Inter-regional Migration Modelling: A Review and Assessment

John Stillwell
School of Geography, University of Leeds, Leeds LS2 9JT, UK
Email: j.c.h.stillwell@leeds.ac.uk

Paper prepared for the 45th Congress of the European Regional Science Association, Vrije Universiteit Amsterdam, The Netherlands, 23-27 August 2005.

Abstract

Population migration involves the relocation of individuals, households or moving groups between geographical locations. Aggregate spatial patterns of movement reflect complex combinations of motivations and there is a huge literature on modelling different types of migration at various spatial scales. This paper, which originates from a study for EUROSTAT to find a generally applicable migration model, endeavours to present a succinct review of the state of the art based around the distinction between explanatory and projection models. Whilst the review inevitably lacks comprehensiveness, it emphasises the distinction between micro and macro approaches, between influences and determinants, and between mathematical and statistical models. It highlights a recent two-stage model developed for use in a policy context in the UK and contrasts this approach with models developed in the context of multi-state demography and used for migration projection in the European Union.

Key words: migration, modelling, deterministic, projection, spatial interaction

1 Introduction

Population is a complex phenomenon not only because of the variety of spatial patterns of movement that occur but because of the myriad of motivations that influence the size and composition of flows between any two discrete areas and the imprecision of the data that are collected. Whilst it is possible to utilise census, survey and registration data to provide some insights into directional patterns and temporal change, there is a conspicuous lack of data directly relating migration flows to the motivational factors that underpin the movements that take place and this has a limiting effect on explanatory analysis. Population censuses tend not to ask people why they moved; registration data normally lack a motivational dimension; and surveys that do ask questions of this type are usually localised and rarely comprehensive. As a consequence, those who seek to explain migration flows are confronted with the task of trying to identify the determinants that are relevant and then to establish which of those explanatory variables are the most important. This is the challenge that has been taken up by many researchers and which has resulted in a plethora of studies involving different modelling and calibration approaches, different measures of the independent migration variable and a wide range of dependent variables.

In contrast to the range of studies that have sought to *explain* internal migration using modelling methods, another strand of pure and applied modelling has attempted to *project* internal migration based on current or historical trends and frequently in the context of the estimation and projection of sub-national populations. This tradition is an important part of regional demography and has its roots in the single region and multi-regional population projection approaches developed in the 1960s and 1970s.

Whilst academic research on migration modelling has embraced both explanatory and projection methods, the application of migration models by national government departments or agencies has in the past tended to focus on the generation of internal migration projections from a migration sub-model independent of the main demographic projection system. The approach adopted in England during the 1980s is a prime example of this with the migration projections prepared by the Department of Environment being fed into the Office of Population Censuses and Surveys' population model (Armitage, 1986). In contrast, some countries have attempted to link explanatory factors into their population projection systems since the 1980s. A good example is the Demographic Regional Economic Model (DREM) developed by the Central Statistics Bureau in Norway which attempts to take into account labour market factors determining migration flows between regions (Stambøl, 1991).

These two approaches provide a broad framework for a review of migration models where emphasis is given to some of the most recent studies in this field. However, it is important to recognise some fundamental theoretical underpinnings to migration modelling (Section 2) and to identify the range of factors that influence migration selectivity and which actually determine migration flows (Section 3). Given that explanatory migration models have been developed in a very wide range of spatial and temporal contexts and it is impossible to present a comprehensive review, Section 4 focuses on the tradition of spatial interaction modelling that has its roots in the application of Newtonian gravitational principles in social science. Distinction is drawn between mathematical and statistical calibration methods of different forms of spatial interaction model and a more detailed summary is provided of a recent state-of-the-art two-stage migration model based on spatial interaction principles and calibrated using statistical regression. Subsequently, in Section 5, in the context of multi-state demographic models, attention is paid in particular to migration models developed and used for the projection of internal migration in countries of the European Union (EU). Some conclusions are drawn in the last section to provide some guidelines for future modelling.

2 Theoretical underpinnings

A key distinction in migration modelling is that between micro and macro approaches (Stillwell and Congdon, 1991), a dichotomy which has parallels in economics and in psychology, for example. *Micro theory* relates to the individual migrating unit (person, group or household) and to the processes underlying the decision of the potential migrant to remain in the current location or to move somewhere else. It involves identification of those factors that influence this decision-making process: in the first instance, whether to stay or to move. Thereafter, it also takes into consideration the subsequent stage in the individual decision-making which involves choice between the alternative destinations available, once the decision to move has been taken. Since the choices at both stages are between discrete options (go or stay; go to i or j), the approach to migration modelling at the micro level is often known as the discrete choice approach (Maier and Weiss, 1991) and has its roots in the axiom of utility maximization since it is peoples' expectations about improving their own prospects in various locations that are at the heart of the decision-making process (Rothenberg, 1977).

The factors bearing on these decisions include both the characteristics of individual persons (such as age, marital status, household status) or wider family units (such as family size and structure) and the wider characteristics of the potential destinations (such as regional relativities of unemployment, wages or house prices). The relationship between migration behaviour and the changes that individuals experience as a consequence of progress through their life courses have been examined by various researchers since Thomas (1938), many in the context of intra-urban residential mobility and involving social psychologists such as Rossi (1955). Since utility is stochastic, a micro model formulation is likely to mean that the probability that an individual will choose destination region i is determined by an expression that compares the attributes of region i *vis a vis* those of the other possible destinations. The model which is calibrated empirically is a multi-nomial logit model. However, some potential destinations may be evaluated similarly because of preferences for certain types of area and this may lead to correlation in the random component of utility functions, leading to a contradiction of the assumptions of the model. Consequently, a number of studies, including Liaw and Ledent (1987), Hughes and McCormick (1989) and Van Wissen and Rima (1988) subdivide the decision-making process conceptually into two parts and consider that people's evaluations of alternative destinations are correlated. Nested logit models or multinomial probit models are used in these studies.

In contrast to micro theory, *macro theory* relates to aggregate migration flows and is more appropriate for setting migration in its labour or housing market context in order to deal with questions such as whether people migrate into areas where jobs are available or where prices are lower rather than the behavioural aspects surrounding the migration decision itself. Macro approaches are therefore concerned with investigating relationships between migration and objectively determined macro variables such as population sizes, unemployment rates or environmental conditions. One theoretical perspective is that embraced by classical models of regional self-balance which suggest that migration is the equilibrating mechanism through which regions achieve adjustment, as, for example, when people move from regions with high unemployment to regions where unemployment is low. Myrdal (1957), however, argued that the selective nature of migration enhances regional differentials and therefore migration is disequilibrating rather than equilibrating. In contrast to the debate over the role of inter-regional migration, the seminal contribution to migration theory offered by Lee (1966; 1969) in identifying push and pull factors influencing aggregate flows of internal migrants between regions has been of fundamental importance to those constructing macro models.

The distinction between micro and macro approaches thus provides a broad classification system for migration modelling and Cadwallader (1989) has articulated a valuable conceptual framework for understanding the relationship between the two approaches by suggesting that there are four sets of relationships:

- between aggregate migration and regional attributes that has been traditionally investigated by macro models;
- between the regional variables defined objectively and the subjective perceptions of those indicators by individual migrants;
- the integration of those perceptions about places into aggregate utility functions; and
- their subsequent translation into aggregate migration flows.

