model of the UK population: Methods, progress and problems. In *Proceedings of the second international conference on e-social science*. Manchester: National Centre for E-Social Science. <http://www.ncess.ac.uk/events/conference/2006/papers>. Accessed 9.09.2008.

Brown, L., & Harding, A. (2002). Social modelling and public policy: Application of microsimulation modelling in Australia. *Journal of Artificial Societies & Social Simulation, 5*(4). <http://jasss.soc.surrey.ac.uk/5/4/6.html>. Accessed 18.04.2008.

Caldwell, S., Clarke, G., & Keister, L. (1998). Modelling regional changes in US household income and wealth: a research agenda. *Environment and Planning C: Government and Policy, 16*, 707–722.

Conte, R., & Gilbert, N. (1995). Introduction. In N. Gilbert & R. Conte (Eds.), *Artificial societies: The computer simulation of social life*. London: UCL Press.

Epstein, J., & Axtell, R. (1996). *Growing artificial societies: Social science from the bottom up*. Washington, DC/London: Brookings Institution Press/MIT Press.

Espindola, A. L., Silveira, J. J., & Penna, T. J. P. (2006). A Harris–Todaro agent-based model to rural–urban migration. *Brazilian Journal of Physics, 36*(3A), 603–609.

Eubank, S., Guclu, H., Anil Kumar, V., Marathe, M., Srinivasan, A., Toroczkai, Z., et al. (2004). Modelling disease outbreaks in realistic urban social networks. *Nature, 429*(6988), 180–184.

Ferguson, N. M., Cummings, A. T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., et al. (2005). Strategies for containing an emerging influenza pandemic in Southeast Asia. *Nature, 437*(7056), 209–214.

Foster, I. (2003). The grid: Computing without bounds. *Scientific American, 288*(4), 78–85.

Fotheringham, A. S., Rees, P., Champion, T., Kalogirou, S., & Tremayne, A. R. (2004). The development of a migration model for England and Wales: Overview and modelling out-migration. *Environment & Planning A, 36*(9), 1633–1672.

Hamnett, C. (2003). Contemporary human geography: Fiddling while Rome burns. *Geoforum, 34*, 1–3.

Harding, A. (2007). Challenges and opportunities of dynamic microsimulation modeling. In *Plenary paper presented to the 1st General Conference of the International Microsimulation Association*, Vienna, 21 August 2007. Available at www.euro.centre.org/ima2007/programme/papers/Challenges_Harding_Paper.pdf. Accessed November 2008.

Hill, T., & Lewicki, P. (2006). *Statistics: Methods and applications. Chaid Analysis*. Tulsa: Statsoft Inc.

Johnston, R. J., Hepple, L., Hoare, A., Jones, K., & Plummer, P. (2003). Contemporary fiddling in human geography while Rome burns: has quantitative analysis been largely abandoned – And should it be? *Geoforum, 34*, 157–161.

Koser, K., & Salt, J. (1997). The geography of highly skilled migration. *International Journal of Population Geography, 3*(4), 285–303.

Loibl, W., & Toetzer, T. (2003). Modeling growth and densification processes in suburban regions-simulation of landscape transition with spatial agents. *Environmental Modelling & Software, 18*(6), 553–563.

Magne, L., Rabut, S., & Gabard, J. (2000). Towards a hybrid macro-micro traffic flow simulation model. In INFORMS Salt Lake City string 2000 conference.

Makowsky, M., Tavares, J., Makany, T., & Meier, P. (2006). An agent-based model of crisis-driven migration. In *Proceedings of the complex systems summer school 2006*. Santa Fe, NM: Santa Fe Institute.

Mozumder, P., & Marathe, A. (2004). Implications of an integrated market for tradable renewable energy contracts. *Ecological Economics, 49*(3), 259–272.

Murphy, B. (2000). SPSD/M. In A. Gupta & V. Kapur (Eds.), *Microsimulation in Government Policy and Forecasting* (pp. 587–592). Amsterdam: North-Holland.

Nakaya, T., Fotheringham, A. S., Clarke, G., & Ballas, D. (2007). Retail modelling combining meso & micro approaches. *Journal of Geographical Systems, 9*, 345–369.

Office for National Statistics (2007). 2004-based sub-national population projections for England (revised), National Statistics Centre for Demography. <http://www.statistics.gov.uk/statbase/product.asp?vlnk=997>. Accessed 17.04.2008.

O'Hare, J. (2000). TRIM3. In A. Gupta & V. Kapur (Eds.), *Microsimulation in Government Policy and Forecasting* (pp. 581–586). Amsterdam: North-Holland.

Orcutt, G. (1957). A new type of socio-economic system. *Review of Economics & Statistics, 58*, 773–797.

Ott, J., & Rao, D. (1985). A chi-square test to distinguish allelic association from other causes of phenotypic association between two loci. *Genetic Epidemiology, 2*(1), 79–84.

Rees, P. (1994). Estimating and projecting the populations of urban communities. *Environment and Planning A, 26*(11), 1671–1697.

Rees, P., & Birkin, M. (1983). Census-based information systems for ethnic groups: A study of Leeds and Bradford. *Environment and Planning A, 16*, 1551–1571.

Rephann, T., & Holm, E. (2004). Economic–demographic effects of immigration: Results from a dynamic spatial microsimulation model. *International Regional Science Review, 27*, 379–410.

Russell, S., & Norvig, P. (1995). *Artificial intelligence: A modern approach*. Englewood Cliffs, NJ: Prentice Hall.

Schelling, T. (1971). Dynamic models of segregation. *Journal of Mathematical Sociology, 1*, 143–186.

Taylor, M., Brice, J., Buck, N., & Prentice-Lane, E. (2005). *British Household Panel Survey User Manual* (Vol. B). Colchester: University of Essex.

Transims Travelogue (1996). Transims travelogue (LAUR-96-1430). <http://transims.tsasa.lanl.gov/PDF_Files/Travel4.pdf>. Accessed 18.04.2008.

University of Essex (2005) [SN: 5151 computer file]. *Institute for Social and Economic Research, British Household Panel Survey; Waves 1–13, 1991–2004*. Colchester, Essex: UK Data Archive.

Unwin, D. J. (2005). Fiddling on a different planet? *Geoforum, 36*, 681–684.

Van Imhoff, E., & Post, W. (1998). Microsimulation methods for population projection. *Population: An English Selection, 10*, 97–138.

Williams, H. C. W. L. (1977). On the formation of travel demand models and economic evaluation measures of user benefit. *Environment & Planning A, 9*, 285–344.

Williamson, P., Birkin, M., & Rees, P. (1998). The estimation of population microdata using data from small area statistics and samples of anonymised records. *Environment and Planning A, 30*, 785–816.

Wilson, A. (1967). A statistical theory of spatial trip distribution models. *Transportation Research, 1*, 253–269.

Wilson, A. G. (1974). *Models for urban & regional planning*. Chichester: Wiley.

Wilson, A. G., & Pownall, C. E. (1976). A new representation of the urban system for modelling and for the study of micro-level interdependence. *Area, 8*, 256–264.