WORKING PAPER 270

An investigation of some alternative model-based approaches to projecting inter-regional migration.

JOHN C.H. STILLWELL

School of Geography
University of Leeds
LEEDS LS2 9JT

February 1980

CONTENTS

CONTENTS cont'd ....

ABSTRACT

The provision of reliable estimates or projections of current and future populations is of major importance to national, regional and local authorities. However, the relative paucity of gross migration data has acted as a brake on the development of migration analysis related to population projection research. Cancellation of the intermediate census in 1976 has meant that aggregate inter-area (regional or local authority) migration flows during the seventies are only available for the year that preceded the 1971 census. The absence of more recent information on directional movements has imposed quite severe limitations on those authorities and individuals whose concern is with population projection, services planning or policy evaluation.

In this paper, several methods for estimating, projecting and forecasting regional in- or out-migration totals and inter-regional migration flows are examined and tested against an observed data base for the five year period, 1966-71. The models which are evaluated and compared produce estimates or projections of migration which are necessary in order to construct complete sets of population accounts for either historical or forecasting periods.

List of Figures:

List of Tables:

List of Tables: cont'd ....

1. INTRODUCTION

Ramprakash (1973) affirms that

> ".... forecasting migration has proved the most interesting and intractable problem in population projections for regions and subdivisions. This is because of conceptual difficulties in defining migration, of data limitations, and of the policy and planning framework in which migration is seen to occur and is influenced."

These factors, together with the complexities involved in analysis of the process of migration, as well as the element of uncertainty that surrounds any assessment of what the future might hold, prompt us to recognise at the outset that migration projection is the least understood and therefore most unpredictable component of regional population projection, requiring 'above all else a sense of humour' in one author's opinion (Morrison, 1973).

It is essential, however, to recognise also that intra-national migration is the principal component of population redistribution, assuming increasing importance at smaller spatial scales in relation to urban and regional development and planning. While the methods and models suggested in this paper have been developed to assist inter-region migration projection, primarily because no suitable projection or forecasting methodology exists, they also represent a response to the demand "for the integration of analytic migration models into the demographic component framework" (Baxter and Williams, 1978).

2. THE DEFINITION OF TERMS AND THE FRAMEWORK

2.1 Some Problems with Definitions

The words prediction, projection, and forecasting have been used in the demographic and geographic literature without consistent discrimination. If we adopt the guidelines proposed by Brass (1974) or Pittenger (1976), a projection can be interpreted as a formal process of measuring a future condition which results if the defined assumptions embodied in the projection method turn out to be empirically valid. Inter-area migration projections or extrapolations may be produced, for example, under alternative assumptions about transition rates. It may be assumed that rates remain constant or that they change in a specific way. When a judgemental statement is made on the basis of an assessment of the plausibility of differing assumptions, the selected projection becomes the forecast. In other words, if the transition rates are assumed to remain constant, a forecast is generated specifically on this basis. The distinction reflects Pittenger's generalisation that ".... all forecasts are projections, but not all projections are forecasts."

A prediction might then be defined as a systematic procedure for calculating projections or forecasts, and distinctions between short, medium and long term population predictions indicate that the most popular time horizons attached to these periods are less than 5 years, 5-15 years and greater than 15 years. The difference between projecting population stocks and migration flows in this context is that population projection can be undertaken for any point in time, while migration, like births and deaths, has to be determined for a time period which, to be compatible with census data in the U.K., normally refers to a one year or five year period. The predictive models developed and tested in this paper are based on a five year projection period as a consequence of the analyses of inter-region migration for the historical period 1961-66 that were described in earlier papers (Stillwell, 1977; 1978).

One particular feature of most projection model studies is that the results, the projected migration flows, are never 'tested' against a set of observed flows. This statement sounds contradictory in the sense that projection implies the unavailability of observed information, but we can envisage circumstances in which projection can be accomplished for a historical period (θ) based wholly on data available for the previous historical period (θ -1), together with a necessary set of assumptions. In this case, 'true' projections can still be generated that can subsequently be compared with observed flows. The temptation is to employ additional known information for the period θ in order to generate a set of more realistic projections, 'pseudo'-projections, that are really estimates rather than projections in the strictest sense. This method is particularly relevant in a historical situation where no observed data on directional migration flows actually exists, but where information on 'other' demographic, economic or policy variables can be employed to help generate the migration distribution.

Two types of 'historical' projection context can therefore be distinguished:

(1) where observed migration flows are available (1966-71)

(2) where observed migration flows are not available (1971-76)

and in each of these periods choice can be made whether to use or to ignore information related to the projection period that would help in determining projected flows. Obviously comparison of the results of projection with observed flows would only be possible for the former period, and it is this period which is adopted as our projection period in the paper. Some of the models are developed on the basis of pre-1966 trends, while others make use of information available for the projection period itself. The time period label (θ) is used in combination with t and t+T, which are general labels used to indicate the initial and final points in any time period. The population variable $K^{i*}(t,\theta)$ therefore represents the population of region i at t, the beginning of the period θ. $K^{*i}(t+T, \theta-1)$ represents the population at t+T, the end of the period θ-1. θ is usually assumed to be the first projection period; θ-1, the latest historical period for which data is known; and θ+1, the second projection period. The three periods θ, θ-1 and θ+1 are equal in length and refer to 1966-71, 1961-66, and 1971-76 respectively in this analysis.

2.2 The Framework Adopted for Projection

The modelling framework (Figure 1) through which inter-region migration flow projections are determined involves three stages, the first two of which are analagous to the trip generation and distribution phases in transport modelling. The first stage is one where aggregate regional out-migration ($\hat{\underline{O}}_i$) and inmigration ($\hat{\underline{I}}_j$) vectors are projected using some combination of input data on population, previous migration, policy variables and births and deaths. Predicted totals are then distributed between regions in the second stage, sometimes utilising the outmigration and inmigration projections as row and column constraints and requiring additional information on inter-regional distances and decay parameters calibrated for a previous period. The projected inter-regional, exist-survive migration matrix, ($\hat{\underline{M}}_{ij}$), can then be used together with population, births and deaths statistics, to provide a complete set of population accounts, ($\hat{\underline{K}}_{ij}$), for the projection period. If a closed system accounting model is adopted, additional estimation of population, births and migration flows for a rest-of-the-world region will be required. The generation of sets of projected accounts is the third and final stage in the framework.

In the next section, a series of alternative distribution models are outlined which have been formulated under the assumption that regional totals of outmigration and inmigration can be projected in an accurate

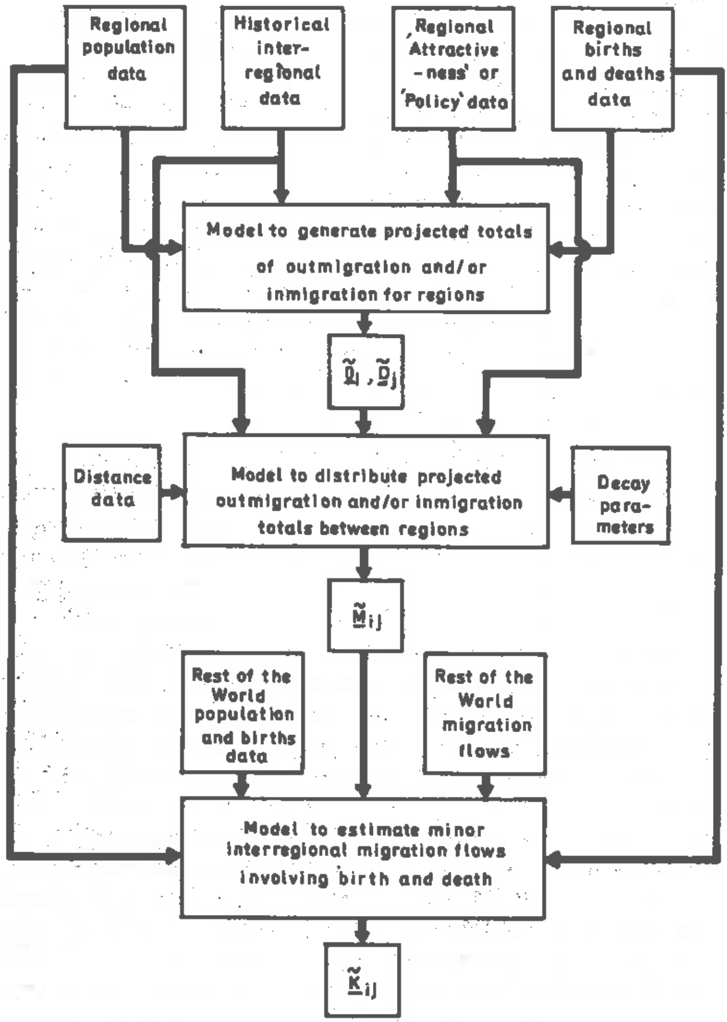


Figure 1: <u>A Framework for Inter-Region Migration Flow Projection</u>

and reliable way. Methods of total outmigration and inmigration projection are themselves reviewed in section 4 and the ability to predict variables that determine or are dependent on migration is also considered. In section 5, the results of the projection models are compared with a set of observed flows over 1966-71 for the regional system that is adopted. The regions illustrated in Figure 2 are identified as groups of aggregated-counties which were used initially by Stillwell (1979) and the inter-county flows have been aggregated to represent inter-regional flows using a consolidation program (Stillwell, 1976). Details of procedures necessary for the adjustment of the observed data for boundary changes occurring between 1966 and 1971 have been relegated to an appendix in order not to detract from the main theme of the paper.

Section 5 also contains some discussion of the preparation of sets of accounts for 1966-71 and 1971-76, and some analysis of the regional population characteristics identified from the accounts matrices for these two periods.

3. INTER-REGION MIGRATION DISTRIBUTION MODELS FOR PROJECTION PURPOSES

There exists a large variety of alternative approaches to modelling inter-area migration in a historical context and these have been reviewed, for example, by Weeden (1973) and Shaw (1975). Types of model have been differentiated by Stillwell (1975) according to the purposes for which they were formulated, their structural characteristics, the techniques that they incorporate, and the nature of migration data that is available for the system of interest concerned.

Migration projection or forecasting models on the other hand, tend to fall into one of two broad categories. The first type of projection model is that associated with probabilistic techniques which may involve Markov assumptions. Aggregate, inter-area flows can be predicted simply and directly using transition rates for the latest 'known' period and population stock statistics for the regions of the system at the start of the projection period. This type of model is particularly popular in forecasting because the only information necessary is the set of transition rates and the base population totals. The approach is based on the hypothesis that the pattern and volume of migration in the future will tend to be a function of past and current trends. Hagerstrand (1957) argues that this effect of historical dependence (the 'beaten path' effect) is created in part by the

Figure 2: The Regional System based on Aggregated-County Boundaries

existence of family contacts or business associations. While Markovian models have been developed in a policy context by Rogers (1966) and Mackinnon (1975), and used to analyse age/sex differences in inter-regional migration in Britain by Joseph (1975), more sophisticated stochastic models with heterogeneous transition matrices and models involving investigation of the distribution of times between moves, are discussed by Ginsburg (1973).

The alternative approach to projection is one that considers the functional relationship between migration and variables that in some way influence the propensity to move. Historical tests of deterministic models have been undertaken by researchers concerned primarily with predicting gross migration flows. These deterministic approaches vary from the use of extended gravity models incorporating variables such as per capita wage levels and unemployment rates (Lowry, 1966; Rogers, 1967; Masser, 1970) to multi-variate regression formulations involving measures such as the ratio of employment in manufacturing industry to the area of new industrial building (Hart, 1970) or to econometric studies involving the lagged response of potential migrants to monetary incentives (Lianos, 1972). The dependence-independence debate centred on the migration-employment relationship has resulted in a number of simultaneous equations approaches; notably those by Okun (1968), Greenwood (1973, 1975a, 1975b), and more recently by Rogers (1976), Ledent (1978), and Gober-Meyers (1978). Attempts have also been made to bring together probabilistic and deterministic theory by advancing a causal interpretation of inter-regional transition rates, based on analysis of the effects of cumulative inertia and of the intrinsic attractiveness of destinations (Cordey-Hayes and Gleave, 1974; 1975).

In the absence of directional migration flows data, analysis has been conducted on some occasions with totals of out-, in- or residual net-migration (Blanco, 1963; Oliver, 1964; Lowry, 1966; Jack, 1970; Morrison and Relles, 1975), although none of these studies seeks to estimate the inter-regional distribution of migrants on the basis of these totals. Most deterministic models have been constructed and tested for explanatory reasons and many have been calibrated using linear regression techniques with log-transformed variables. However, Senior (1979) has illustrated the problems of exaggeration which occur with simple gravity models when used in forecasting contexts, and it has become evident that constrained spatial interaction models (Wilson, 1974) are more suitable for this purpose because balancing factors can be introduced which ensure internal consistency.

Constrained spatial interaction models based on entropy-maximising methodology are derived so that "errors of estimate generated by the method are tolerably small and without systematic bias" (Lowry, 1966). This is one of the conditions that Lowry recognised as being necessary for a satisfactory forecasting model. Two other conditions quoted by Drewe (1971) are that "values for independent variables needed to implement the forecast can be obtained by some reasonable procedure; and that the fitted parameters are applicable to the particular context of the forecast as well as to the context in which they were chosen". One of the objectives of this paper is to establish to what extent different procedures for estimating independent variables are reasonable. The issue of the applicability of calibrated parameters to forecasting is also considered. However, a selection of distribution models is outlined initially to illustrate the variety of 'demands' for information made by alternative models. We begin with a simple rates model.

3.1 Transition Rates Models

Inter-regional transition rates (m_{ij}) are calculated for a historical time period as:

$$m_{ij}\,(\theta-1) = M_{ij}\,(\theta-1)\,/K^{i*}(t,\theta-1),\ i \neq j \qquad (1)$$

where $K^{i*}(t,\theta-1)$ is the population of region i at t, the commencement of the historical period (θ-1), as represented in the accounts notation of Rees and Wilson (1973), and where $M_{ij}(\theta-1)$ is the flow of exist-survive migrants between region i and region j during the historical period θ-1. Transition rates defined in this way can be applied to generate a projected migration distribution ($\hat{M}_{ij}$) for a future period from:

$$\hat{M}_{ij}(\theta) = m_{ij}(\theta-1)\ K^{i*}(t,\theta),\ i \neq j \qquad (2)$$

The symbol ∿ is used throughout the paper to identify a projected variable.