Data availability has been a major constraint on micro-behavioural modelling. There are relatively few national or regional surveys of migration motivation that provide spatial data on individual person, family of household decisions with regard to migration. In contrast, attempts to model the macro relationships between migration and factors deemed to be influential are more commonplace because of the availability of aggregate data on migration from censuses and registers and of explanatory variables from various sources.

3 Selective influences and determinants of migration

Another important distinction is between those characteristics of individuals or households that are indicative of higher or lower propensities to migrate and those factors that actually determine whether a move takes place and which destination is selected. Age is a typical example of the former since it does not itself determine migration but people of different ages have very different migration propensities. Job opportunities exemplify the latter; the opportunity to work elsewhere may well be the driving factor behind the migration of many in the labour force ages but may only be one of a combination of determinants. We can examine the range of selective influences on migration and the determinants of migration separately but should acknowledge that they are not mutually exclusive. A comprehensive review of determinants across a spectrum of dimensions is found in Champion *et al.* (1998).

3.1 Selective influences

Demographic characteristics have a major influence on migration propensities. Age is a variable that changes for the individual in a regular and irreversible way over the life course, whilst sex is fixed at birth and persists. Migration intensities vary in a familiar way with age in most developed countries at different spatial scales as demonstrated by work by Rogers and Castro (1981). The shape of the age-specific migration rate schedule reflects a number of life course transitions (e.g. leaving home, getting married, having children, retirement, *et cetera*) whose sequence and associated housing needs and distance of movement has been carefully documented by Warnes (1992). Moreover, the relationship between migration rates and age has been modelled by as a function of five components associated with childhood, employment, retirement, old age and a constant (Rogers *et al.,* 1978; Rogers and Castro, 1981; Rogers and Willekens, 1986). The model of migration intensity at exact age *a* for any zone has the form:

$$m^a = b_1 \exp(-\alpha_1 a) + b_2 \exp\{-\alpha_2(a-\mu_2) - \exp(-\lambda_2(a-\mu_2))\} + b_3 \exp\{-\alpha_3(a-\mu_3) - \exp(-\lambda_3(a-\mu_3))\} + c \qquad (1)$$

where the profile of the schedule is defined by 7 parameters ($\alpha_1, \alpha_2, \mu_2, \lambda_2, \alpha_3, \mu_2, \lambda_3$) and the level of the schedule is determined by the remaining parameters (b_1, b_2, b_3, c). This model was operationalised in a general computer program called MODEL by Rogers and Planck (1984). It has been fairly widely used for smoothing erratic data and disaggregating from broad to narrow age bands and the methodology was used in the sub-national population projection model for England in the 1980s following a design by Bates and Bracken (1987).

Migration differentials between males and females are much less distinct than those between age groups but differences will be distinguishable when comparing migration intensities and patterns in certain contexts. Female rates may rise faster than males after age 16 to a slightly earlier peak than men and then decline at a rate slightly below men until retirement age. Thereafter, particularly in older old age, female rates may exceed those for males again. Differences in migration profiles are also evident between single, married, widowed and divorced groups (Devis, 1983).

Since Sjaastad's (1960) pioneering work on the human investment approach to migration analysis, our understanding of how age composition change influences migration has improved considerably. Plane (1992) explored the effects of demographic change on migration in the USA through an examination of migration rates of different age groups and cohorts over time, and the effect on total migration flows of the ageing of regional populations. Plane and Rogerson (1991) have borrowed Easterlin's (1980) relative cohort size hypothesis to explain migration levels. Baby boom generations, for example, experience more competitive conditions on entry into the labour market and hence fewer job opportunities tend to depress migration levels in comparison with smaller birth cohorts. Pandit (1997) has carried out a set of time series tests on the efficacy of the cohort size hypothesis *vis a vis* the business cycle hypothesis in the USA and tentative interpretations of the influence of birth cohort effects on net-migration for selected zones in the UK and Australia have been undertaken by Stillwell *et al.* (2001).

Whilst regular demographic influences have been shown to occur in migration intensities in different countries, fewer cross-national comparisons have been undertaken that have focused on the differences in migration propensities between various ethnic groups, social classes, those with different educational qualifications, those in different classifications of economic activity or those with different housing tenure characteristics. Various national studies have been carried out, including that by Owen and Green (1992), whilst Fielding (1992) has used data from the Longitudinal Study in the UK to demonstrate the extent to which inter-regional migration selects persons in higher-level occupations. Another key selective influence on migration is housing tenure. Hughes and McCormick (1981; 1985; 1987) and Boyle (1993), for example, have argued that the management of public housing is responsible for discouraging longer distance migration of tenants in Britain, whilst those in owner-occupied housing are known to be more likely to move over longer distances than those in council housing.

This short review of selective influences serves to highlight the important issue that surrounds modelling based on data sets of aggregate migration flows. It is very likely that these bundles of individuals will conceal streams of selective migrants influenced because of their demographic, social and housing circumstances, but also motivated to move for a host of different reasons, some of which are now considered.

3.2 Determinants

Gravity variables. Lee's classic study published in 1966 conceptualises migration as involving origins, destinations and the links between them. The characteristics of the origin may act as 'push' factors for potential out-migrants whilst the attributes of the destination reflect 'pull' factors that entice migrants to a particular destination. The separation of origins and destinations imposes a cost on migration and the term 'impedance' is often used to refer to the frictional effect of distance on migration. These factors were represented in the early formulations of the gravity model as gravity variables and were measured by the total populations of the origin and destination zones and the physical distance between them.

Considerable debate has centred on the measurement of distance since it may be argued that physical distance does not reflect either the social costs of moving or time costs that may not be proportional to distance. Moreover, measures of physical distance can vary from Euclidian distances between zone centroids to road mileage distances or network-weighted distances calculated on the basis of shortest surface so as to take account of the effect of estuaries. Areas that share a boundary tend to have more migration between them because this migration will include a proportion of short-distance moves from one side of the boundary to the other. Consequently, several studies have used contiguity variables that take the value of 1 for zones that share boundaries with each other, and 0 for other pairs of zones. Finally, it is apparent that areas situated in high population-density regions are likely to be less attractive as destinations to migrants, everything else being equal, because of increased spatial competition between destinations (Fotheringham, 1986). A destination accessibility or potential variable for zone i which measures the degree of spatial competition faced by a destination zone from nearby destination zones may be created and this competing destinations effect can be extended by incorporating a set of 'regional' variables that describe each zone and its neighbours (ODPM, 2002).

Economic variables. Longer distance migrants tend to have a higher probability of changing their place of work as well as their place of usual residence when they migrate. These migrants, together with their partners or families in some cases, are more likely to be influenced by relative regional economic prosperity. Young aspiring business executives are attracted to dynamic regions where economic growth is relatively buoyant and where company development appears to be successful. Consequently, variables that measure levels of prosperity, such as GDP per capita, or the number of new business registrations, together with those that identify how conditions are changing over time, are likely to be important influences, though these may be ‘picked up’ by variables that characterize the labour market.

Labour market variables. On many occasions, levels of prosperity are reflected in the conditions of the job market. Labour market factors are seen as potentially important both in prompting out-migration from an area as well as in influencing people's destination choices. They include measures such as levels of employment and changes in jobs as well as unemployment rates. The relationship between migration and unemployment remains unclear, depending, in part, on the state of the economy overall.

Housing market variables. Housing factors form a critical element underlying migration patterns, but they have complex interactions with migration and need especially careful treatment. On the one hand, some housing measures such as high house prices and low vacancy rates can reflect the strong economic performance of an area and indeed of neighbouring areas within commuting distance. On the other hand, these factors can directly influence the opportunities for in-migration, with high house prices acting as a deterrent and high vacancy rates as an attraction. The size, composition and quality of the housing stock can also influence both the level and the type of migration. Most obviously, the number of new houses constructed or the number of housing demolitions are likely to be very important determinants of migration. Housing tenure is also known to affect migration patterns, most notably the well-documented problems that people moving between local authority areas have in accessing council housing.

Environment variables. In the current era within advanced economies, environmental factors play a major role in people's residential moves, both in prompting exits from areas and in acting as 'pull' factors. The term is used here in its broadest sense, covering all the physical, economic, social and political aspects that affect both the everyday quality of life and the longer-term trends in life chances. This category can therefore be considered to include most of the factors mentioned under the other headings above, insofar as they bear upon the overall quality of an area but it also includes variables relating to derelict and vacant land, the proportion of new housing on brownfield land, population density, settlement size and level of urbanisation, crime and anti-social behaviour, climate and air quality, sports and leisure facilities as well as to 'bright lights' and good schools.

Policy variables. Public policy variables relevant to migration behaviour include not only direct interventions such as migration incentives and migration policy (such as the distribution of asylum seekers from reception centres to allocated dwelling spaces) but also indirect influences through the uneven effects of government grants, local taxes, defence spending, higher education expansion and the amount and location of land approved for house building. In general, it may be more satisfactory to estimate the role of public policy, past or anticipated, by reference to variables representing the aspects that public policy seeks to alter. For instance, the migration impact of a regional development initiative can be assessed by reference to, for instance, the number of extra jobs, while the impact of a policy that alters the availability of land for house-building can be studied via changes to the number of housing completions. Finally, in contrast to public policy, many organisations and private companies have staff recruitment and mobility policies (both internal and international) that result in inter-regional migration and for which it is very difficult to obtain any detailed information.

This review demonstrates the complexity that surrounds the phenomenon of inter-regional migration. This is accentuated by the fact that on many occasions, people tend to move and to choose their destinations on the basis of a unique combination of reasons. In concluding this section, it is also essential to recognise that different variables will be required to explain or project migration at different levels of aggregation. Thus, for example, trends over time in the overall migration intensities for one country are likely to be associated with fluctuating economic conditions, interest rates or mortgage rates, whilst regional out-migration rates for one period of time may be associated with regional prosperity indicators relating to each region and region indicators relative to the national average. On the other hand, the explanation of origin-destination migration flows between any two regions would need to embrace not only the characteristics of the two regions concerned and their distance apart, but also their characteristics *vis a vis* the characteristics of the regions that make up the rest of the system. These distinctions are taken up in the next section in which we consider alternative modelling approaches.

4 Migration models: explanatory approaches

Non-demographic models use additional non-demographic information for explaining and predicting migration patterns. There are many different forms that could be classified under this umbrella and which might be applied to explain regional differences in out-migration, the relative attractiveness of destination regions for in-migration, the net balance between out-migration and in-migration across a set of regions, or the spatial distribution of migrant flows between origins and destinations. The initial focus of the review that follows in this section is on the latter, and more specifically on the distinction between mathematical and statistical models that incorporate gravity variables.

4.1 Early gravity models

Ravenstein (1885) recognised the importance of the frictional effect of distance on migration in formulating his laws of migration back in the nineteenth century but migration models based on gravitational features were first developed in the 1940s (Zipf, 1946). These models incorporated terms measuring the masses of each origin and destination and of the distance between them and were calibrated statistically using log-linear regression techniques. Modifications were made to these early Newtonian gravity models by introducing parameters to weight the influence of the origin and destination factors and by experimenting with alternative distance functions.

4.2 Spatial interaction models: mathematical formulations

One of the shortcomings of these early approaches was the inability of the OLS regression formulation to predict interaction that was consistent with observed flows from each origin and to each destination. This was remedied by the introduction of so-called balancing factors (Wilson, 1967) to ensure internal consistency within the model and the derivation of the same model based on entropy-maximising techniques (Wilson, 1970). Wilson's family of four models of spatial interaction between any two zones i and j take the following general form:

M_{ij} = Scaling factor (or balancing factors)

* Origin out-migration (or attractiveness factor)

* Destination in-migration (or attractiveness factor)

* Distance function (with distance decay parameter) (2)

where a scaling factor is used when no observed out-migration or in-migration totals are known so that the sum of all the flows predicted in the origin-destination matrix is constrained to the total number of migrations observed in the system (the so-called unconstrained case).

Attractiveness factors are used as proxies for mass terms when out-migration or in-migration totals are unknown. When out-migration or in-migration totals are available, balancing factors replace the scaling factor to ensure that the row or column elements of the predicted matrix are consistent with the observations. The doubly constrained model of migration between regions i and j incorporated balancing factors for both origins and destinations (A_iB_j), mass terms (O_iD_j) and the distance function (d_{ij}) used is typically either a power function (as shown below) or an exponential function ($\exp(-\beta d_{ij})$):

$$M_{ij} = A_i B_j O_i D_j d_{ij}^{-\beta} \quad (3)$$

These mathematical models are calibrated using search routines that generate an optimum distance decay parameter by iteration from a given starting value on the basis of a measure such as the convergence between the predicted and observed mean migration distance (Stillwell, 1991). This approach was extended with the calibration of zone-specific distance decay parameters by Stillwell (1978) and the incorporation of a competing destinations variable to remove the effect of spatial structure by Fotheringham (1983; 1991). More recently, Fotheringham *et al.* (2001) have shown how the competing destinations model makes explicit the linkage between spatial choice behaviour at different levels in the spatial hierarchy.

Whilst models of this type have been used typically to estimate missing information in a historical context, less commonplace are examples of the application of these types of spatial interaction models for migration projection. One example is the model developed by Rees *et al.* (1990) to project ward populations in Swansea. In the context of projection, independent projections are required of out-migration and in-migration that may be derived from the extrapolation of historical trends or may be connected with projected explanatory variables as summarised by Stillwell (1991). Examples of studies in which projections of migration were tested against observed data are few and far between.

4.3 Spatial interaction models: statistical formulations

In parallel to the development of mathematically calibrated spatial interaction models, statistical modelling of inter-regional migration has also evolved and new forms of model have been introduced from the baseline gravity model specification outlined by Congdon (1991) in which the variables are log transformed and which has the form:

$$\log(M_{ij}) = b_0 + b_1\log(P_i) + b_2\log(P_j) + b_3\log(d_{ij}) + \varepsilon_{ij} \quad (4)$$

where b_0 is the constant and b_1, b_2 and b_3 are the regression coefficients associated with the relevant population terms and distance, and where ε_{ij} is the random error term associated with each interaction. This general liner model formulation is equivalent to an unconstrained spatial interaction model in the Wilson family.

One of the earliest developments was that by Lowry (1966) who extended the set of independent variables on the right-hand side of the equation and there have been a large number of studies subsequently that have sought to identify the most important determinants of migration. One of the basic assumptions of the linear model is that the observations are independent of one another and that the relationship between migration and the predictor variables is the same across each zone in the system of interest. The recognition that there are likely to be local variations in parameters has led to the application of geographically weighted regression (GWR) (Fotheringham *et al*., 2002) and the re-specification of the model in the following form:

$$\log M_{ij}(g) = b_0(g) + b_1(g)\log(X_i) + b_2(g)\log(Y_j) + \varepsilon_{ij} \qquad (5)$$

where Xi and Yj represent explanatory variables, and (g) indicates that the parameters are to be estimated at a location whose co-ordinates are given by the vector g.