On some occasions, it is appropriate to calculate a 'true' admission rate as:

$$m_{ij}(\theta-1) = M_{ij}(\theta-1)\,/K^{*j}(t+T,\theta-1),\ i \neq j \qquad (3)$$

but this is less useful in projection since the end-of-the-period destination region population is frequently unknown. When a final population total is available, a 'pseudo-admission rate' may be used, which is defined as:

$$m_{ij}(\theta-1) = M_{ij}(\theta-1) / K^{j*}(t,\theta-1),\ i \neq j \qquad (4)$$

While the accounts-based model developed by Rees and Wilson (1977) normally includes an estimation of all flows between the regions of the system of interest and a residual rest-of-the-world region, the transition rates model defined here applies only to the 'internal' migration flows between the regions of England and Wales and Scotland. This is so that comparison can be made between the projections generated by models that do not cater for external flows. Emigration and immigration flows are assumed to be exogeneous inputs to the accounts-based model that is used later in the paper. Intra-regional flows are also excluded from the analysis.

3.2 Growth Factor Models

The growth factor methodology involves the use of ratios between projected and observed totals of outmigration and inmigration. The model, adapted from Wilson (1974) has the form:

$$\hat{M}_{ij}(\theta) = g_{ij}\ M_{ij}(\theta-1),\ i \neq j \qquad (5)$$

where the growth factors for regions i and j are combined into:

$$g_{ij} = a_i\ b_j,\ i \neq j \qquad (6)$$

and defined individually as:

$$a_i = \frac{\hat{O}_i\ (\theta)}{O_i(\theta-1)} \qquad (7)$$

$$b_j = \frac{\hat{D}_j(\theta)}{D_j(\theta-1)} \qquad (8)$$

where $\hat{O}_i(\theta)$ = the projected total outmigration and survival from region i for period θ.

$\hat{D}_j(\theta)$ = the projected total inmigration and survival to region i during the period θ.

$O_i(\theta-1)$ = the observed total outmigration and survival from region i for period θ-1.

$D_j(\theta-1)$ = the observed total inmigration and survival to region i during the period $\theta-1$.

The model equation (5) can be extended to ensure that the following constraint equations hold for flows in the projection period:

$$\sum_{j \neq i} \hat{M}_{ij}(\theta) = \hat{O}_i(\theta) \tag{9}$$

and

$$\sum_{i \neq j} \hat{M}_{ij}(\theta) = \hat{D}_j(\theta) \tag{10}$$

and therefore that:

$$\sum_{i \neq j} \sum_{j \neq i} \hat{M}_{ij}(\theta) = \sum_i \hat{O}_i(\theta) = \sum_j \hat{D}_j(\theta) \tag{11}$$

These conditions are implicit in the historical case. Balancing factors are then introduced as:

$$A_i = \hat{O}_i \Big/ \sum_{j \neq i} B_j\, g_{ij}\, M_{ij}\,(\theta-1),\ i \neq j \tag{12}$$

$$B_j = \hat{D}_j \Big/ \sum_{i \neq j} A_i\, g_{ij}\, M_{ij}\,(\theta-1),\ j \neq i \tag{13}$$

and the model equation can be rewritten as

$$\hat{M}_{ij}(\theta) = A_i\, B_j\, g_{ij}\, M_{ij}\,(\theta-1),\ i \neq j \tag{14}$$

A multiplicative assumption is used for defining g_{ij} rather than the additive assumption of the simple average factor method. The square root of the product is sometimes taken for dimensional reasons, although this is not important when the balancing factors are incorporated.

3.3 Doubly Constrained Spatial Interaction Models

Several versions of the doubly constrained spatial interaction model have been calibrated and tested with inter-regional migration data for the 1961-66 period. Six of these models may be redefined as potential projection models as follows:

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; d_{ij}^{-\beta_*^{(1)}(\theta-1)}(\theta-1), \; i \neq j \quad (15)$$

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; e^{-\beta_*^{(2)}(\theta-1)d_{ij}\;(\theta-1)}, \; i \neq j \quad (16)$$

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; d_{ij}^{-\beta_i^{(1)}(\theta-1)}(\theta-1), \; i \neq j \quad (17)$$

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; e^{-\beta_i^{(2)}(\theta-1)d_{ij}\;(\theta-1)}, \; i \neq j \quad (18)$$

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; d_{ij}^{-\beta_j^{(1)}(\theta-1)}(\theta-1), \; i \neq j \quad (19)$$

$$\hat{M}_{ij}(\theta) = A_i \; B_j \; \tilde{O}_i \; \tilde{D}_j \; e^{-\beta_j^{(2)}(\theta-1)d_{ij}\;(\theta-1)}, \; i \neq j \quad (20).$$

where the time period label (θ) has been omitted from the $\tilde{O}_i$, $\tilde{D}_j$ and the balancing factor terms. Model equations (15) and (16) are standard doubly constrained models with negative power and negative exponential decay functions respectively. Model equations (17) to (20) are versions with either origin- or destination-specific parameters. All the equations assume that the set of inter-regional distance (d_{ij}) values that are adopted (Table 1) remain unchanged and that the beta values (Table 2) calibrated for 1961-66 (θ-1) are appropriate for the following 5 year period (θ). Both these variables might however change over time. The d_{ij} terms might change if alternative population centroids were selected or if the road network altered significantly, and both of these features might then affect the propensity to migrate over distance. Although the distances and beta values are assumed to be the same for both periods in the initial projection runs, it might be appropriate to make certain changes for projection runs in subsequent projection periods.

The historical analysis reported by Stillwell (1978) also included an investigation of the effect of modelling flows between contiguous regions separately from flows between non-contiguous regions. It was found that this form of disaggregation made considerable improvements to the overall model fit. One additional problem associated with using a contiguity/non-contiguity projection model, defined either with a negative power function as:

1	N Newcastle	North									
2	NW Manchester	129	North West								
3	YH Leeds	92	40	Yorkshire & Humberside							
4	EM Leicester	181	87	96	East Midlands						
5	WM Birmingham	201	80	109	39	West Midlands					
6	EA Cambridge	229	154	144	68	100	East Anglia				
7	SE London	274	184	190	98	110	54	South East			
8	SW Bristol	285	159	194	116	88	143	116	South West		
9	W Cardiff	299	172	209	137	102	174	154	44	Wales	
10	S Glasgow	143	212	210	299	289	350	394	366	370	Scotland

Table 1: <u>Matrix of Road Mileage Distances between Selected Regional Centroids</u>

REGION	Calibrated Beta Values, 1961-66 (and Mean Migration Lengths)					
	(1) β_*	(2) β_*	(1) β_i	(2) β_i	(1) β_j	(2) β_j
1. North			1.3226	0.0069	1.2428	0.0066
			(193.2)		(180.2)	
2. North West			0.6073	0.0061	0.5835	0.0062
			(136.1)		(134.6	
3. Yorkshire and Humberside			1.0602	0.0104	0.8679	0.0089
			(123.6)		(124.5)	
4. East Midlands			0.7997	0.0079	0.4823	0.0046
			(98.5)		(112.7)	
5. West Midlands			0.7410	0.0074	0.4985	0.0047
			(100.3)		(116.6)	
6. East Anglia			0.6333	0.0057	1.0762	0.0093
			(96.9)		(90.7)	
7. South East			0.6800	0.0046	0.6084	0.0039
			(142.2)		(170.5)	
8. South West			0.6638	0.0050	1.0718	0.0073
			(129.8)		(129.9)	
9. Wales			0.5276	0.0046	0.6777	0.0052
			(139.4)		(153.1)	
10. Scotland			0.6949	0.0025	0.1776	0.0006
			(309.1)		(306.4)	
All Regions	0.7723	0.0052				
	(145.9)					

Bracketed values represent mean migration lengths

Table 2: Generalised, Origin-specific and Destination-specific Distance Decay Parameters and Mean Migration Lengths for the Regions, 1961-66

$$\hat{M}_{ij}^{Q}(\theta) = A_i^Q B_j^Q \hat{O}_i^Q \hat{D}_j^Q d_{ij}^{Q\;-\beta_*^Q(\theta-1)}(\theta-1), \; i \neq j \quad (21)$$

$$, \; Q = C,N$$

or with a negative exponential function as:

$$\hat{M}_{ij}^{Q}(\theta) = A_i^Q B_j^Q \hat{O}_i^Q \hat{D}_j^Q e^{-\beta_*^Q(\theta-1) d_{ij}^Q(\theta-1)}, \; i \neq j \quad (22)$$

$$, \; Q = C,N$$

is that independent projections of total outmigration and inmigration require to be disaggregated according to whether they occurred between contiguous or non-contiguous regions. This disaggregation might be achieved crudely by splitting up the aggregate $\hat{O}_i$ and $\hat{D}_j$ projections according to the observed contiguous and non-contiguous flow proportions evident in the preceding historical period.

3.4 Singly Constrained Spatial Interaction Models

Production or attraction constrained variants of the family of spatial interaction models with alternative distance functions, can be defined for inter-regional migration distribution as follows:

$$\hat{M}_{ij}(\theta) = A_i \hat{O}_i \hat{W}_j^{(2)} d_{ij}^{\;-\beta_*^{(1)}(\theta-1)}(\theta-1), \; i \neq j \quad (23)$$

$$\hat{M}_{ij}(\theta) = A_i \hat{O}_i \hat{W}_j^{(2)} e^{-\beta_*^{(2)}(\theta-1) d_{ij}(\theta-1)}, \; i \neq j \quad (24)$$

$$\hat{M}_{ij}(\theta) = B_j \hat{W}_i^{(1)} \hat{D}_j d_{ij}^{\;-\beta_*^{(1)}(\theta-1)}(\theta-1), \; i \neq j \quad (25)$$

$$\hat{M}_{ij}(\theta) = B_j \hat{W}_i^{(1)} \hat{D}_j e^{-\beta_*^{(2)}(\theta-1) d_{ij}(\theta-1)}, \; i \neq j \quad (26)$$

One advantage of this type of model is that it is unnecessary to provide projections for both $\hat{O}_i$ and $\hat{D}_j$ at the same time; and therefore in these equations, $\hat{W}_i^{(1)}$ and $\hat{W}_j^{(2)}$ refer to projected 'attractiveness' factors that may be measured using variables such as origin and destination population size.

Eight singly constrained migration models have been calibrated on the basis of 1961-66 inter-regional data. They incorporate either base population totals or numbers in employment* as origin and destination region attractiveness factors. The decay parameters and coefficients of determination are indicated in Table 3.

Model Equation	Attractiveness Factors			
	Population size		Employment total	
	β *	r^2	β *	r^2
(23)	0.8114	0.65516	0.7972	0.57276
(24)	0.0051	0.64989	0.0051	0.56377
(25)	0.5641	0.81322	0.5328	0.75726
(26)	0.0035	0.79895	0.0033	0.74337

Table 3: <u>Distance Decay Parameters and Coefficients of Determination associated with Historical Singly Constrained Inter-regional Migration Models, 1961-66</u>

The 'best-fit' model of this sub-set appears to be an attraction-constrained version where the attractiveness of the origin region is measured by population size, although in general, the goodness-of-fit statistics are inferior to those associated with doubly constrained models (Stillwell, 1978). It should also be understood that the attractiveness variables and their method of measurement may require refinement in relation to the demand for a policy sensitive distribution model.

The models that have been suggested in this section for projecting the distribution of inter-region migration are dependent upon certain exogenous data inputs. Regional totals of projected outmigration and in-migration in particular are required by each type of model except the initial transition rates model. The methods by which alternative projected totals can be obtained are now considered.

---

* the regional employment totals are the numbers of economically active persons in employment in 1966 (G.R.O., 1968).

4. PROJECTING TOTAL OUTMIGRATION AND TOTAL INMIGRATION FOR REGIONS

It is possible to envisage a number of alernative approaches to the projection of total regional out- and in-migration. In this section, three methodologies are investigated. The first, based on historical rates, is perhaps the most traditional approach. The second method is referred to as the 'net migration' approach and makes use of residual net migration information and functional relationships that appear to exist between aggregate migration rates. The third approach is an attempt to introduce some regional development indicators into the analysis and to investigate the potential for building a policy-sensitive projection model.

4.1 Total Migration Rates Models

The most straightforward method of projecting totals of outmigration and inmigration is to apply an historical aggregate rate to a start-of-the-projection-period population. In the case of outmigration, this would be:

$$\tilde{O}_i(\theta) = o_i(\theta-1)K^{i*}(t,\theta) \quad (27)$$

where

$$o_i(\theta-1) = O_i(\theta-1)/K^{i*}(t,\theta-1) \quad (28)$$

and

$$O_i(\theta-1) = \sum_{j\neq i} M_{ij}(\theta-1) \quad (29)$$

and the corresponding inmigration rate model would be:

$$\tilde{D}_j(\theta) = d_j(\theta-1)K^{j*}(t,\theta) \quad (30)$$

where

$$d_j(\theta-1) = D_j(\theta-1)K^{j*}(t,\theta-1) \quad (31)$$

and

$$D_j(\theta-1) = \sum_{i\neq j} M_{ij}(\theta-1) \quad (32)$$

This method is analogous to the transition rates approach described in section 3.1 and relies on the usual assumption that the historical total rate coefficients remain constant for the projection period. The formulation for outmigration can be redefined in terms of the probability of outmigration conditional upon survival ($p_i^{o/s}$) as:

$$\tilde{O}_i(\theta) = p_i^{o/s}(\theta-1)p_i^{s}(\theta-1)K^{i*}(t,\theta) \quad (33)$$

where

$$p_i^{o/s}(\theta-1) = O_i(\theta-1)/K^{\varepsilon(i)\sigma} \quad (34)$$

and

$$p_i^{s}(\theta-1) = K^{\varepsilon(i)\sigma}/K^{\varepsilon(i)*} \quad (35)$$

the probability of survival. $K^{\epsilon(i)\sigma}$ represents those persons who survive the period in region i. If the probability of survival in the projection period, $p_i^S(\theta)$, can be estimated exogenously, this can be plugged into the model in place of $p_i^S(\theta-1)$. Survival and locational probabilities have been identified and used for projection by Rees (1977).

The shortcoming of this simple rates methodology is that the observed rates will not remain constant from one period to the next. The observed outmigration and inmigration rates for the historical period 1961-66, and for our so-called 'projection' period, 1966-71, are illustrated in Table 4 and the significant changes can be identified. Changes of the inmigration rate range from +0.4 per thousand for the West Midlands to +19.2 per thousand for East Anglia, whereas the outmigration rates for the North and Scotland both appear to have declined by 0.5 per thousand, while the East Midlands has experienced the largest outmigration rate increase of 10.3 per thousand.

The changes that are evident from the available data may be explained in a variety of ways but the more important issues in relation to projection, is whether an alternative methodology can be developed to produce total outmigration and inmigration projections that are more accurate than those generated using the historical rates model. One alternative methodology may be to incorporate additional known information on net migration in some way. This approach is discussed in section 4.2. A second alternative approach might involve making use of the functional relationship between rates of outmigration and inmigration.

Studies by Miller (1967) of metropolitan area migration in the U.S. and by Cordey-Hayes and Gleave (1975) of city-region migration in Britain, showed that prosperous regions experiencing high rates of aggregate inmigration also tend to experience high rates of aggregate outmigration, while depressed regions exhibit lower rates of out- and inmigration. Cordey-Hayes and Gleave argued that the reasons for this are related to local employment conditions, differential mobility levels and the availability of information on opportunities outside the area. They considered that an area in which there are many job vacancies and low unemployment will have a rapid job turnover because of the freedom of risk in changing job. A high rate of inmigration, an influx of persons with 'more mobile' characteristics, will subsequently generate a counterstream.