The restrictive assumptions associated with the log-normal model have also led to the emergence of new statistical models based on the Poisson distribution (Congdon, 1991; Flowerdew, 1991). In the log-normal model, the error term and hence the dependent variable are assumed to be log-normally distributed continuous variates and the variance of the errors is constant regardless of the size of the estimation flow. The issue here is that the migration dependent variable is likely to be measured in discrete units (integer counts of persons) and follows a discrete probability distribution. This is also particularly important when there is likely to be a large number of small flows in the origin-destination matrix and a much smaller number of large flows. In terms of model structure, the Poisson regression equation becomes:

$$M_{ij} = \exp(b_0 + b_1\log P_i + b_2\log P_j + \log d_{ij}) + \varepsilon_{ij} \qquad (6)$$

In generalised linear modelling, a likelihood ratio statistic is used to assess how well the model fits the data. This statistic is called the deviance (D), and as the number of flows and the size of the flows increases, the deviance converges to the chi-squared distribution. Thus, the size of the deviance can be used to assess the goodness-of-fit of the model.

Scholten and Van Wissen (1985) compared the performance of spatial interaction models with log-linear approaches and concluded that using log-linear models with historical interaction parameters performed better than other approaches in terms of model fit and prediction. Flowerdew (1991) demonstrated that the possibilities of fitting Poisson regression models on quite large data sets using GLIM and the Poisson regression approach has been developed and adopted in several studies since then including Flowerdew and Lovett (1988), Amrhein and Flowerdew (1992), Bohara and Krieg (1996) and Boyle *et al.* (1998). More recently, the application of origin-specific Poisson models calibrated using GWR has been undertaken by Nakaya (2001) and similar models have been used to compare interregional migration in Japan and Britain by Yano *et al.* (2003). The Poisson approach is considered further in the next section.

4.4 Two-stage migration modelling

As indicated earlier, there is evidence that individuals often conceive of their migration as a two-stage process with worsening conditions at an origin eventually reaching a threshold level at which they decide to leave and then conditions at various locations being examined in order to decide on a suitable destination. This principle underpins the development of two-stage migration models. A state-of-the-art example is that constructed for the Office of the Deputy Prime Minister (ODPM) in the UK and involving the calibration of a policy-sensitive model of internal migration and the development of a user-friendly planning support system known as MIGMOD (MIGration MODeller). Detailed accounts of the work are provided in ODPM (2002), Champion *et al.* (2002), Rees *et al.* (2004) and Fotheringham *et al*. (2004).

MIGMOD approach. The central features of this approach are the separate modelling of: out-migration from each area based on a set of determinant variables (Stage 1); and the distribution of migrants between destinations also based on a set of determinants(Stage 2). The project also involved the development of an operational, user-friendly combination of Stages 1 and 2, enabling the model user to quickly set up and run a range of 'what if?' scenarios, to view the large volume of inputs and outputs, and to develop a selection of scenarios of determinant variables reflecting desired policy options. The data used in the model was from a time series of movement events from 1983-84 to 1997-98 recorded when National Health Service (NHS) patients re-register with doctors in different Family Health Service Authority (FHSA) (Stillwell, 1994). Age groups were eventually chosen for the model corresponding to childhood/schooling ages (0-15), the ages at which adolescents leave home for higher education (16-19), the ages at which students leave higher education for their working and partnership careers (20-24), the ages when they look for career advancement (25-29), the family formation ages (30-44), the later working ages (quiescent in terms of migration)(45-59), and the retirement and older ages (60+). Thus, the state-space involved calibrating the two-stage migration model for seven age groups and two sexes. The options of calibrating the out-migration model for each origin or for clusters of origins were ruled out in favour of an 'all origins together' calibration. Consequently, 14 separate models were calibrated. However, the situation for the destination choice model was different and an origin-specific distribution model was adopted, allowing the determinants to have different influences on the outcomes for each origin. Thus, it was possible to calibrate 98×14 or 1,372 separate models (for just one year). Data were obtained for 139 potential determinants of out-migration and 69 potential determinants of migration destination choice. Some explanatory variables were cross-sectional; others were available as a time series; some variables were lagged. In addition to variables measuring the characteristics of each zone, national variables were included and also regional variables were calculated for the Stage 1 model that were designed to capture the possible pull effects on out-migration caused by conditions elsewhere.

Stage 1: Out-migration model. The volume of out-migration from an origin i is predicted as:

$$O_{it}^{m} = omr_{it}^{m} P_{it}^{m} \tag{7}$$

where O_{it}^{m} is the total out-migration of migrant group m from zone i in time interval t, omr_{it}^{m} is the out-migration rate from origin i in time unit t for migrant group m (one of the 14 age-sex groups) and P_{it}^{m} is the population of migrant group m at risk of migrating from origin i during time interval t. The general form of the out-migration rate is as follows:

$$omr_{it}^{m} = f(X_{it/t-1}^{m}, Y_{it/t-1}^{m}, Z_{t-1}) \tag{8}$$

where omr_{it}^{m} is the out-migration rate for migrant group m from zone i in time interval t, $X_{it/t-1}^{m}$ is a vector of origin attributes in either year t or t-1 (lagged by one year); $Y_{it/t-1}^{m}$ is a vector of distance-weighted attributes describing the situation in other areas in either year t or t-1 and Z_{t-1} is a vector of attributes describing the national economic situation as it affects the overall volume of migration in year t-1 (lagged by one year).

Much debate focused on which specific form the model should take. Multiplicative or additive? Logged (the multiplicative option) or unlogged variables? Should non-linear forms of the variables such as quadratic forms be considered? After a number of experiments to identify optimum solutions, the final form of the model related the adjusted out-migration rate for migrant group m in zone i at time t to a new series of independent variables comprising cross-sectional (X), regional (Y) and national (Z) indicators, plus quadratic terms, a linear time trend (T) and a dummy (LD) for London:

$$omr_{it}^{m} = \kappa^{m} + \Sigma_p \alpha_p^{m} X_{pit/t-1}^{m} + \Sigma_q \beta_q^{m} Y_{qit/t-1}^{m} + \Sigma_r \gamma_r^{m} Z_{rt-1}^{m} + \Sigma_p \delta_p^{m} (X_{pit/t-1}^{m})^2 + \Sigma_q \eta_q^{m} (Y_{qit/t-1}^{m})^2 + \Sigma_r \theta_r^{m} (Z_{rt-1}^{m})^2 + \psi^{m} T_t + \zeta^{m} LD_i + \varepsilon_{it}^{m} \tag{9}$$

where ε_{it}^m is the error term for each zone, time and migrant group combination. A detailed review of the out-migration model can be found in Fotheringham *et al.* (2004).