Region	Inmigration (per thousand)			Outmigration (per thousand)		
	1961-66	1966-71	Difference	1961-66	1966-71	Difference
North	35.7	42.5	+6.8	50.7	50.2	-0.5
North West	32.1	35.2	+3.1	35.3	41.5	+6.2
Yorkshire and Humberside	39.5	42.2	+2.7	44.9	53.6	+8.7
East Midlands	62.4	70.7	+8.3	47.9	58.2	+10.3
West Midlands	41.9	42.3	+0.4	45.3	50.1	+4.8
East Anglia	95.8	115.0	+19.2	61.6	67.6	+6.0
South East	33.3	37.0	+3.7	34.7	40.1	+5.4
South West	89.7	98.5	+8.8	58.7	64.9	+6.2
Wales	42.2	48.1	+5.9	42.1	45.7	+3.6
Scotland	16.9	23.7	+6.8	32.3	31.8	-0.5

Table 4: <u>Observed Rates of Aggregate Regional Inmigration and Outmigration. 1961-66 and 1966-71</u>

Correlation analysis confirms that the relationship between rates of regional outmigration and inmigration for 1961-66 is significant at the 99% level of confidence, and the equation for the regression line plotted in Figure 3 can be defined as:

$$o_i(\theta-1) = 28.5 + 0.34\, d_i(\theta-1) \qquad (36)$$

Use can be made of equation (36) to generate estimates of outmigration rates based on available inmigration rates information. The same relationship might also be adopted for other time periods, and assumptions of this sort would permit projected outmigration to be estimated from projected inmigration, given the availability of the latter. One of the problems associated with this type of approach is that, although the correlation between the two variables is significant, the explanation of this relationship, in terms of the operation of the labour market turnover process, is much less straightforward. Rogers (1978) points out that there exists a system of structural relationships between variables such as outmigration, inmigration, employment opportunities, unemployment conditions, and age/sex population characteristics, for which equations may be defined simultaneously. The problem then involves identification of 'true' equations in a situation where underlying theory has not yet been established. Rogers' argument rests on the premise that in using a superficial relationship between outmigration and inmigration rates, no attempt is being made at explanation of the structural mechanics involved, and therefore changes in 'hidden' variables may produce unforeseen effects on outmigration or inmigration that the simple equation omits. The estimation of the structrual equations relating demographic and econometric variables, referred to by Rogers as a 'demometric' approach, is used to indicate that the outmigration-inmigration rate correlation does not invalidate traditional push-pull theory in the way that Cordey-Hayes and Gleave intimate. But it is the complexities involved in the definition of the correct set of simultaneous equations as well as in the provision of data to test them out, that are the major obstacles to developing what would become a comprehensive demographic-economic model (including lagged variables) of the type suggested by Ledent (1978), and applied to a multi-regional system. At this preliminary stage, it is intended to observe historical relationships between total migration rates and assume that they will continue into the future. However, the econometric argument is picked up again in section 4.3.

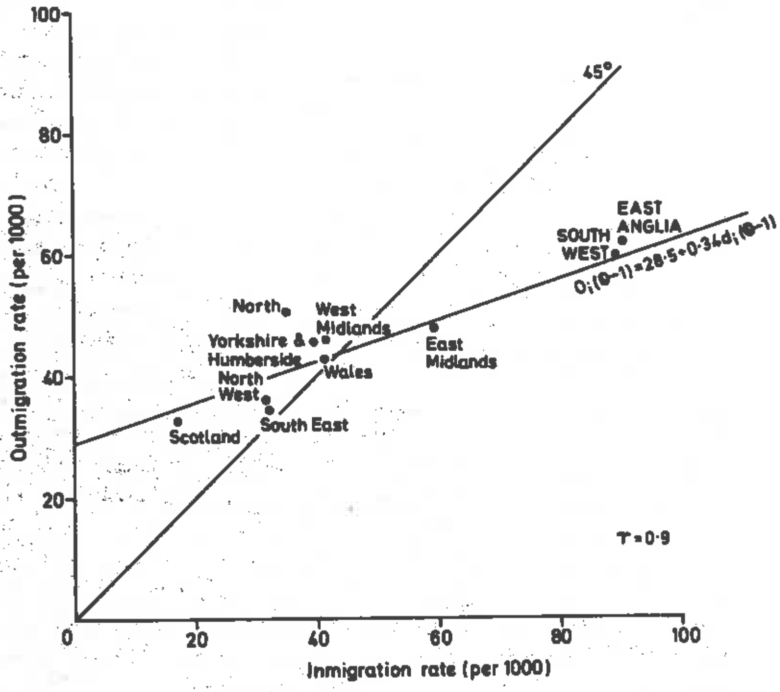


Figure 3 <u>The Relationship Between Regional Outmigration and Inmigration Rates for 1961-66</u>

4.2 A 'Net Migration' Approach to Regional Outmigration and Inmigration

4.2.1 The Derivation of Residual Net Migration Estimates

It is unfortunate that the census, the most reliable and comprehensive data source in this country, only provides researchers with periodic information about migration. The result is analysis based on the five year or one year migration question. However, an indirect method of estimating annual net migration makes use of the Registrar General's annual statistics for births, deaths and mid-year population totals (G.R.O., O.P.C.S., annual). January population totals may be interpolated and a series of net migration figures calculated as the residuals between population change and natural increase or decline for each region. Although the vital statistics involved require adjustment to counter the effects of changes in boundaries, a time series of crude birth rates (b_i), death rates (d_i), natural change rates (ni_i), population changes rates (Δp_i) and net migration rates (nm_i) can be determined. The Registrar General statistics for counties can be aggregated to provide an annual series of net migration rates, 1961 to 1971, for the regions adopted in the study.

The importance of the migration component of population change is evident in each of the graphs of Figure 4. While the birth rate schedules for each region show general decline following the 'turning point' of 1964, and death rates appear to maintain a relative stability over the decade, the net migration schedules exhibit considerable fluctuation vis à vis natural increase. Since the residual migration statistics are estimated net values, the volume of the total migration into or out of each region is concealed. However, the characteristics of the regional net migration time series profiles can be distinguished and theories can be advanced that may explain variation over time. Such theories might involve extensive analysis of regional development and policy during the sixties.

4.2.2 Net Migration Analysis: Preliminary Discussion

Given the uncertainty associated with the net migration estimation and adjustment methods, the relevant concern in this section is whether we can make use of this annual information in preparing projections. There are a number of potential areas of investigation. Firstly, there are questions related to whether long term trends in net migration exist and whether trend extrapolation or curve-fitting techniques would be suitable for projecting net migration. The linear function is the simplest of many available functions that can be fitted using least squares regression.

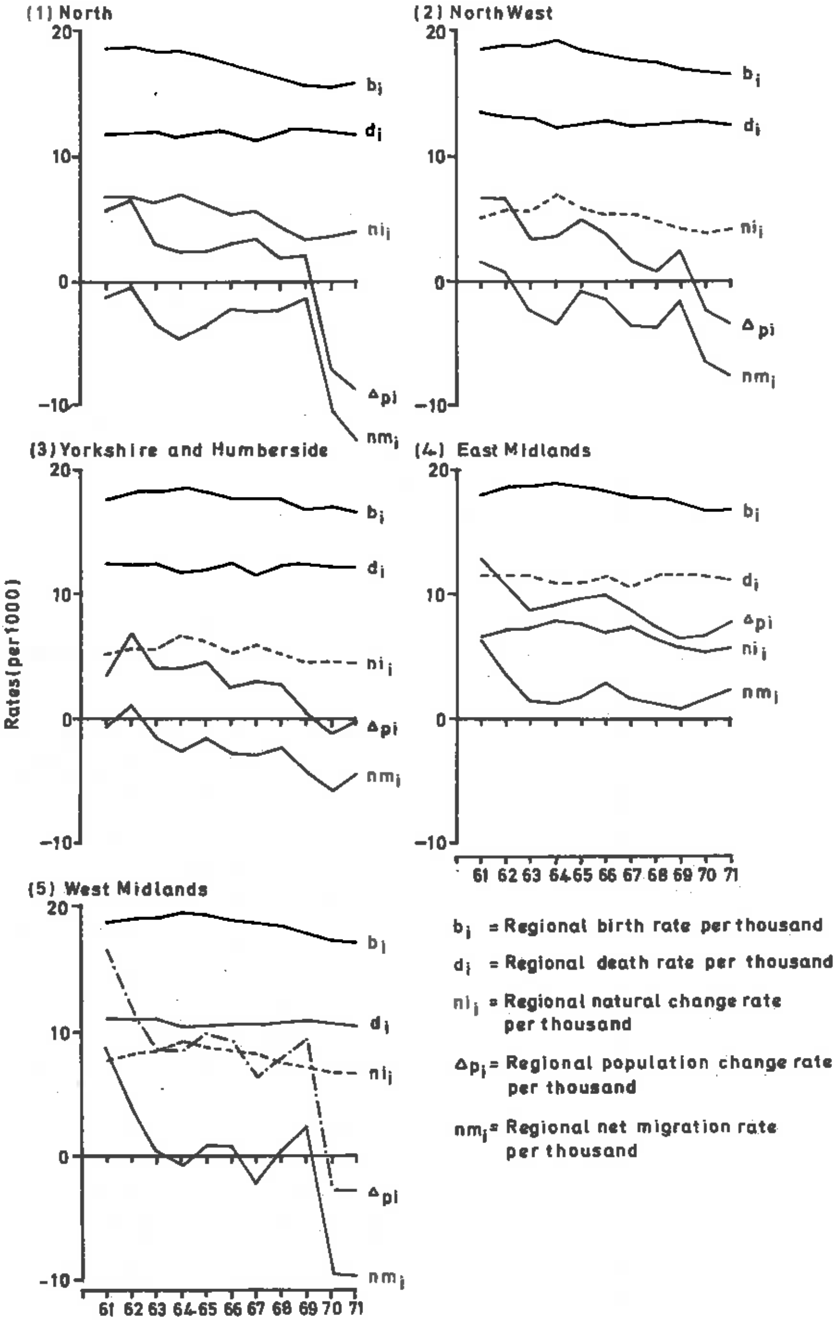


Figure 4: The Components of Regional Population Change, 1961 to 1971

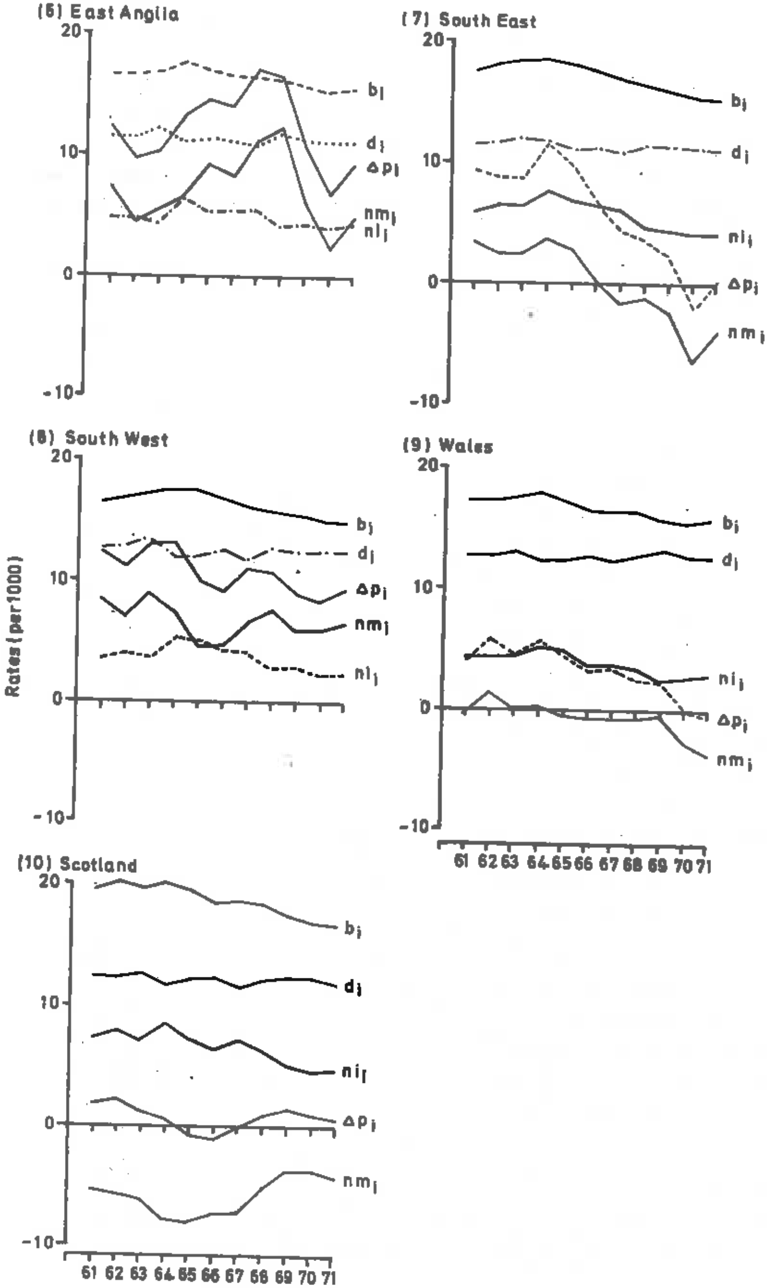


Figure 4: The Components of Regional Population Change, 1961 - 1971

More sophisticated time series techniques have been used to model and forecast populations by Saboia (1974), whose conclusions reflect the necessity to have large numbers of observations, a requirement that rules out similar analysis with the net migration series identified here. The other disadvantage of a time series extrapolation approach is the inability to allow migration projections to respond to changes in policy variables. This leads to a second set of issues concerned with whether historical netmigration fluctuations can be connected with changes in macro-economic variables whose values reflect policy changes from year to year; and subsequently, whether functional relationships could be utilised for producing policy-related projections. One uncertainty in this strategy is whether net migration is the correct variable to be using in this context, since the ultimate objective is the generation of separate projections of regional out- and in-migration. Difficulties involved in establishing consistent time series data for 'policy' variables suggest that a cross-sectional approach related directly to 'census' out-migration and inmigration would perhaps be more advantageous than the indirect net migration approach, although Morrison and Relles (1975) indicate the way in which time series analysis can be conducted in a forecasting context. The method that they adopt and the difficulties involved with time series analysis in relation to regional net migration projection in Britain, are reviewed in section 4.3. A third line of enquiry is one where annual net migration is re-estimated and aggregated to be consistent and comparable with observed 'census' net migration, and where the parameters of relationships between rates of net migration and rates of in- and out-migration may be used to generate predictions.