Stage 2: Destination choice model. This involved the calibration of a migration destination model that distributes the total number of out-migrants from zone i to each of the destination zones based on the characteristics of each destination zone and the separation between the origin and each destination. The model to be calibrated has the general form:

$$M_{ij}^m = O_i^m \prod_p X_{pj}^{\alpha pim} d_{ij}^{\beta im} / \sum_j \prod_p X_{pj}^{\alpha pim} d_{ij}^{\beta im} \qquad (10)$$

where O_i^m is the volume of out-migration of type m from origin zone i; X_{pj} is an attribute of zone j that affects the choice of j by migrants from i; and d_{ij} is the distance between i and j. The X variables are raised to powers, α_{pim}, specific to each variable p, origin i and migrant group m, while the distance variable is raised to the power β_{im} specific to each origin i and migrant group m. The parameters of this model indicate the sensitivity of migration flows to particular destination characteristics; they indicate what features of a destination make it attractive to migrants and which features make it unattractive. For example, a relatively large score on an attribute with a positive parameter estimate would make a destination attractive to migrants, *ceteris paribus.* The model is calibrated separately for each of the origins and each of the 14 migrant groups. For any origin, $O_i^m / \sum_j \prod_p X_{pj}^{\alpha pim} d_{ij}^{\beta im}$ will be a constant (k_i^m) so that the origin-specific model is then simply,

$$M_{ij}^m = k_i^m \prod_p X_{pj}^{\alpha pim} d_{ij}^{\beta im} \qquad (11)$$

Poisson regression was therefore preferred over OLS since it assumes the conditional mean of the migrant variable has a Poisson distribution and avoids the need for making some approximation to zero flows. An initial set of 69 explanatory variables was reduced to 27 following a qualitative assessment of each of the variables in the data set and an examination of multicollinearity amongst the independent variables.

MIGMOD provides an example of how the form and content of a model may evolve over the duration of a research project. The model was constructed in two phases and many changes were made to the original specifications to overcome difficulties.

5 Migration models: demographic approaches

The importance of internal migration as a component of population change has been widely recognised by those responsible for creating sub-national population estimates and projections. Consequently, a second genre of approaches to modelling migration has developed within the field of multi-state demography whose aim has been to generate projections of migration flows without involving the type of detailed explanatory factors discussed in Section 2. Wilson (2001) provides a detailed review of the evolution of multi-regional demography, with a clear specification of the model equations.

5.1 Multi-state population projection modelling

The earliest population projections were usually produced using a cohort component model which, in the case of a single region, involved the estimation of the population at the beginning of a projection period, the projection of the number of births during the future time period and the survival of those in existence or being born during the period. Early examples of uni-regional models include those developed by Bowley (1924) in Britain, Weibol in the Netherlands (de Gans, 1999) and Whelpton (1936) in the USA. Leslie (1945; 1948) re-wrote the uni-regional model in matrix notation whilst others (e.g. Plane and Rogerson, 1994) demonstrated how the model could be expanded to include net-migration either in the form of flows or rates.

As far as modelling the migration component was concerned, it was the development of multi-regional demography in the mid-1960s that heralded the proper specification of inter-zonal flows rather than net-migration balances in projection models. Andrei Rogers (1966; 1967; 1968) pioneered the development of the Leslie matrix for a multi-region system and the creation of multi-region life tables (Rogers, 1973). He also provided the theoretical rationale for the use of migration flows rather than net balances in Rogers (1990). An alternative approach to the Rogers' multi-regional survival model known as accounts-based modelling was developed during the 1970s by Rees and Wilson (1977). Rees and Wilson constructed accounts-based models for transition data (involving the migration of those in existence at one point in time who were living at another address at an earlier point in time) in the first instance before applying similar techniques to movement migration (counts of moves taking place in a period irrespective of existence at the beginning or end points) (Rees, 1984). Willekens and Drewe (1984) brought the Rogers and Rees approaches together by switching from a dependence in the model on the multi-regional life table to period-cohort rates.

Thus, demographic models have developed from models requiring little information about migration to models requiring maximum information about migration, i.e., from aggregate net-migration balances, through migration pool, to migration flow information disaggregated by single year of age and sex. Population projection modelling has become more sophisticated as the migration component has been specified with more precision. Within this demographic modelling context, there are two key questions that relate to the internal migration component. The first of these is how to incorporate some form of change into the parameters that govern the intensity and pattern of migration during the projection period. The second is how to deal with the problem of huge data arrays when the origin-destination-time-age-sex dimensions are cross-classified. We discuss briefly each of these issues in turn, before reviewing some of the more recent research undertaken in the context of the development of multi-state models for European NUTS regions.

5.2 Temporal variability in model parameters

Many multi-regional population projection models do not in fact include any temporal variability in the origin-destination migration intensities upon which the model is based; they adopt the Markovian assumption that migration intensities will not change from one period to the next. However, there are some approaches that do try to build in some temporal variance. Plane and Rogerson (1986) discuss the use of causative matrices of ratios which link matrices of Markov intensities from one time period to another in the same way that it is possible to extrapolate from a geometric regression based on two data points. Feeney (1973), on the other hand, adjusts the Markov migration intensity by allowing the distribution of out-migrants to vary over time. This works by adjusting the base period intensity using the ratio of the destination region's share of the national population (excluding the origin region) at the start of the projection period to the same share recorded in the base period. The model is written as a probability of migrating between an origin and a destination in a projection period. An alternative probability approach is that termed the destination-population-weighted (DPW) model (Plane, 1982) which incorporates a balancing term to ensure that the probabilities sum to one. Some authors, including Fielding (1992) and Courgeau (1995) have suggested defining an origin-destination migration intensity based on the populations of both the origin and destination. This measure of migration velocity can be used in the same way as the traditional Markovian intensities but would require an adjustment if temporal variation was required. Pioneering work on the temporal stability of migration was undertaken in the 1980s in the Netherlands by Baydar (1983) who decomposed migration flows into an overall component or the total number of migrants in year t (N_t), a generation component or the probability of out-migration from region i in year t (o_{it}), and a distribution component or the probability of in-migrating to region j given origin i (p_{ijt}):

$$M_{ijt} = N_t \, o_{it} \, p_{ijt}, \quad i \neq j \tag{12}$$

and used a log-linear model to calibrate the parameters which quantify the time dependence of the different variables and thus identified the most stable and volatile components.

5.3 Shrinking dimensionality

The second issue revolves around the necessity to shrink large dimensional multi-regional models since the modern form of a demographic sub-national migration model is the multi-state model that uses migration flow information by age, sex, region of out-migration and region of in-migration. In its pure form, the multi-state migration model is highly descriptive: it has a separate parameter for every piece of information of the migration pattern. This means that the data requirements for the full multi-dimensional model are very large indeed. The creation of population projections for a system of 30 regions with 100 age groups and two sexes in any one year would involve 30(origins) x 29(destinations) x 100(age groups) x 2(sexes) = 174,000 flows. Research by van Imhoff *et al.* (1997) has shown how far it is possible to simplify (shrink) the structure of the multi-regional model before the resulting loss of information and accuracy becomes unacceptable.

5.4 Poisson modelling in a multi-state projection context

From a methodological point of view the multi-state model can be viewed as an accounting structure for a spatial interaction model. Both developments have converged using the framework of the Poisson regression model. The approach by van Imhoff *et al.* is particularly relevant here since their study was conducted in the context of the development of regional population projections at NUTS 2 level across the EU. Moreover, the projection method parallels that of the MIGMOD approach by separating the modelling into two stages: (i) the projection of out-migration by age and sex from each region; and (ii) the allocation of this pool of out-migrants to destinations. The second stage is known as a 'migrant pool' model because in-migration to destinations depends only on the size of the pool and not on the composition of the pool by region of origin. In the framework of log-linear modelling, the pool model corresponds to a hypothesis of independence between the origin and destination.