4.2.3 Net Migration and Total Rate Relationships

Estimates of residual net migration forregions during the census period 1961-66 have been determined by a procedure of consolidation and adjustment of the relevant annual statistics presented in Figure 4. The two definitions of net migration can be reconciled using the framework provided by the population accounts. In Table 5 the census net migration rates defined as:

$$nm_i^c(\theta-1) = \frac{D_i(\theta-1)-O_i(\theta-1)}{K^{i*}(t,\theta-1)} \qquad (37)$$

where K^{i*} refers to a census defined usually resident population at t, the start of the historical period θ-1 (1961-66), are compared with two sets of residual net migration rate estimates defined as:

	NET MIGRATION RATES (per 1,000)		
Region	nm_i^c	nm_i^{R1}	nm_i^{R2}
North	-15.1	-16.3	-13.0
North West	-3.2	-9.1	-5.2
Yorkshire and Humberside	-5.4	-4.0	-4.9
East Midlands	14.5	17.8	12.3
West Midlands	-3.4	3.4	9.3
East Anglia	34.2	35.5	35.1
South East	-1.4	4.4	15.0
South West	30.9	36.3	36.7
Wales	0.09	-3.3	2.1
Scotland	-15.4	-26.9	-34.1

Table 5: Census and Registrar General Residual Net Migration Rates (per thousand) for Regions, 1961-66

$$nm_i^R(\theta-1) = \frac{(K^{*i}(t+T,\theta-1)-K^{i*}(t,\theta-1))-(K^{\beta(i)*}(\theta-1)-K^{*\delta(i)}(\theta-1))}{K^{i*}(t,\theta-1)} \quad (38)$$

where $K^{\beta(i)*}(\theta-1)$ represents total births and $K^{*\delta(i)}(\theta-1)$, total deaths during θ-1. nm_i^{R1}, in which a census defined usually-resident population is employed, is distinguished from nm_i^{R2}, in which an interpolated Registrar General home population is used. The correlation coefficients describing the relationship between nm_i^C and nm_i^{R1} (Figure 5) and nm_i^{R2} respectively, are 0.969 and 0.967, both significant at a 99% confidence level.

The linear regression equation relating census net migration to residual net migration in 1961-66 is:

$$nm_i^C(\theta-1) = 0.5 + 0.81\ nm_i^{R1}(\theta-1) \quad (39)$$

and the differences between the sets of figures for each region are accounted for by the inclusion, in the residual estimate, of international migrations and of migrations of persons who are born and/or who die during the period. The census figure consists entirely of inter-regional exist-survive migrants.

Functional relationships between rates of 'census' net migration and rates of both out- (Figure 6) and in-migration (Figure 7) may also be identified for the regions, and the regression equations can be written as:

$$o_i(\theta-1) = 43.8 + 0.43\ nm_i^C(\theta-1) \quad (40)$$

and

$$d_j(\theta-1) = 43.8 + 1.43\ nm_j^C(\theta-1) \quad (41)$$

Correlation coefficients of 0.762 and 0.968 are both significant at a 99% confidence level. Equations of this type cannot be used to explain the underlying relationships between structural variables associated with the process through which migration takes place, for reasons discussed in the section 4.1, yet they do allow us to establish a connection between estimates of residual net migration and both out- and in-migration, so that in a projection context, the following sequence of general functional relationships ($y = f(x)$):

(1) $$\tilde{nm}_i^C(\theta) = f(\tilde{nm}_i^{R1}(\theta)) \quad (42)$$

(2) $$\tilde{d}_i(\theta) = f(\tilde{nm}_i^C(\theta)) \quad (43)$$

and then, either

(3) $$\tilde{o}_i(\theta) = f(\tilde{nm}_i^C(\theta)) \quad (44)$$

or

(4) $$\tilde{o}_i(\theta) = f(\tilde{d}_i(\theta)) \quad (45)$$

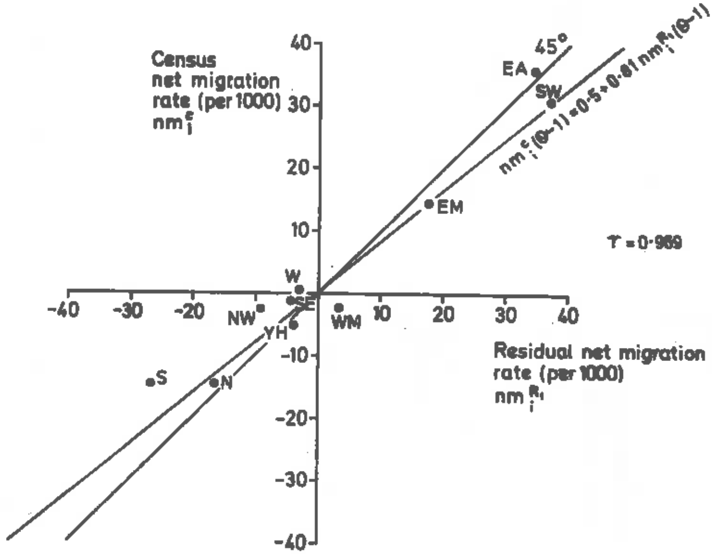


Figure 5: A Comparison of Census (nm_i^c) and Residual (nm_i^{R1}) Net Migration Rates, 1961-66

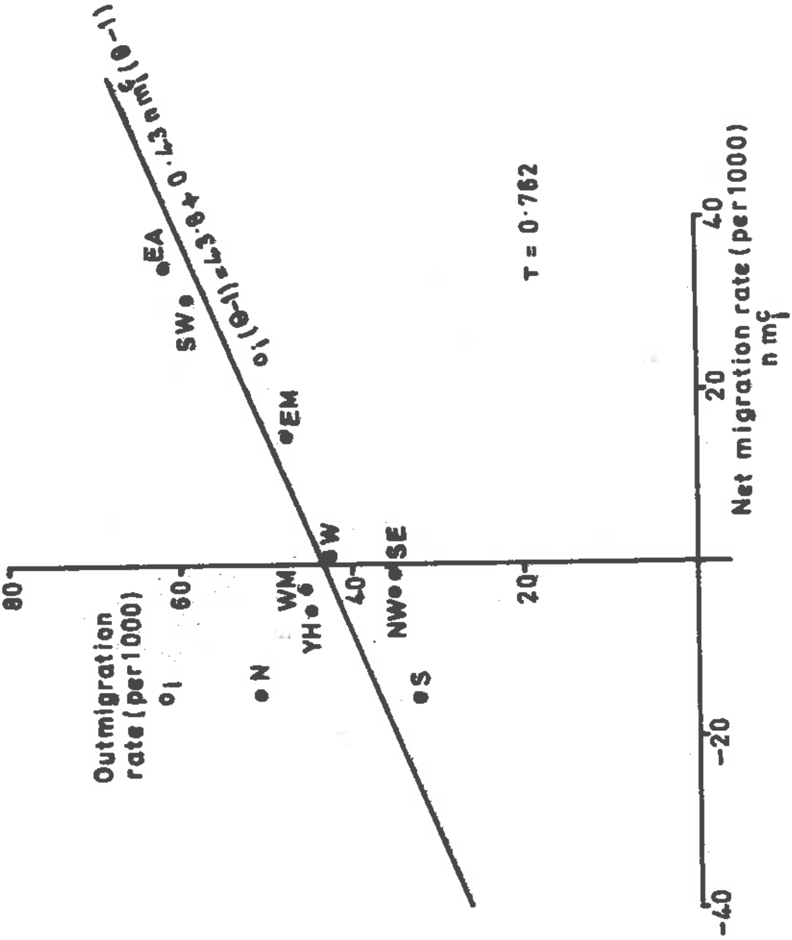


Figure 6: <u>The Relationship Between Regional Outmigration and Net Migration Rates 1961-66</u>

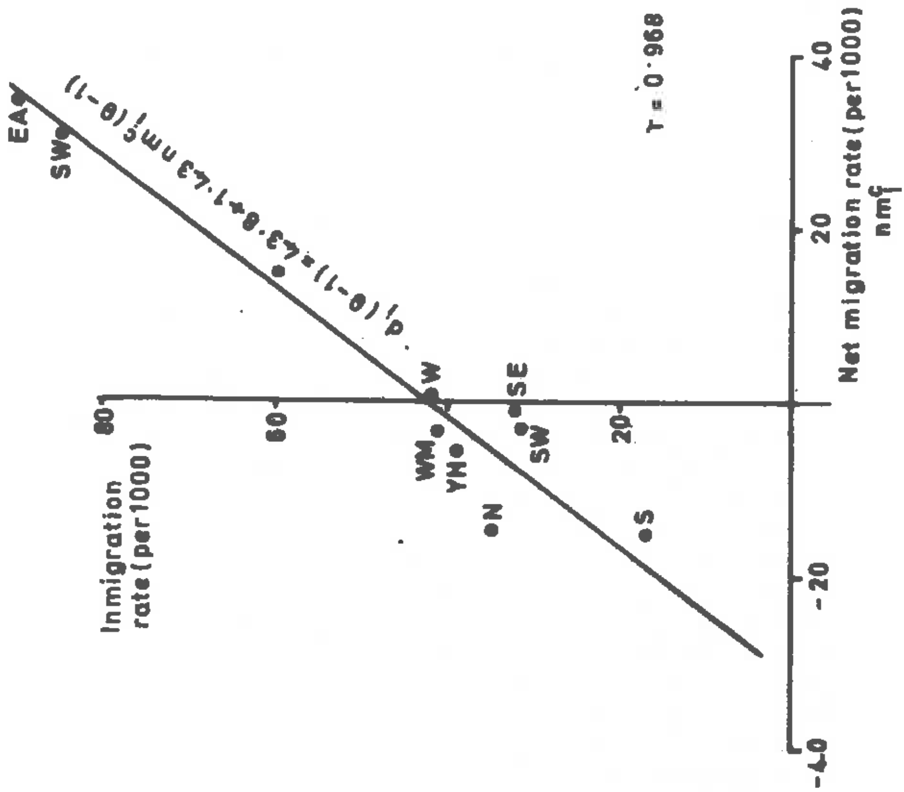


Figure 7: <u>The Relationship Between Regional Inmigration and Net Migration Rates 1961–66</u>

can be followed and the rate projections converted to out- and in-migration flows by multiplication with a base population total. The aggregated residual net migration statistics can therefore be used to determine total outmigration and total inmigration through regression analysis in which observed coefficients are assumed to remain constant from one period to the next.

4.2.4 $\tilde{O}_i$ and $\tilde{D}_j$ Determined from Net Migration Estimates and Projections

In order to determine $\tilde{O}_i$ and $\tilde{D}_j$ in the projection period, θ, we can make use of the observed annual net migration information in the latter half of the decade (Figure 4), and estimate outmigration and inmigration from residual net migration figures that have been aggregated for the years concerned (for 1966-71). This approach involves making use of the rate relationships discussed previously, the initial stage of which is the estimation of census net migration from the estimated residual net migration figures presented in Table 6, using the parameters of equation (39). However, the residual net migration information would not be available in a 'true' projection context so it is necessary to consider the potential for projecting net migration itself. One advantage of using annual residual migration is that some indications of changes in net balance can be implied from the time series data that are not observable from the five year or one year cross-sectional estimates of net migration obtained from the census. The issue then becomes one of whether we can believe that the trends which the time series statistics reflect, will continue into the future. Time series linear regressions on the net migration rates for each region for the first six years of the decade (Figure 8) allow extrapolated average values to be projected for 1968 which can be multiplied up to give five year rates comparable with the estimated rates for 1966-71 (Table 6). The correlation coefficient for the estimated and projected rates is 0.95 which is significant at the 99% confidence level.

Examination of Figure 8 indicates some cases where the extrapolated regression line fits the observed data points for the second half of the decade quite closely (the North West and Yorkshire and Humberside), but in general the annual figures from the trend extrapolations are not particularly competent projections, because of the 'surprising' trend changes in certain regions, although in some cases the deviations of observed points on either side of the trend line cancel one another out

Region	Estimated net migration	Rate (per 1000) $nm_i^R(\theta)$	Projected net migration	Rate* (per 1000) $\tilde{nm}_i^R(\theta)$
North	-90,443	-27.3	-81,412	-24.6
North West	-116,996	-17.6	-132,212	-19.9
Yorkshire and Humberside	-98,451	-22.9	-90,545	-21.13
East Midlands	17,176	4.5	-4,410	-1.15
West Midlands	-62,179	-12.5	-110,956	-23.3
East Anglia	65,528	42.1	72,741	49.7
South East	-118,925	-7.1	60,362	3.73
South West	110,427	27.2	58,529	16.8
Wales	-14,601	-5.4	-13,069	-4.95
Scotland	-163,622	-31.2	-244,983	-47.3

*$\tilde{nm}_i^R(1966\text{-}71) = 5(\tilde{nm}_i^R(1968))$

Table 6: Estimated and Projected Residual Net Migration, Aggregated for 1966-71.

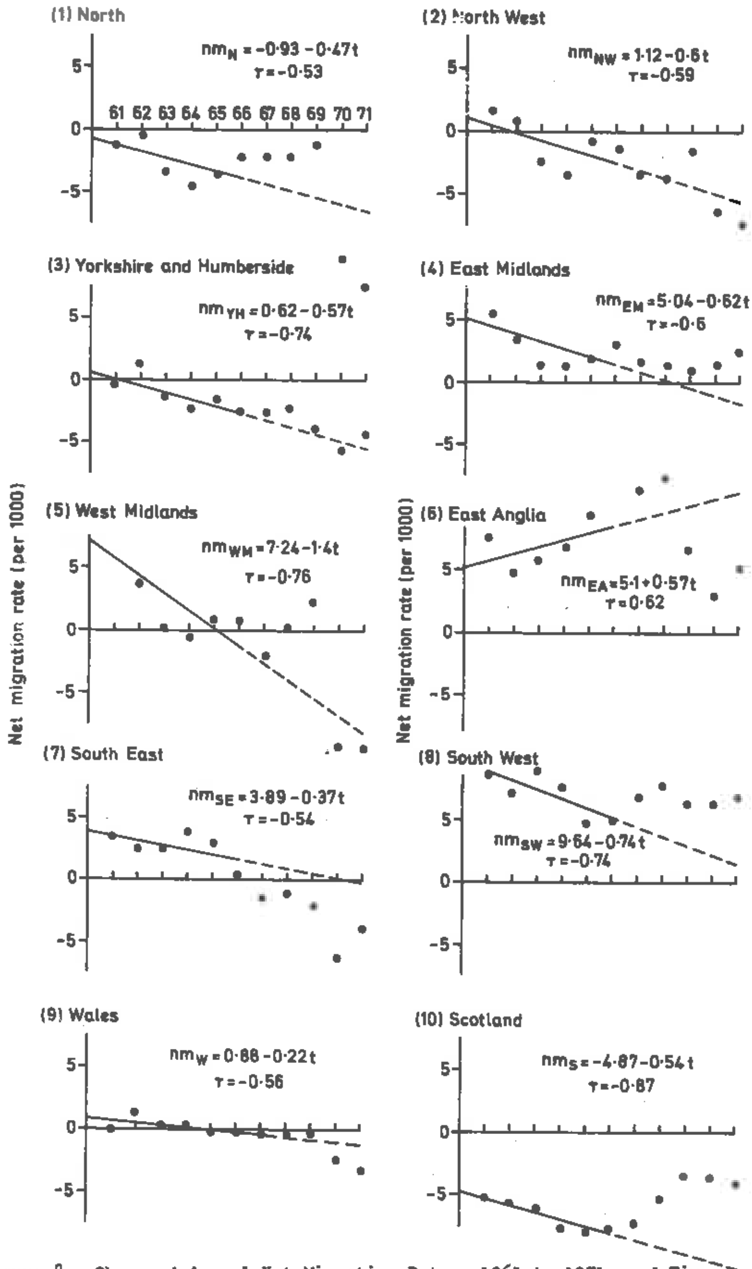


Figure 8: Observed Annual Net Migration Rates, 1961 to 1971, and Time Series Regression Relationships Based on 1961 to 1966 Annual Data, with Extrapolation to 1971

to provide an average projection which is relatively accurate. The trend lines and therefore the projected values for net migration are most inaccurate for the regions of the South East, East Midlands, the South West and Scotland. In the South East, an observed net loss is underpredicted as a small net gain. The projections for the East Midlands indicate increasing net losses towards 1971 with an average net loss overall, but in fact the observed values are positive for the whole of the decade, with a change towards increasing net gain by 1971. Net migration gains in the South West are considerably underpredicted, while net migration losses for Scotland are severely overpredicted. The results produced by a time series approach of this type are therefore subject to considerable criticism.