The approach assumes that interregional migration is classified along five dimensions referred to by letters: O (representing region of origin); D (region of destination); A (age); S (sex); and T (time period). Consequently, the observed count of migrants (or moves when registration data are being used) is represented by M_{ijast} where i and j are particular regions, a refers to one age group, s refers to males or females and t refers to one time period. The objective is to develop a model that describes each migration flow (or its corresponding rate) as the product of a limited number of parameters and then to examine the relative significance of the parameters. This approach therefore seeks to answer questions such as: Are the parameters representing sex more important than those representing age? How important is the origin effect? Is the time trend significant? It also allows the significance of relationships between dimensions to be identified, the so-called interaction effects, e.g. between particular origins and destination regions or between certain age groups and sex.

The Generalised Linear Modelling (GLIM) framework provides a suitable context for estimating the parameters of this type of model and log-linear regression models can be calibrated using a maximum likelihood algorithm available in the GLIM software package. The Poisson model is particularly useful because it produces unbiased parameter estimates, even in the case of over-dispersion in the data set. The parameter values are automatically normalised in GLIM and there is always a one-to-one correspondence between the number of parameters and the degrees of freedom in any model. Unlike the MIGMOD approach in which separate log-linear models are fitted for each age group for males and females, log-linear modelling using GLIM in this context can make use of the complete data sets and therefore it can take a long time to calibrate all the parameters. Once calibrated, the goodness of fit of a model is measured in GLIM using a deviance statistic. If an additional variable is introduced and there is an overall reduction in deviance, the latter reduction gives a measure of the importance of the new variable. It is usual in this type of modelling to begin from the null model (or grand mean model) in which M_{ijast} is estimated as simply the average flow (or rate) in the system. This serves as the baseline against which the results of other models (incorporating coefficients to identify effects of origin, destination, age, sex and time) can be compared. The deviance is also equal to the entropy statistic, which is used frequently in spatial interaction modelling.

Van Imhoff *et al.* (1997) calibrated models for the Netherlands, Italy and the UK specifically for the purpose of investigating to what extent the full multi-dimensional migration matrix could be simplified without seriously affecting the performance of the model. Their results indicate that a model of reasonable fit should contain at least the following interactions: origin-destination (OD), age-origin (AO), age-destination (AD) and sex-age (SA). In other words, the best model requires interactions among age, sex and origin and similarly between age, sex and destination, but the origin-destination effects are independent of age and sex. It was also found that time interacts with the main effects only (i.e. with age/sex, with origin, and with destination) and the remaining components (e.g. age/sex origin, age/sex destination, origin/destination, can be held constant). The absence or presence of interactions with the time dimension is crucial for using the model in internal migration projections. For making assumptions about internal migration the time invariant components need not be taken into account, and explicit hypotheses are only necessary about the time varying components. In a subsequent article (Van der Gaag *et al.,* 2000) explicit hypotheses were made about each of the time interactions in the model for projection purposes. For the time trend of the origin effects O as well as for the time trend of the destination effect D, three scenarios were proposed: (1) convergence, (2) divergence or (3) status quo. Convergence in the origin dimension implies that all origin-specific out-migration rates converge towards a common level, whereas divergence implies the reverse process in which existing differences become larger. Convergence in the destination dimension implies that the attractiveness of all zones converge towards a level which is proportional to their population size; divergence implies the opposite process whereby existing differences, standardised by their population size, enlarge. These convergence-divergence scenarios were used in the sub-national population projections for the EU at the NUTS 2 level in 1995.

6 Conclusions

The following conclusions have been drawn from this review of migration modelling. Firstly, the review indicates that there is a long tradition of modelling internal migration and a wide variety of approaches that can be differentiated into those based on *micro* or individual decision-making and those that deal with *macro* effects on aggregate flows. It is macro approaches to migration modelling that are applicable in the context of this study. Secondly, state-of-the-art models divide the migration process into two parts: the first part models the out-migration from each origin; and the second part models *destination choice*. Certain models generate a pool of out-migrants which are allocated to different destinations, but it is preferable to allocate migrants from each origin to each destination because there are frequently important factors that link certain origin and destination pairs.

Thirdly, it is important to recognise the difference between those causal factors or variables that *determine migration* (such as marriage or job opportunities) and those factors that have a *selective influence* on migration (such as age or social class). It is essential to develop models of out-migration and destination choice that are age-specific and which divide the aggregate flow into appropriate life course groups. Sex is less important but should be incorporated if possible but there are likely to be severe data constraints on any further disaggregation by composition and this is why relatively few models of more disaggregated migration have been calibrated.

We must recognise that migration flows, even of specific age groups, involve bundles of individuals motivated to migrate between regions for different *combinations of reasons*. In some cases, there is anecdotal evidence to support the importance of a particular variable in determining a chosen destination (such as a good school) but very little research has been undertaken to quantify its influence. Gravity variables that include the size of the origins and destinations and the intervening distances between origins and destination have proven to be important determinants in past studies but statistical relationships (signs, significance) between migration and many explanatory variables (unemployment, wages) have turned out to be specific to the system of spatial units being used and the national socio-economic conditions prevalent at the time. It is tempting to try and build a model containing a large number of explanatory variables but this makes huge demands on data collection, problems of autocorrelation and lack of clarity in interpretation. It is a well-known axiom of migration that more people migrate over shorter distances than longer distances. Consequently, zone size is very important since systems with smaller zones are likely to pick up more residential migrants who are not changing their jobs and will have very different motivations from those moving job as well as house.

From the developments in multi-state demography has emerged another genre of modelling internal migration that seeks to identify those *demographic influences* which have an important influence on the stability of migration flows over time and which distinguishes those direct and interaction effects between origin, destination, age, sex and time dimensions that are most important and should be incorporated within a general model, even though different variants are applicable in different countries.

Approaches to macro migration modelling have various alternative formulations and make use of *different mathematical or statistical calibration techniques*. Several studies have emphasised the benefits of the use of the general linear modelling approach in fitting explanatory models of migration and the application of the Poisson model has been used both in modelling sub-national migration to explain the relative effects of exogenous explanatory variable on out- or in-migration, and on origin-destination migration. It has also been used to investigate the importance of using fewer parameters than those suggested in a full multi-state model and of using migration flows dependent on other endogenous variables, for instance population size or composition. Standard software (GLIM) is available to calibrate Poisson models although there may be some difficulties in using this package when dimensions are big and the number of cases are being modelled is enormous.

It is clear that internal migration is influenced by various explanatory determinants and that demographic dimensions such as age and sex are important selective influences, but it is also clear that migration is a phenomenon that experiences *historical dependence*. The two types of migration modelling that have been identified in this review, gravity-based models and demographic models, might be usefully brought together to provide a hybrid approach that allows the impacts of both dimensions to be evaluated. Having said this, modelling historical flow patterns and projecting what will happen in the future should not necessarily be considered to require the same model. A good explanatory model of migration distribution probabilities may prove much less effective in a projection context, in comparison with a model based on historical flows, for example, simply because of the inadequacies of the projection of the independent variables. However, one of the key features of a projection model may be to test out how sensitive migration is to policy measures such as job creation or house-building programmes. Consequently, experience suggests that a modelling system would be particularly useful if it provided users with the means to experiment with alternative scenarios based on policy related variables whilst also allowing for results to be simulated under a 'do nothing' assumption. Furthermore, some testing of a projection model against observed data should be undertaken where data permits.

References

Amrhein, C.G. and Flowerdew, R. (1992) The effect of data aggregation on a Poisson regression model of Canadian migration, *Environment and Planning A*, 24: 1381-1391.

Armitage, R. (1986) Population projections for English local authority areas, *Population Trends*, 43(Spring): 31-40.

Bates, J.J. and Bracken, I. (1987) Estimation of migration profiles for local authority areas in England, 1971-1981, *Environment and Planning A*, 19(4): 521-35.

Baydar, N. (1983) Analysis of the temporal stability of migration in the context of multi-regional forecasting, *Working Paper No 38,* Netherlands Interuniversity Demographic Institute, Voorburg.

Bohara, A.K. and Krieg, R.G. (1996) A zero-inflated Poisson model of migration frequency, *International Regional Science Review*, 19(3): 211-22.

Bowley, A.L. (1924) Births and population change in Great Britain, *The Economic Journal,* 34: 188-192.

Boyle, P. (1993) Modelling the relationship between tenure and migration in England and Wales, *Transactions, Institute of British Geographers*, 18: 359-376.

Boyle, P. J., Flowerdew, R. and Shen, J. (1998) Analysing local-area migration using data from the 1991 Census: the importance of housing growth and tenure, *Regional Studies,* 32(2): 113-132.

Cadwallader, M. (1989) A conceptual framework for analysing migration behaviour in the developed world, *Progress in Human Geography*, 13(4): 494-511.

Champion, A.G., Bramley, G., Fotheringham, A.S., Macgill, J. and Rees, P.H. (2002) A migration modelling system to support government decision-making, in Stillwell, J. and Geertman, S. (eds.) *Planning Support Systems in Practice*, Springer Verlag, Berlin, pp.257-278.

Champion, A.G.., Fotheringham, A.S., Rees, P., Boyle, P.H. and Stillwell, J.C.H. (1998) The *Determinants of Migration Flows in England: a Review of Existing Data and Evidence*, Department of Geography, University of Newcastle, for the Department of Environment, Transport and the Regions.

Congdon, P. (1991) An application of general linear modelling to migration in London and the South East, Chapter 7 in Stillwell, J.C.H. and Congdon, P. (eds.) *Migration Models: Macro and Micro Perspectives*, Belhaven Press, London, pp. 113-136.

Courgeau, D. (1995) Migration theories and behavioural models, *International Journal of Population Geography*, 1: 19-27.

De Gans, H.A. (1999) *Population Forecasting 1895-1945: The Transition to Modernity,* Kluwer Academic, Dordrecht.

Devis, T. (1983) People changing address: 1971 and 1981, *Population Trends*, 32: 15-20.

Easterlin, R.A. (1980) *Birth and Fortune. The Impact of Numbers on Personal Welfare*, Basic Books, New York.

Feeney, G. (1973) Two models for multiregional population dynamics, *Environment and Planning*, 5: 31-43.

Fielding, A.J. (1992) Migration and social change, in Stillwell, J.C.H., Rees, P.H. and Boden, P. (eds.) *Migration Processes and Patterns Volume 2: Population Redistribution in the United Kingdom*, Belhaven, London.

Flowerdew, R. (1991) Poisson regression modelling of migration, in Stillwell, J.C.H. and Congdon, P. (eds.) *Migration Models: Macro and Micro Approaches*, Belhaven Press, London, pp. 92-112.

Flowerdew, R. and Lovett, A. (1988) Fitting constrained Poisson regression models to interurban migration flows. *Geographical Analysis,* 20: 297-307.

Fotheringham, A.S. (1983) A new set of spatial interaction models: the theory of competing destinations, *Environment and Planning A*, 15: 15-36.

Fotheringham, A.S. (1986) Modelling hierarchical destination choice, *Environment and Planning A*, 18: 401-418.

Fotheringham, A.S. (1991) Migration and spatial structure: the development of the competing destinations model, in Stillwell, J. and Congdon, P. (eds.) *Migration Models: Macro and Micro Approaches,* Belhaven, London, pp. 57-72.

Fotheringham, A.S., Brunsdon, C. and Charlton, M. (2002) *Geographically Weighted Regression: The Analysis of Spatially Varying Relationships*, Wiley, Chichester.

Fotheringham, A.S., Nakaya, T, Yano, K., Openshaw, S. and Ishikawa, Y. (2001) Hierarchical destination choice and spatial interaction modelling: a simulation experiment, *Environment and Planning A*, 33: 901-920.

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

Hughes, G. and McCormick, B. (1981) Do council housing policies reduce migration between regions?, *Economic Journal*, 91: 919-937.

Hughes, G. and McCormick, B. (1985) Migration intentions in the UK. Which households want to migrate and which succeed?, *Economic Journal, Conference Supplement,* 95: 113-123.

Hughes, G. and McCormick, B. (1987) Housing markets, unemployment and labour market flexibility in the UK, *European Economic Review*, 31: 615-645.

Hughes, G. and McCormick, B. (1989) Does migration reduce differentials in regional unemployment rates?, in van Dijk, J., Folmer, H., Herzog, Jr, H.W. and Schlottmann, A.M. (eds.) *Migration and Labour Market Adjustment*, Kluwer, Dordrecht, pp. 85-108.

Lee, E.(1966) Theory of migration, in Hauser, P.H. and Duncan, O.D. (eds.) *The Study of Population,* University of Chicago Press, Chicago, pp. 282-297.

Lee, E. (1969) A theory of migration, in Jackson, J.A. (ed.) *Migration*, Cambridge University Press.