Nevertheless, they can be applied to generate projected values for outmigration and inmigration rates using the rate relationships defined in equations (42) to (45). In Table 7, two projected inmigration rates are defined on the basis of equations (42) and (43). $\tilde{d}_j(1)$ is the projected inmigration rate for 1966-71 based on the observed residual net migration estimates, while $\tilde{d}_j(2)$ is the projected inmigration rate based on the net migration rate projections dervied by the time series analysis. $\tilde{o}_i(1)$ and $\tilde{o}_i(2)$ are the outmigration rate projections based on the functional relationship with projected inmigration rates $\tilde{d}_j(1)$ and $\tilde{d}_j(2)$ respectively. The sequence of estimation:

$$\begin{matrix}\text{residual}\\ \text{net migration}\end{matrix} \rightarrow \begin{matrix}\text{census}\\ \text{net migration}\end{matrix} \rightarrow \text{Inmigration} \rightarrow \text{Outmigration}$$

was adopted in order to utilise the mostsignificant relationships between the migration rate variables.

The rate projections for inmigration and outmigration generated using this net migration extrapolation methodology are less accurate than the rate projections based on the assumption of no change (Table 4). In most cases the difference between the projected rates and the observed rates (Table 7) indicates substantial underprojection (except for the South East), although the outmigration rate values seem to be more accurate than the inmigration rate projections. Although the projected results are somewhat disappointing in terms of accuracy and might perhaps be improved by the adoption of a more sophisticated time series extrapolation method or by an alteration of the sequence of functional relationships that is followed, one important function of the approach that should be emphasised here is that it serves to illustrate the way in which net migration data

REGION	Inmigration Rates				Outmigration Rates			
	$\tilde{d}_j(1)$	Difference from observed d_j	$\tilde{d}_j(2)$	Difference from observed d_j	$\tilde{o}_i(1)$	Difference from observed o_i	$\tilde{o}_i(2)$	Difference from observed o_i
North	12.9	-29.6	19.6	-22.9	32.9	-17.3	35.2	-15.0
North West	24.1	-11.1	21.5	-13.7	36.7	-4.8	35.8	-5.7
Yorkshire and Humberside	18.1	-24.1	20.1	-22.1	34.7	-18.9	35.3	-18.3
East Midlands	49.7	-21.0	43.2	-27.5	45.4	-12.8	43.2	-15.0
West Midlands	30.1	-12.2	17.5	-24.8	38.7	-11.4	34.5	-15.6
East Anglia	93.3	-21.7	102.1	-12.9	60.2	-7.4	63.2	-4.4
South East	36.2	-0.8	48.8	11.8	40.8	0.7	45.1	5.0
South West	75.9	-22.6	63.9	-34.6	54.3	-10.6	50.2	-14.7
Wales	38.2	-9.9	38.8	-9.3	41.5	-4.2	41.7	-4.0
Scotland	8.5	-15.2	5.0 *	-18.7	31.4	-0.4	30.2	-1.6

$\tilde{d}_j(1)$, $\tilde{o}_i(1)$ are rates based on estimated net migration, 1966-71

$\tilde{d}_j(2)$, $\tilde{o}_i(2)$ are rates based on projected net migration, 1966-71

* A minimum rate of 5 per thousand is assumed for Scotland (the net rate extrapolation suggests negative inmigration!)

Table 7: Projected In- and Out-migration Rates (per thousand), 1966-71, Based on Net Migration Data; and Differences from the Observed Values

can be incorporated into the projection exercise. The next section investigates some alternative methods of projecting out- and in-migration that incorporate other types of information.

4.3 Alternative Approaches to Regional Outmigration and Inmigration Projection

4.3.1 A Review of Deterministic Models

In contrast to the rates approaches outlined in the previous two subsections, an alternative strategy for projecting outmigration and inmigration makes use of deterministic models in which macro-economic variables that stimulate the transfer of migrants between regions, are identified. Existing econometric models that have been calibrated and tested with historical data, differ in the emphasis placed on 'push' factors as opposed to 'pull' factors acting as stimulants to mobility. A number of studies in the U.S. have been reviewed by Greenwood (1975c). Although increasing job vacancies or profitable income differentials are proven inducements for migration to favourable destinations, there is much less concensus about the nature and role of 'push' factors. Some research (Miller, 1973) suggests that the relationship between outmigration and economic conditions at the origin is important, yet many studies, including that of Lowry (1966), emphasise the insensitivity of outmigration to the economic characteristics of the origin zone.

Certainly it has proved possible to identify and measure variables that reflect the economic 'climate' of the destination zone and to recognise that these conditions both determine and are determined by the numbers of responsive inmigrants. Growth in employment opportunities, for example, will tend to attract inmigrants to fill vacancies not immediately taken up by the indigenous unemployed. Subsequently, lagged effects on both employment and migration will result from the operation of the employment-migration multiplier. Ideally it would be useful to construct, within an inmigration projection model, a method of making some allowance for changes in leading regional indicators, so that sets of inmigration projections could be produced that are based on alternative policies. The corollary of this would then be a method of projection catering for the indirect effects of the migration multiplier. The paucity of data however, has been the main constraint on research in this area.

One of the major difficulties associated with modelling aggregate immigration as a function of a set of explanatory variables is the variety of purposes for which people migrate: migration heterogeneity. The influence of housing factors, environmental factors, education, retirement, employment and incomes, are all either more or less important depending upon the mix of migrant types partially associated with the spatial scale at which migration is taking place. A suitable regression equation for age-group(s) inmigration might be defined as:

$$D_j^s = \sum_i M_{ij}^s = a^s + \sum_n b_n^s V_{jn}^s \qquad (46)$$

where a^s is the age-group regression intercept, and b_n^s is the n'th regression coefficient related to the n'th variable describing the destination, V_{jn}^s, which *influences* inmigration in this age group. Regression analysis based on further disaggregation by sex or socio-economic group would expose variables related in some unique way with the migrants concerned.

Although some age/sex disaggregated data is available from the census which would allow this type of analysis to be conducted, we rely here on the generalisation that factors vary in influence according to the distance or scale at which migration is occurring. There is no denying the severity of the aggregation problem, but interregional movement will usually be influenced by economic factors such as employment growth or unemployment change, which imply a migration involving a change of job as well as a change of home, while other factors such as the availability of housing, education or social services are of more importance to intra-regional and especially to intra-urban migrants. In the last instance, a local migration involving a change of house is not necessarily associated with a change of job. Baxter and Williams (1978) have distinguished the more important factors influencing long and short distance migration respectively, and on the basis of this dichotomy, a crude aggregate model might be formulated that determines (longer distance) regional migration according to a set of prospective macro-economic variables.

Donovan (1971) and Morrison and Relles (1975) have both conducted this type of analysis using linear regression techniques, and a brief review of their work is perhaps appropriate at this point. Donovan's multiple regression analysis involved relating regional inmigration totals (excluding Scotland) to combinations of the following regional variables for 1960-61 and 1961-66:

		1960-61	1961-66
Outmigration	(O_i)	✓	✓
Average schedule E Earnings, 1959-60	$(\overline{Y}_j)$	✓	✓
Male Earnings relative to National Average, 1964-65	$(\overline{Y}'_j)$		✓
Employment change, 1948-58	(ΔE_j)	✓	✓
Employment change, 1961-66	$(\Delta E'_j)$		✓
Mean number of employees, 1966	$(\overline{E}_j)$		✓
Mean estimated annual unemployment, 1954-61	$(\overline{U}_j)$	✓	✓
Manufacturing Service Index, 1961	(S_j)		✓

Outmigration levels were included as independent, predictor variables for inmigration because outmigrants can be seen as releasing vacancies which are then taken up by inmigrants. Donovan produced 'best-fit' equations for 1960-61 as:

$$D_j(1960\text{-}61) = a+b\Delta E_j \quad (r = 0.76) \qquad (47)$$

and for 1961-66 as:

$$D_j(1961\text{-}66) = a+b_1\overline{Y}'_j+b_2\overline{E}_j+b_3\overline{U}_j+b_4S_j \quad (r = 0.98) \qquad (48)$$

in which the employment change variable generated the most significant fit for the single year period, while the optimum combination of variables for 1961-66 consisted of indices of earnings, employment, unemployment and manufacturing services. No explanation was offered for why this particular combination proved satisfactory and effects relating to the extent of multicollinearity were not investigated beyond a presentation of the matrix of correlation coefficients.

This type of cross-sectional analysis was also undertaken by Morrison and Relles using the Continuous Work History Sample (C.W.H.S.) organised by the U.S. Social Security Administration as the data base for estimating migration flows between U.S. cities. This sample of employed persons, available on an annual basis, is particularly suited to analysis of labour force migration. In a series of historical tests for annual periods 1961 to 1965, the variables (identified as important) were the change in agricultural employment (ΔE_j), the outmigration rate, the percentage of the 1960 population that resided outside the region in 1955, the percentage of the 1960 labour force between the ages of 18 and 34, and the percentage of the region's labour force in professional occupations. The latter three variables were assumed to index the differential presence of chronic movers (Morrison, 1971), a highly migratory age group, and a highly migratory occupation group respectively.

The results indicated that for any one year, employment growth and inmigration rate were significantly related, but the regression coefficients showed fluctuation from year to year. On the assumption of distinct trends in migration over time, lagged in- and out-migration rates (defined as the average of the rates for the previous two years) were shown to be effective independent variables for explaining the current inmigration rate, and these substituted for the structural variables designed to measure it initially. The best fit model that Morrison and Relles tested for annual periods, (θ)*, may be rewritten as:

$$d_j(\theta) = a+b_1\Delta E_j(\theta)+b_2o_j'(\theta)+b_3d_j'(\theta) \qquad (49)$$

where

$$o_j'(\theta) = \frac{o_j(\theta-1)+o_j(\theta-1)}{2} \qquad (50)$$

and

$$d_j'(\theta) = \frac{d_j(\theta-1)+d_j(\theta-1)}{2} \qquad (51)$$

in which the values of the dependent and independent variables were log-transformed. Correlation coefficients varied between 0.78 and 0.9 for different years but all were significant at the 99% level of confidence.

Migration projection using deterministic models of this form rely on two factors in particular. Firstly it is necessary to have projected values of the independent variables themselves, and secondly, it is essential that the coefficients calibrated for historical data sets are suitable for application in a projection period. Fluctuation in the relationship between employment change and inmigration rate over time has already been mentioned as having been identified by Morrison and Relles, who propose a forecasting model that makes some allowance for these short term changes.

*The time period labels used in the review of Morrison and Relles' work represent calendar year periods.

It is in the forecasting context that these authors make fullest use of the time series data. They show that, for any one region, residuals from the inmigration time series trend line behave in a fashion consistent with the residuals from the trend in employment. Inmigration rate residuals ($\dot{d}_j$) in any one year can therefore be predicted as a function of employment residuals ($\dot{E}_j$) using:

$$\dot{d}_j(\theta+1) = a+b\dot{E}_j(\theta+1) \qquad (52)$$

where the parameters have been calculated using the residuals data for all regions and all years prior to $\theta+1$, and the employment residuals $\dot{E}_j(\theta+1)$ have been estimated as the difference between the known employment total for $\theta+1$ and the total extrapolated on the basis of the time series employment data. The inmigration residuals can then be combined with the region's extrapolated long term inmigration (d_j^*) value to yield a projected inmigration rate for the period as:

$$\tilde{d}_j(\theta+1) = a+bd_j^*(\theta+1)+\dot{d}_j(\theta+1) \qquad (53)$$

The inmigration rate forecast for the following future period ($\theta+2$) would include the most-recent rate prediction and would exclude the least-recent value so that the coefficient would change from year to year as the forecasting proceeds. The method is dependent, therefore, on the availability of employment statistics in the projection period.

The approach that Morrison and Relles adopt in their forecasting analysis confirms that the projected rate of inmigration is, first and foremost, dependent upon its historical rate predecessors. However, the model that they develop is only feasible because of the availability of consistent time series data for outmigration and inmigration as well as for employment. The method is not directly applicable in a British regional context primarily because of the lack of a consistent time series of employment data, and also because of the lack of a migration time series other than the residual net migration series, and it becomes increasingly evident why first attempts at projection are based on simple rates models of the type outlined in section 4.1, making use of the only published inmigration data that is available.

The major criticism of the rates model rests with the long-standing argument that historical rates will not necessarily remain constant in the projection period. Morrison and Relles counter this by using the extrapolated long term trend rate for inmigration, and by introducing an employment variable to cater for short term fluctuations. Although the lack of time series employment data prevents us from adopting this approach directly with inmigration for regions in Britain, this does not mean to say that explanatory variables should be excluded altogether from a projection model of migration. If we wish to work towards the construction of policy sensitive models, then future migration should be regarded as a function of the extrapolated system conditions in which migration is seen to occur, as well as future 'push' and 'pull' regional characteristics. These regional characteristics can be defined themselves according to future economic and social conditions and policy variables.