Leslie, P.H. (1945) On the use of matrices in certain population mathematics, *Biometrika*, 33: 183-212.
Leslie, P.H. (1948) Some further notes on the use of matrices in population mathematics, *Biometrika,* 35: 145-213.
Liaw, K-L. and Ledent, J. (1987) Nested logit model and maximum quasi-likelihood method: a flexible methodology for analysing interregional migration patterns, *Regional Science and Urban Economics*, 17: 67-88.
Lowry, I. (1966) *Migration and Metropolitan Growth: Two Analytical Reports*, Chandler, San Francisco.
Maier, G. and Weiss, P. (1991) The discrete choice approach to migration modelling, Chapter 2 in Stillwell, J.C.H. and Congdon, P. (eds.) *Migration Models: Macro and Macro Approaches*, Belhaven Press, London, pp.17-33.
Myrdal, G. (1957) *Rich Lands and Poor*, Harper and Row, New York.
Nakaya, T. (2001) Local spatial interaction modelling based on the geographically weighted regression approach, *GeoJournal*, 53: 347-358.
ODPM (2002) *Development of a Migration Model*, Report prepared by the University of Newcastle upon Tyne, the University of Leeds, and the Greater London Authority/London Research Centre, ODPM, London.
Owen, D. and Green, A. (1992) Migration patterns and trends, in Champion, A.G. and Fielding, A.J. (eds.) *Migration Processes and Patterns Volume 1: Research Progress and Prospects*, Belhaven, London.
Pandit, K. (1997) Cohort and period effects in US migration: how demographic and economic cycles influence the migration schedule, *Annals of the Association of American Geographers*, 87: 439-450.
Plane, D.A. (1982) An information theoretic approach to the estimation of migration flows, *Journal of Regional Science*, 22: 441-456.
Plane, D.A. (1992) Demographic influences on migration, *Regional Studies*, 27(4): 375-383.
Plane, D.A. and Rogerson, P.A. (1986) Dynamic flow modelling with inter-regional dependency effects: an application to structural change in the US migration system, *Demography*, 23: 91-104.
Plane, D.A. and Rogerson, P.A. (1991) Tracking the baby boom, the baby bust, and the echo generations: how age composition regulates US migration, *Professional Geographer*, 43: 416-30.
Plane, D.A. and Rogerson, P.A. (1994) *The Geographical Analysis of Population with Applications to Planning and Business,* Wiley, New York.
Ravenstein, E.G. (1885) The laws of migration, *Journal of the Royal Statistical Society*, 48(II): 167-227.
Rees, P.H. (1984) Spatial population analysis using movement data and accounting methods: theory, models, the MOVE program and examples, *Working Paper 404/Computer Manual 23*, School of Geography, University of Leeds, Leeds.
Rees, P.H., Fotheringham, A.S. and Champion, A.G. (2004) Modelling migration for policy analysis, Chapter 14 in Stillwell, J. and Clarke, G. (eds.) *Applied GIS and Spatial Analysis*, Wiley, Chichester, pp. 259-296.
Rees, P.H., Stillwell, J.C.H., Boden, P., Duley, C., Ramsden, C., Bristow, M., Patel, I. and Rees, G. (1990) *SWIS: An Information and Projection System for Ward Populations and Housing for the City of Swansea, Manual and Report*, Report for GMAP Ltd, University of Leeds, Leeds.
Rees, P.H. and Wilson, A.G. (1977) *Spatial Population Analysis*, Edward Arnold, London.
Rogers, A. (1966) Matrix methods of population analysis, *Journal of the American Institute of Planners,* 32: 40-44.

Rogers, A. (1967) Matrix analysis of interregional population growth and distribution, *Papers of the Regional Science Association*, 18: 17-196.
Rogers, A. (1968) *Matrix Analysis of Interregional Population Growth and Distribution*, University of California Press, Berkeley.
Rogers, A. (1973) The multi-regional life table, *Journal of Mathematical Sociology*, 3: 127-137.
Rogers, A. (1990) Requiem for the net migrant, *Geographical Analysis*, 22: 283-300.
Rogers, A. (1995) *Multiregional demography,* Wiley, London.
Rogers, A. and Castro, L.J. (1981) Model migration schedules, *RR-81-30*, International Institute of Applied Systems Analysis, Laxenburg, Austria.
Rogers, A., Raquillet, R. and Castro, L.J. (1978) Model migration schedules and their applications, *Environment and Planning A,* 10: 475-502.
Rogers, A. and Planck, F. (1984) MODEL: a general program for estimation of parameterized model schedules of fertility, mortality, migration and marital status and labour force status transitions, *Working Paper 83-102*, IIASA, Laxenburg, Austria.
Rogers, A. and Willekens, F. (eds.) (1986) *Migration and Settlement: A Multiregional Comparative Study,* Reidel, Dordecht.
Rossi, P. (1955) *Why Families Move: A Study in the Social Pschycology of Urban Residential Mobility*, Free Press, Glencoe, Illinois.
Rothenberg, J. (1977) On the microeconomics of internal migration, in Brown, A.A. and Neuberger, E. (eds.) *Internal Migration: A Comparative Perspective*, Academic Press, New York, pp. 183-208.
Scholten, H.J. and van Wissen, L.J.G. (1985) A comparison of the loglinear interaction model with other spatial interaction models. In: P. Nijkamp, H. Leitner and N. Wrigley (eds.) *Measuring the Unmeasurable*, Martinus Nijhoff Publishers, Dordrecht, pp. 177-196.
Sjaastad, L. (1960) The relationship between migration and income in the United States, *Papers of the Regional Science Association*, 6: 37-64.
Stambøl, L. (1991) Migration projection in Norway: a regional demographic-economic model, Chapter 15 in Stillwell, J.C.H. and Congdon, P. (eds.) *Migration Models: Macro and Micro Perspectives*, Belhaven Press, London, pp. 287-308.
Stillwell, J.C.H. (1978) Interzonal migration: some historical tests of spatial interaction models, *Environment and Planning A*, 10: 1187-1200.
Stillwell, J.C.H. (1991) Spatial interaction models and the propensity to migrate over distance, Chapter 3 in Stillwell, J.C.H. and Congdon, P. (eds.) *Migration Models: Micro and Macro Approaches*, Belhaven Press, London, pp. 34-56.
Stillwell, J.C.H. (1994) Monitoring intercensal migration in the United Kingdom, *Environment and Planning A*, 26: 1711-1730.
Stillwell, J.C.H., Bell, M., Blake, M., Duke-Williams, O. and Rees, P.H. (2001) A comparison of net-migration flows and migration effectiveness in Australia and Britain: Part 2, Age-related migration patterns, *Journal of Population Research*, 18(1): 19-39.
Stillwell, J.C.H. and Congdon, P. (1991) (eds.) *Migration Models: Macro and Macro Approaches*, Belhaven Press, London.
Stillwell, J.C.H., Duke-Williams, O. and Rees, P.H. (1995) Time series migration in Britain: the context for 1991 Census analysis, *Papers in Regional Science: Journal of the Regional Science Association*, 74(4): 341-359.
Thomas, D.S. (1938) *Research Memorandum of Migration Differentials*, Social Science Research Council, New York.

Van der Gaag, N., van Imhoff E. and van Wissen, L. (2000) Internal migration scenarios and regional population projections for the European Union, *International Journal of Population Geography*, vol. 6(1): 1-19.

Van Wissen, L.J.G. and Rima, A. (1988) *Modelling Urban Housing Market Dynamics,* Studies in Regional Science and Urban Economics, North Holland-Amsterdam.

Warnes, A. (1992) Migration and the life course, in Champion, A.G. and Fielding, A.J. (eds.) *Migration Processes and Patterns Volume 1: Research Progress and Prospects*, Belhaven, London.

Whelpton, P.K. (1936) An empirical method of calculating future population, *Journal of the American Statistical Association*, 31: 457-473.

Willekens, F. and Drewe, P. (1984) A multi-regional model for regional demographic projection, in ter Heide, H. and Willekens, F. J. (eds.) *Demographic Research and Spatial Policy: The Dutch Experience*, Academic Press, London, pp. 309-334.

Wilson, A.G. (1967) A statistical theory of spatial distribution models, *Transportation Research,* 1: 253-269.

Wilson, A.G. (1970) *Entropy in Urban and Regional Modelling*, Pion, London.

Wilson, A.G. (1972) *Papers in Urban and Regional Analysis*, Pion, London.

Wilson, T. (2001) A new subnational population projection model for the United Kingdom, *Unpublished PhD Thesis*, School of Geography, University of Leeds, Leeds.

Yano, K., Nakaya, T., Fotheringham, A.S., Openshaw, S. and Ishikawa, Y. (2003) A comparison of migration behaviour in Japan and Britain using spatial interaction models, *International Journal of Population Geography*, 9(5): 419-431.

Zipf, G.K. (1946) The P1P2/D hypothesis: on intercity movement of persons, *American Sociological Review*, 11: 677-86.