4.3.2 Time Series Net Migration and Regional Indicators

The next relevant question is whether we can make any headway at all towards investigating the way in which migration fluctuates over time in relation to regional factors. One type of broad approach, similar in some ways to that developed by Morrison and Relles, would require annual time series data on specific regional indicators and would involve an attempt to relate this information to the annual estimates of residual net migration that have been discussed earlier. The identification of certain common features of the trend schedules over time might suggest the prospective shape of the net migration trajectory in the future, under certain policy assumptions. Some measure of immigration or outmigration might then be obtained from the net migration projection using the relationships defined in section 4.2. A preliminary investigation of the fluctuation in certain selected indicators has been undertaken, but the results provide little evidence to justify the acceptance of this type of approach, and there are a number of problems, besides the aggregation issue, that obstruct interpretation.

Three indicators have been specified on the basis of available regional statistics; employment, unemployment and housing construction. Statistics for the first variable, employment, are published in the Department of Employment Gazette and summarised in the Historical Abstract of Labour Statistics (Department of Employment and Productivity, 1971). These statistics refer to the number of employees (employed and unemployed) at each mid-year, and the average annual total of registered unemployed persons. The difference between these two counts for each region provides a mid-year measure of the employees in employment in each region. The boundaries of the regions to which these figures apply are the Standard Region boundaries as defined in the 1966 Census, and the differences between this system and our own regional system have been outlined elsewhere (Stillwell, 1979). The figures for the early 60's are not included for the Standard Regions of Yorkshire and Humberside, the East Midlands, the South East, and East Anglia, because of the inconsistency of the information resulting from changed administrative boundaries in 1965. Since a new method of calculating numbers of employees was introduced in 1965, the previous year is used as the base year (1964 = 100) for an annual, regional employment index ($EI_i(\theta)$) defined as:

$$EI_i(\theta) = 100 \left[\frac{E_i(\theta)-\overline{U}_i(\theta)}{E_i(1964)-\overline{U}_i(1964)} \right] \qquad (54)$$

where $E_i(\theta)$ is the number of employees (employed and unemployed) in region i at the mid-point of year θ*,

and $\overline{U}_i(\theta)$ is the annual average total of registered unemployed persons in region i.

The values of the index are plotted for each region in Figure 9. Pre-1964 values are based on the original method of calculating numbers of employees while post-1964 values are based on the new method. The base year of the index for the regions of Yorkshire and Humberside, the East Midlands, the South East and East Anglia, is 1965.

The second regional indicator selected was unemployment. The unemployment percentages ($\overline{u}_i$) plotted in Figure 10 for each of the Standard Regions are the average annual rates that are published in the Abstract of Regional Statistics (Central Statistical Office, 1972). In the case of unemployment figures up to the year 1965, the same rate values are published for Yorkshire and Humberside and the East Midlands, and for the South East and East Anglia. These figures have been included on the graphs presented in Figure 10.

The third regional indicator selected represents a measure of housing construction. Statistics of the number of dwelling completions in the public and private sectors are published in the Abstract of Regional Statistics, and an annual regional index of new dwellings ($NDI_i(\theta)$) can therefore be defined as:

$$NDI_i(\theta) = 100 \left[\frac{ND_i(\theta)}{ND_i(1966)} \right] \qquad (55)$$

where $ND_i(\theta)$ is the total number of dwellings completed in year θ in region i and θ refers again to a calendar year period.

Annual values of the index have been plotted for the Standard Regions in Figure 11.

Do these three indicators contribute in any way to the explanation of the variation in the net migration rate schedules that were illustrated in Figure 4?

*θ in this case is again used as a single year period label rather than a five year period label.

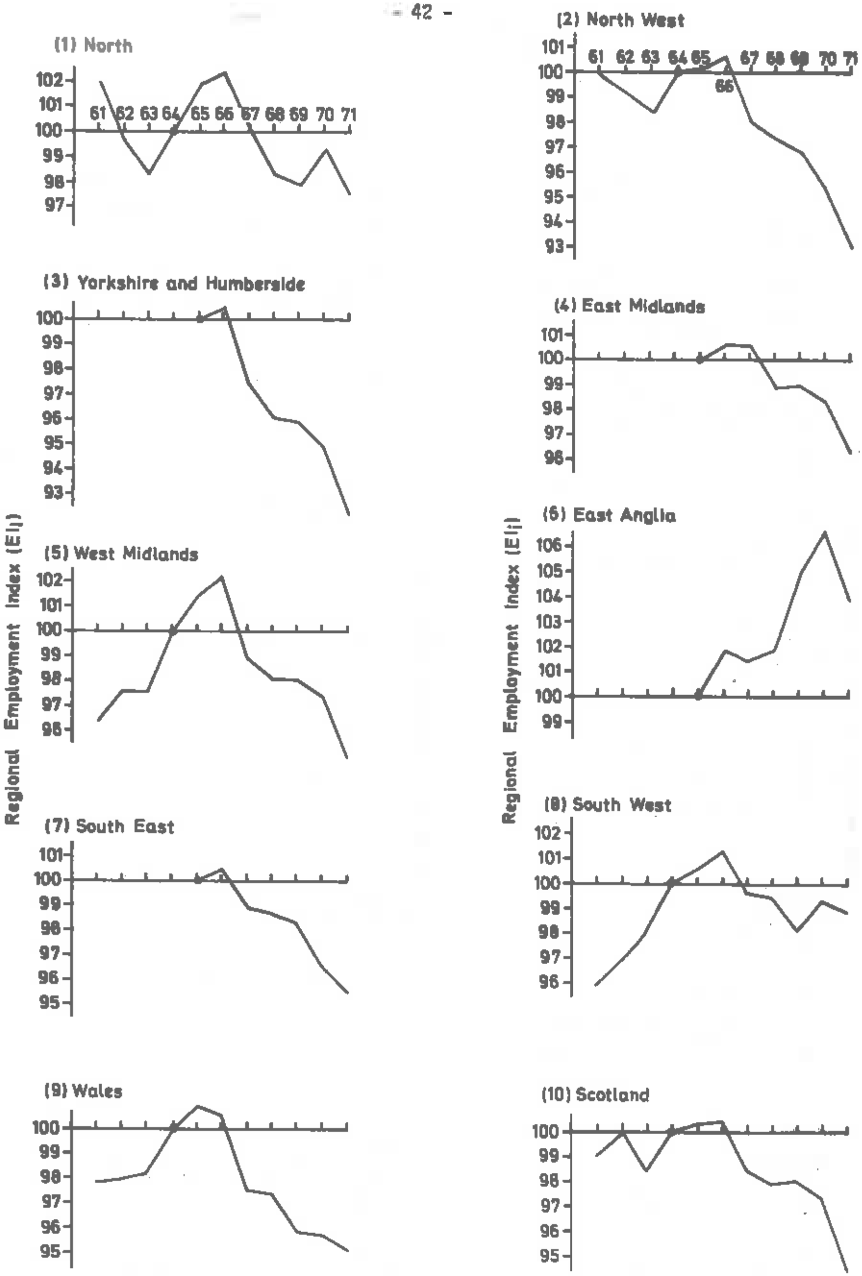


Figure 9: Mid-year Regional Employment Index: Regional Schedules, 1961 to 1971
(1964 Index = 100, except for YH, EM, SE, and EA, where 1965 Index = 100)

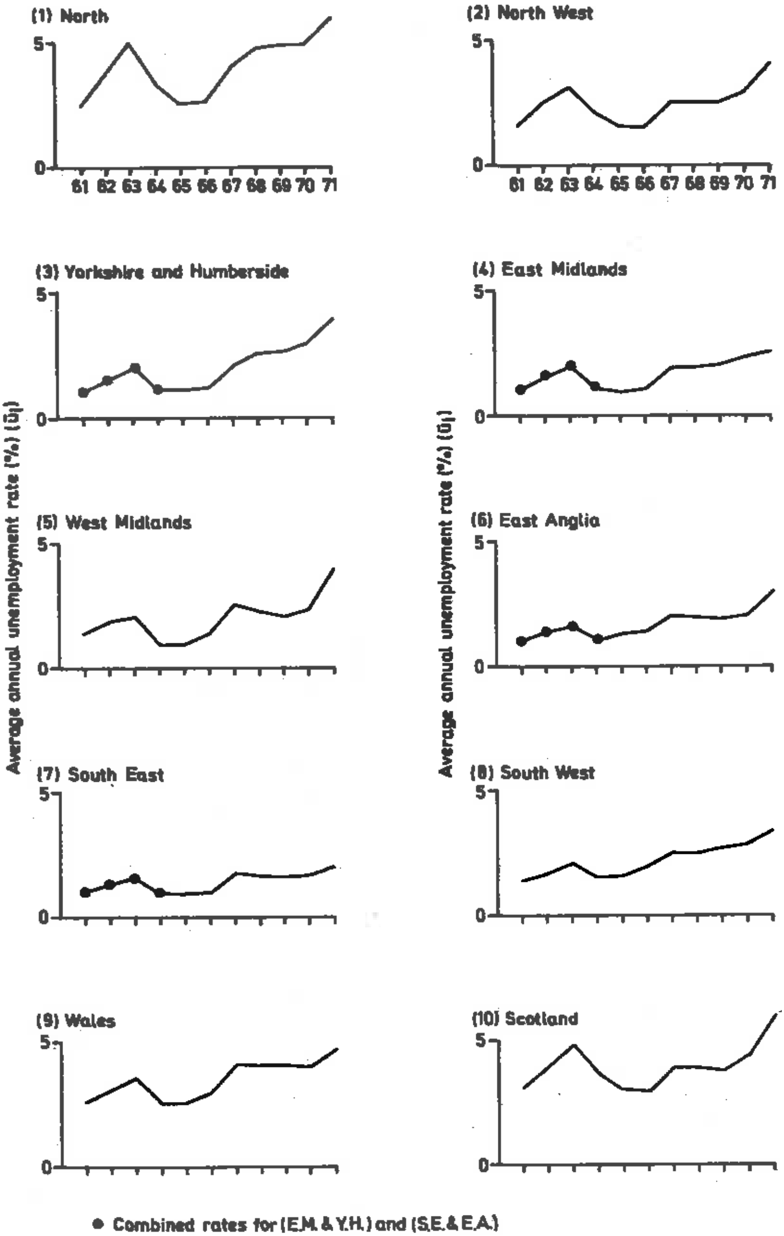


Figure 10: Average Annual Unemployment Rates in % Terms: Regional Schedules, 1961 to 1971

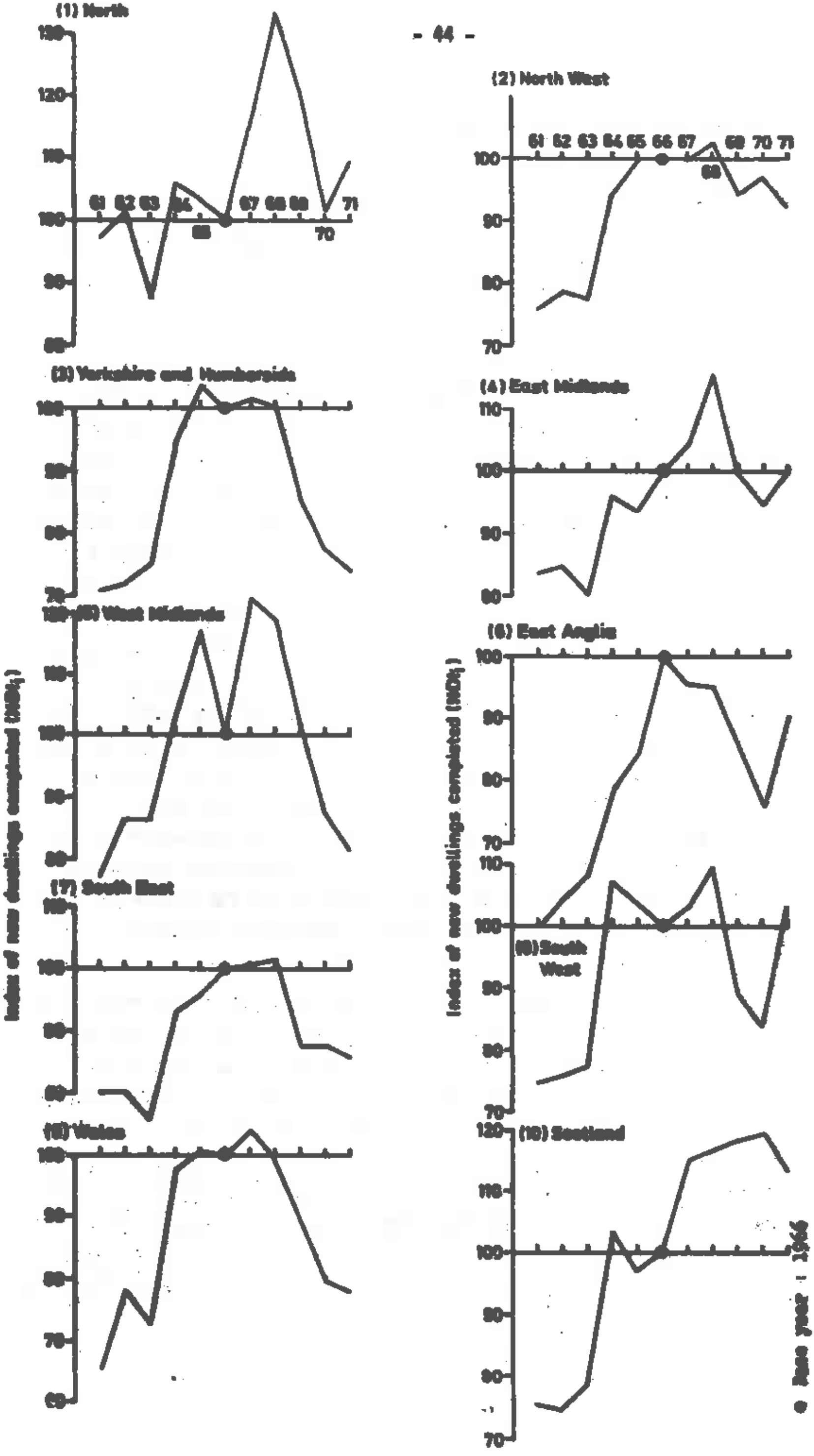


Figure 11: Index of New Dwellings Completed (Public and Private Sectors), Regional Schedules, 1961 to 1971

Regional net migration and employment trends

The employment index that has been defined in equation (54) shows a general turnabout in the middle of the decade with marked declines in employment towards the end of the decade in every region except East Anglia. Certainly projection of the numbers of employees in the late sixties based on trends in employment up to 1966 would clearly be misleading for regions where consistent data is available.

The hypothesis that is relevant in this particular context is that employment is a functional determinant of net migration. In other words, migration is assumed to be employment-led. It might therefore be expected that high regional employment totals would be associated with lower regional net outmigration or higher net immigration, whereas low regional employment would tend to be associated with higher net outmigration or lower net immigration. The relationship is confused because increasing net migration gain might be due to higher net immigration or lower outmigration levels, both of which are affected in different ways by employment considerations. The identification of a functional relationship between annual net migration and employment is not generally evident from the net migration rate and employment index schedules presented in Figures 4 and 9. The schedules for the Northern region and for Scotland have been redrawn in Figure 12 and 13 by way of example, to illustrate that while certain changes in net migration in the former region can be interpreted in terms of the employment trends (lower employment associated with higher net outmigration towards the end of the decade), the trends in net migration and employment in Scotland appear contradictory, with lower levels of employment occurring at the same time as reductions in net outmigration.

The validity of any interpretation of the relationship between these particular variables is a matter for conjecture. There are problems involved in interpretation due to the existence of time lags in the reaction of migration to employment change, and due to the occurrence of boundary changes in the system of interest. The calculation of annual employment change rates would be a more suitable regional indicator than annual employment levels, although the features of employment change schedules that can be implied from the employment index graphs, suggest that general relationships are again difficult to identify. Further research in this area is required, before any definite conclusions about the value of the approach can be drawn.

Regional net migration and unemployment rates

Most of the schedules presented in Figure 10 show a general trend pattern of increasing regional unemployment to 1963, a period of lower unemployment rates in the mid-sixties and then increasing rates in the years through to 1971. The type of relationship that might be expected to occur between unemployment and net migration is one in which net outmigration appears relatively low or net inmigration appears relatively high when unemployment rates are low. Conversely, when unemployment is relatively high, net outmigration is relatively high or net inmigration is relatively low.

An interpretation of the relationship between unemployment and net migration might be possible for a region like the North West (Figure 14), where years of high unemployment tend to be associated with years of higher net outmigration, but the trends for the South West (Figure 15) for example again present contradictions. Complexities associated with time lags should also be taken into account.

Regional net migration and dwelling completions

The argument here is that migration is dependent upon the number of dwellings that become available in a region. Consequently a relatively high number of dwellings would mean relatively high net inmigration or relatively low net outmigration. As with employment and unemployment, the relationship is also dependent upon the situations in other regions. Nevertheless, the relatively high number of dwellings completions in the mid-1960's in the West Midlands (Figure 16) might have been partially responsible for the temporary slowdown in the trend towards increased net outmigration. On the other hand, dwelling completions in the South East (Figure 17) appear to have had little effect on the net migration series.

Undoubtedly the profiles of the respective indices reflect trends in the regional economies, but these trends are not necessarily directly related to the trends in net migration that occur over the same period. Some instances where relationships might prove valid have been mentioned together with some examples of anomalies. The problems involved in interpretation and understanding the reasons for particular changes are considerable and there is no further involvement in the time series net migration analysis in this section. We turn instead to an approach which makes use of census employment data and which deals directly with inmigration projection.

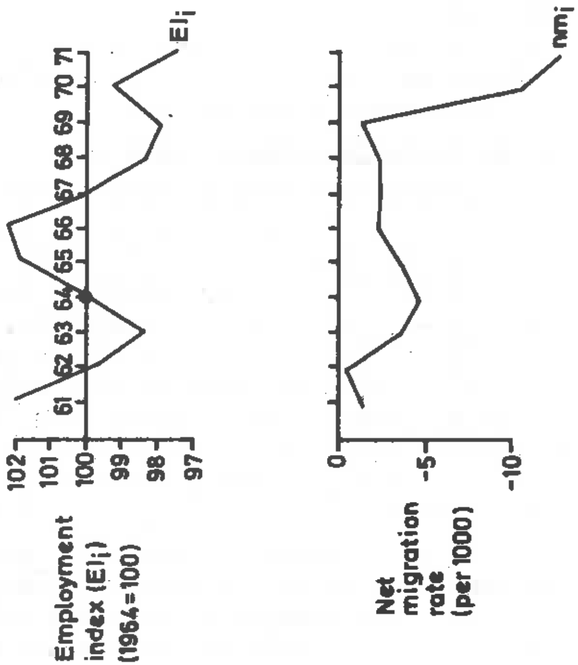


Figure 12: <u>Employment Index and Net Migration Rate Trends for the Northern Region 1961-71</u>

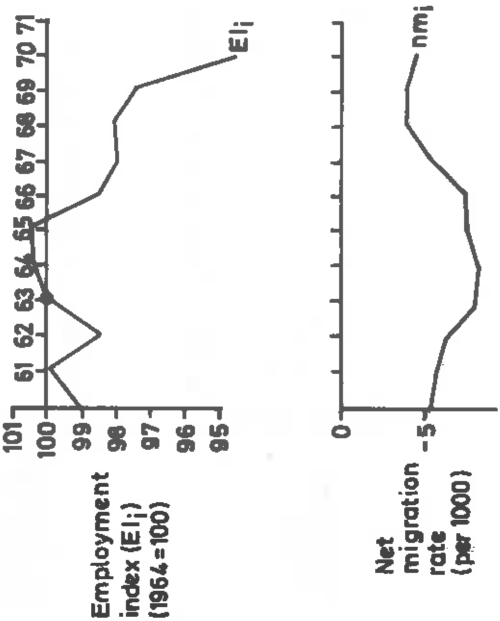


Figure 13: Employment Index and Net Migration Rate Trends for Scotland, 1961-71

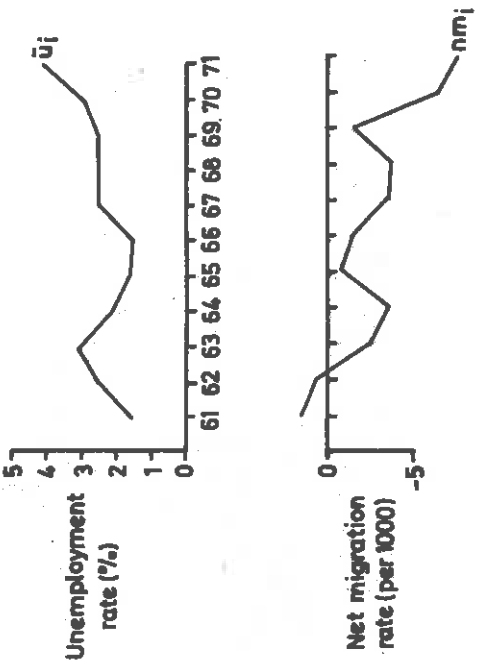


Figure 14: <u>Unemployment Rate and Net Migration Rate Trends for the North West, 1961-1971</u>

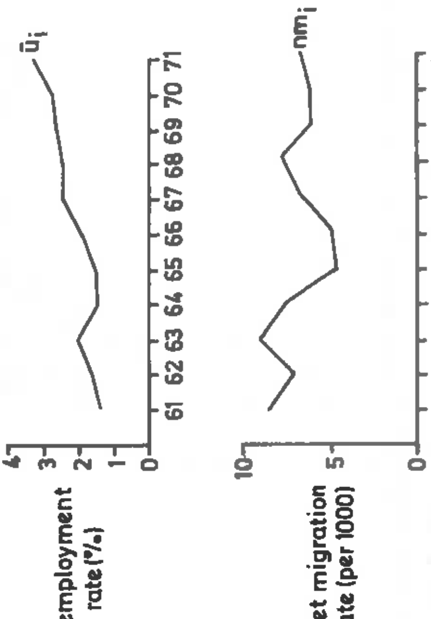


Figure 15: <u>Unemployment Rate and Net Migration Rate Trends for the South West, 1961-1971</u>

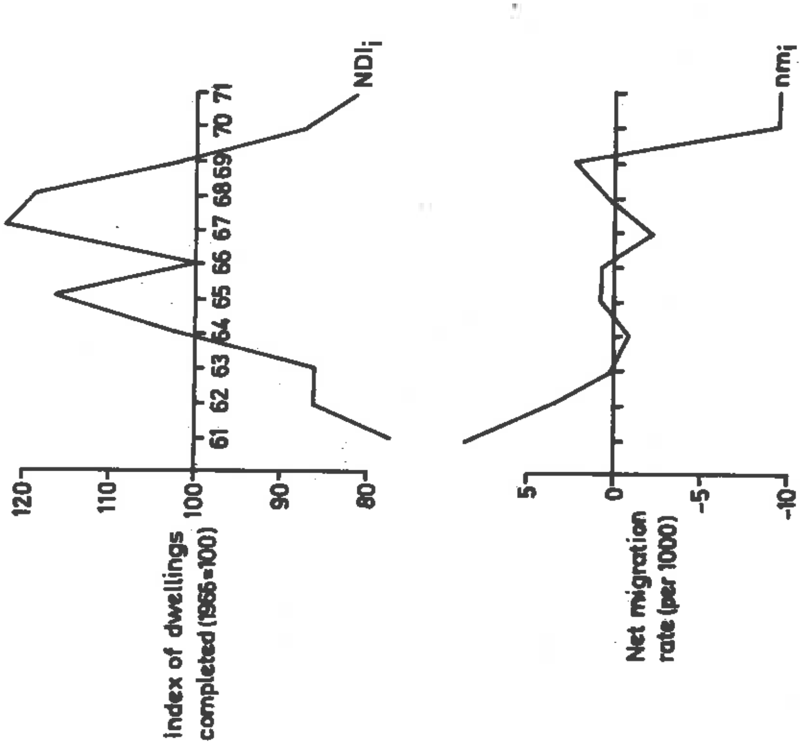


Figure 16: New Dwellings Index and Net Migration Rate Trends for the West Midlands, 1961-1971

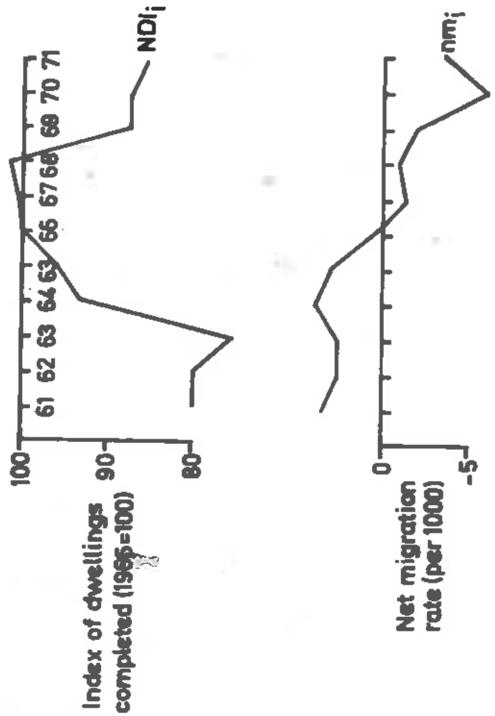


Figure 17: <u>New Dwellings Index and Net Migration Rate Trends for the South East, 1961-1971</u>

4.3.3 Inmigration Rates and Employment Change

The general level of mobility in Britain rose during the period 1961 to 1971. The increase in the propensity to migrate between the regions during this period is confirmed by the observed census statistics for 1961-66 compared with those for 1966-71. Changes in the aggregate inmigration and outmigration rates have been presented in Table 4. The overall inter-regional migration rate rose from 2076710/51286851 = 40.492 per thousand in 1961-66 to 2453675/53047112 = 46.255 per thousand in 1966-71 and a more convincing demonstration of rising migration rates can be obtained from a comparison of the 1960-61, 1965-66 and 1970-71 rates. So, while regional employment remained in decline between 1966 and 1971 and while regional unemployment continued to rise, migration between the regions was occurring at a considerably higher rate than that experienced during the previous five years. One method of incorporating this growth into a migration projection model would be to assume a general increase in the rate of all migration between any two regions, regardless of changing patterns of regional attractiveness. This increase in the general level of mobility might be of the order of 5%, 10% or 15% perhaps. Each historical inmigration rate would therefore require to be upgraded by an assumed or an estimated constant ($\hat{g}$), so that:

$$\tilde{d}_j(\theta) = \hat{g}d_j(\theta-1) \tag{56}$$

where $\hat{g}$ = (1 + the estimated proportional increase in the overall migration rate.)

However, changes in the observed rates between the two five year periods (Table 4) indicate substantial regional variation which suggests some regions are obviously more attractive as migration destinations than others, and therefore experience higher rates of inmigration. Therefore changes in rates of regional inmigration are seen to be partly a function of the general rise in the migration propensity, and partly a function of differing regional attractiveness, which, it may be argued is primarily dependent upon the changing distribution of employment in the country. We might argue, in other words, that one measure of regional attractiveness is that which is based upon the regional share of total employment and the way in which this share changes over time. The development of an approach to projection along these lines is analogous to the approach adopted in shift-share analysis.

Totals of economically-active persons in employment have been obtained from the 1961, 1966 and 1971 Censuses (G.R.O., 1965; 1968; O.P.C.S., 1975) for the aggregated-counties defined in Figure 2. The figures where boundary change involves more than 200 persons have been adjusted by applying the same correction factors that were used to adjust the vital statistics used in determining net migration. Married females in employment are not included, and the 1966 estimates are upgraded by a constant, 1.0142143 (Rees and Wilson, 1977), to allow for underenumeration. The data for the three respective census dates has then been aggregated to give regional employment totals, absolute changes in employment and rates of change in employment (Table 8). The regional shares of total employment and the shifts that occur during the periods are outlined in Table 9. The patterns of changes in total employment illustrating the growth in the early sixties (positive balances) followed by substantial declines (negative balances) in all regions except East Anglia and the South West, are in conformity with the annual figures presented in Figure 9. Furthermore when the regional employment shares are calculated for each year from the annual data, the differences between the 1961 and 1966 percentages and between the 1966 and 1971 percentages (Table 10) show conformity with the two sets in Table 9 calculated from the census data.

The relationship between regional inmigration and 'shifts' in regional employment shares for 1961-66, is illustrated in Figure 18. The correlation coefficient is significant at a 95% confidence level and the regression equation is:

$$d_j(\theta-1) = 48.9 + 43.5\, ES_j(\theta-1) \qquad (57)$$

where

$$ES_j(\theta-1) = \frac{E_j(t,\theta-1)}{E_*(t,\theta-1)} - \frac{E_j(t+T,\theta-1)}{E_*(t+T,\theta-1)} \qquad (58)$$

$E_j(t,\theta-1)$ is the employment in region j at the beginning-of-the-period $\theta-1$, and total employment is defined as:

$$E_*(t,\theta-1) = \sum_j E_j(t,\theta-1) \qquad (59)$$

The effects on inmigration of policy related to changes in the employment distribution may be investigated through this type of analysis. It would seem to offer a more satisfactory and direct approach than the time series analysis using net migration.

Region	Total Ec. Active in Employment			Δ Employment 1961-66	Δ Employment Rate 1961-66	Δ Employment 1966-71	Δ Employment Rate, 1966-71
	1961	1966	1971				
1. North	1,361,500	1,424,515	1,364,049	63,015	0.0463	-60,466	-0.0424
2. North West	3,028,687	3,091,163	2,910,740	62,476	0.0206	-180,423	-0.0584
3. Yorks. & Humberside	2,150,910	2,180,013	2,067,312	29,103	0.0135	-112,701	-0.0517
4. E. Midlands	846,020	949,132	923,279	103,112	0.1219	-25,853	-0.0272
5. W. Midlands	2,908,455	3,095,433	2,992,930	186,978	0.0643	-102,503	-0.0331
6. East Anglia	630,775	692,374	716,020	61,599	0.0977	23,646	0.0342
7. South East	7,803,050	8,078,978	7,885,109	275,928	0.0354	-193,869	-0.0240
8. South West	1,294,597	1,611,495	1,611,970	316,898	0.2448	475	0.0003
9. Wales	1,077,480	1,114,530	1,087,160	37,050	0.0344	-27,370	-0.0246
10. Scotland	2,216,370	2,278,290	2,168,085	61,920	0.0279	-110,205	-0.0484

Table 8: Regional Employment: Totals for 1961, 1966, and 1971; Changes in Total Employment 1961-66 and 1966-71; Rates of Change

Region	% Share of Total Employment			Change in % Share of Total Employment	
	1961	1966	1971	1961-66	1966-71
1. North	5.84	5.81	5.75	-0.03	-0.07
2. North West	12.99	12.61	12.27	-0.38	-0.34
3. Yorks. & Humberside	9.22	8.89	8.71	-0.33	-0.18
4. E. Midlands	3.63	3.89	3.89	0.26	0
5. W. Midlands	12.47	12.63	12.61	0.16	-0.02
6. East Anglia	2.71	2.82	3.02	0.11	0.2
7. South East	33.46	32.95	33.23	-0.51	0.28
8. South West	5.55	6.57	6.79	1.02	0.22
9. Wales	4.62	4.55	4.58	-0.07	0.03
10. Scotland	9.51	9.29	9.14	-0.22	-0.15
Total	100%	100%	100%		

Table 9: Regional Employment: % Shares, 1961, 1966, 1971; and Shifts 1961-66, 1966-71

Region	% Employment Changes (Mid Year)	
	1961-66	1966-71
1. North	-0.11	0.02
2. North West	-0.42	-0.27
3. Yorks. & Humberside	-0.03	-0.26
4. East Midlands		0.07
5. West Midlands	0.19	-0.17
6. East Anglia	0.58	0.21
7. South East		0.23
8. South West	0.1	0.19
9. Wales	-0.04	0
10. Scotland	-0.25	-0.04

Table 10: <u>Regional Employment: % Share Changes 1961-66 and 1966-71, Calculated from Annual Data</u>

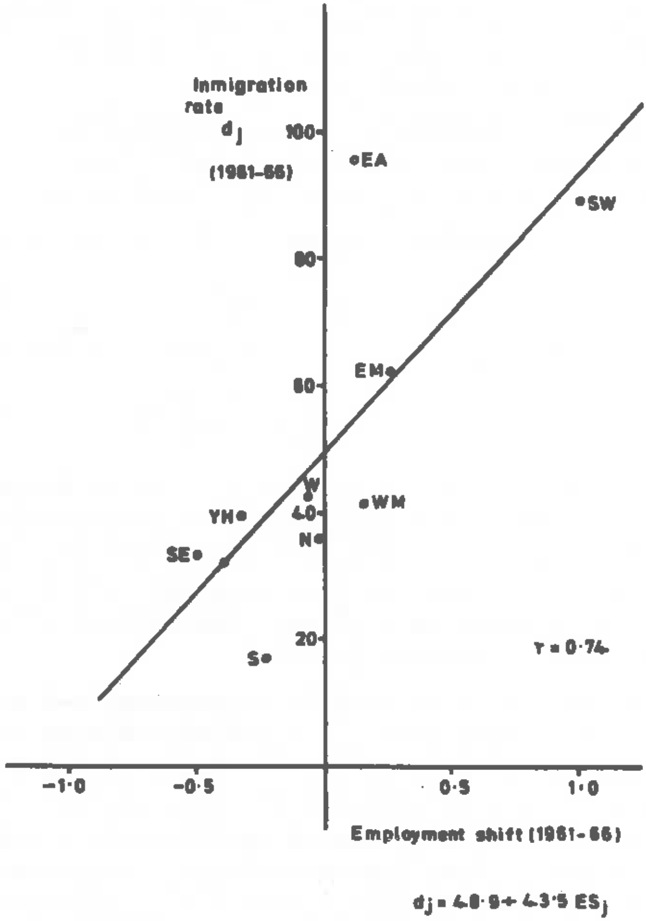


Figure 18: The Relationship Between Regional Immigration Rates and Regional Employment Share Shifts, 1961-66

4.3.4 A Projection Model for Regional Inmigration

The results of previous studies together with the conclusions from the preliminary analysis with regional indicators suggest that a projection model for regional inmigration should therefore be defined on the basis of three features in particular:

(1) the rate of inmigration in the previous historical period,
(2) the general increase in the propensity to migrate,
(3) the changing distribution of employment between regions.

These three features are combined in an inmigration rate model which may be written as:

$$\tilde{d}_j(\theta) = \hat{g}(\theta)d_j(\theta-1)\tilde{ES}_j(\theta) \tag{60}$$

where $\tilde{ES}_j$ is the projected employment shift multiplier redefined as:

$$\tilde{ES}_j(\theta) = \frac{E_j(t,\theta)}{E_*(t,\theta)} \bigg/ \frac{E_j(t+T,\theta)}{E_*(t+T,\theta)} \tag{61}$$

One aim of the model is to test the impact on the rate of inmigration of differing rates of growth in general mobility and changing regional employment shares. It can therefore be used as a technique to generate a set of projections based on different assumptions. The decision to adopt one set of assumptions as the most likely conditions will result in the eventual production of a forecast.

The formulation of an inmigration rate projection model based on a multiplicative relationship between the independent variables is presented as a methodological alternative to the more familiar econometric formulation in which an additive structure is adopted and where additional parameters estimated using multivariate regression analysis, have explanatory significance. Models with a multiplicative structure have been used widely in dynamical systems analysis, while the models developed by Donovan and Morrison and Relles are typical of the traditional econometric approach. The model equation (60) has been defined for purely predictive rather than explanatory purposes, although an additive analogue might easily be tested at a later date.

The remainder of this section is devoted to investigation of the effects on regional inmigration rate projections of making different assumptions about mobility and employment. The simplest situation is one in which the general level of mobility is assumed to grow at a particular rate. For illustrative purposes, it might be assumed that there is a

5% ($\hat{g}$ = 1.05) or 10% ($\hat{g}$ = 1.10) increase in the rate of migration. Table 11 indicates some alternative 1966-71 inmigration rates that have been projected on this basis. The projections have been adjusted by a regional employment shift multiplier derived in the first case, for 1961-66, and in the second case, for 1966-71. These multipliers are based on the information presented in Table 8. It is obviously the selection of the 'most likely' growth level and employment shift mulitpliers that is critical in generating an accurate projection, and unless the objective of the exercise is to experiment with certain assumptions, perhaps geared to policy measures, the analysis is dependent upon historical information. Two of the example projections in Table 11 assume that the historical employment shift multipliers for 1961-66 will be the same in 1966-71. The other two example projections actually make use of the 1966-71 employment information. In a 'true' projection context, the shift multipliers of the latest historical period will be employed if no satisfactory employment shift multiplier projection method is available.

Likewise, future changes in the general level of mobility tend to be projected on the basis of past trends. In the above examples, a specific assumption of 5% or 10% increase in the migration rate was used. However, in this 'pseudo'-projection context, the best fit growth factor can be derived from observed information. The 'best-fit' measure of the increase in the general rate of migration may be calculated from:

$$\hat{g}(\theta) = \frac{\sum_{ij} M_{ij}(1966\text{-}71)}{\sum_i P_i(1966)} \bigg/ \frac{\sum_{ij} M_{ij}(1961\text{-}66)}{\sum_i P_i(1966)} \qquad (62)$$

which becomes

$$\hat{g}(\theta) = \frac{2076710}{51286851} \bigg/ \frac{2453675}{53047112} \qquad (63)$$

$$\hat{g}(\theta) = 1.1423 \qquad (64)$$

and which suggests a 14% increase in the general rate of migration between 1961-66 and 1966-71. The application of this growth factor to the 1961-66 inmigration rates, and adjustment by the 1966-71 employment shift multipliers produces another set of projected rates which are illustrated in Table 12 and which show a much closer approximation to the observed rates. The residuals between the projected and the observed values indicate underprediction in the regions of the North and Scotland, and overprediction in every other case. It would of course be possible to use the observed

Region	$\hat{g}(\theta)\, d_j(\theta-1)$		Employment Shift Multiplier		Projected 1966-71 Immigration Rate per thousand			
	$\hat{g} = 1.05$	$\hat{g} = 1.1$	$ES_j(\theta-1)$	$ES_j(\theta)$				
	(a)	(b)	(c)	(d)	(a)(c)	(a)(d)	(b)(c)	(b)(d)
1. North	37.5	39.3	0.99486	0.98967	37.3	37.1	39.1	38.9
2. North West	33.7	35.3	0.97075	0.97303	32.7	32.8	34.3	34.3
3. Yorks. & Humberside	41.5	43.5	0.96421	0.97975	40.0	40.7	41.9	42.6
4. E. Midlands	65.5	68.6	1.07163	1.00	70.3	65.5	73.5	68.6
5. W. Midlands	43.9	46.1	1.01283	0.99842	44.5	43.8	46.7	46.0
6. East Anglia	100.6	105.4	1.04059	1.07092	104.7	107.7	109.7	112.9
7. South East	34.9	36.6	0.98476	1.0085	34.4	35.2	36.0	36.9
8. South West	94.2	98.7	1.18378	1.0335	111.5	97.4	116.8	102.0
9. Wales	44.3	46.4	0.98485	1.0066	43.6	44.6	45.7	46.7
10. Scotland	17.7	18.6	0.97687	0.98385	17.3	17.4	18.2	18.6

Table 11: <u>**Projected Regional Immigration Rates, 1966-71; Based on Assumed Growth Factors and Employment Shift Multipliers**</u>

Region	$d_j(\theta-1)$	$\hat{\varepsilon}(\theta)\ d_j(\theta-1)$ ($\hat{\varepsilon} = 1.1423$)	Employment Shift Multiplier $ES_j(\theta)$	Immigration Rates per thousand	
				Projected $\tilde{d}_j(\theta)$	Observed $d_j(\theta)$
1. North	35.7	40.8	0.98967	40.4	42.5
2. North West	32.1	36.7	0.97303	35.7	35.2
3. Yorkshire & Humberside	39.5	45.1	0.97975	44.2	42.2
4. East Midlands	62.4	71.3	1.00000	71.3	70.7
5. West Midlands	41.9	47.9	0.99842	47.8	42.3
6. East Anglia	95.8	109.4	1.07092	117.2	115.0
7. South East	33.3	38.0	1.00850	38.3	37.0
8. South West	89.7	102.5	1.03350	105.9	98.5
9. Wales	42.2	48.2	1.00660	48.5	48.1
10. Scotland	16.9	19.3	0.98385	19.0	23.7

Table 12: <u>The Projection of Regional Immigration Rates, 1966-71; Based on a Calculated Growth Factor and the 1966-71 Employment Shift Multipliers</u>

regional rates and the general rate of migration increase to determine the precise employment share shifts that would be needed to bring the two estimates into line. These multipliers might then be used in projection analysis for subsequent periods.

If no further information beyond 1966 was available, methods of using other historical migration data would be required. It might be convenient, for example, to make use of the single year migration data for 1960-61 and 1965-66 as a means of projecting for 1966-71. If it was assumed that the difference between the general migration rates for these two single year periods was the same as the difference between the 1965-66 and 1970-71 rates, then on the basis of linear interpolation, a single year rate could be estimated for each year through to 1971 and these one year rates could be aggregated to generate a five year figure. However, there are several conceptual problems associated with the aggregation of single year rates to represent a five year period (Rees, 1977) and boundary changes present further complications.

It is appropriate at this stage to summarise and compare the alternative projections of inmigration and outmigration that have been produced. The methods and their associated problems have been discussed in some detail. In the next section, the projected flow totals are reviewed and some are used to examine the application of distribution models in a projection context.

5. MODEL-BASED MIGRATION PROJECTIONS FOR THE REGIONAL SYSTEM

In the first part of this section, alternative flow projections of total regional inmigration and outmigration for 1966-71 are summarised and compared. The distribution models introduced in section 3 are applied in predictive mode using the observed out- and in-migration totals for 1966-71 and one set of the pseudo-projected totals. The results are compared against the observed interregional flow matrix for this period. A set of forecast outmigration and inmigration totals for 1971-76 is prepared in the second part of the section and the two best-fit distribution models of the recalibrated series are then used to obtain forecasts of the inter-regional migration flows in 1971-76. Characteristics of the historical accounts for 1966-71 and the forecasted accounts for 1971-76 are considered in the third section.

5.1 A Summary and Comparison of Migration Flow Projections for 1966-71

5.1.1 Total Inmigration and Outmigration Flow Projections

Several alternative in- and out-migration rate projections have been generated for 1966-71 in the previous sections. Table 13 contains a summary of the projected inmigration flow totals determined by applying the rates to start-of-the-period regional populations. Nine different inmigration totals have ben projected. These are as follows:

$\tilde{D}_j(1)$ - based on the historical inmigration rate for 1961-66 as defined in equation (31).

$\tilde{D}_j(2)$ - based on the regression relationship with the outmigration rate for 1961-66 as defined by:

$$d_j(\theta-1) = -57.5 + 2.35\ o_j(\theta-1) \qquad (65)$$

$\tilde{D}_j(3)$ - based on the net migration 'estimate' for 1966-71 and the regression relationships defined in equations (39) and (41).

$\tilde{D}_j(4)$ - based on the net migration 'projection' for 1966-71 and the regression relationships defined in equations (39) and (41).

$\tilde{D}_j(5)$ - based on the assumption of 5% growth in general mobility and the application of 1961-66 employment shift multipliers.

$\tilde{D}_j(6)$ - based on the assumption of 5% growth in general mobility and the application of 1966-71 employment shift mulipliers.

$\tilde{D}_j(7)$ - based on the assumption of 10% growth in general mobility and the application of 1961-66 employment shift multipliers.

$\tilde{D}_j(8)$ - based on the assumption of 10% growth in general mobility and the application of 1966-71 employment shift multipliers.

$\tilde{D}_j(9)$ - based on the observed growth in general mobility rate and the application of 1966-71 employment shift multipliers.

Table 14 contains the projected outmigration flow totals defined as follows:

$\tilde{O}_i(1)$ - based on the historical outmigration rate for 1961-66 as defined by equation (28).

$\tilde{O}_i(2)$ - based on the regression relationship with the inmigration rate for 1961-66 as defined in equation (36).

$\tilde{O}_i(3) \rightarrow \tilde{O}_i(9)$ - each of these projected outmigration flow totals is based on the corresponding projections of inmigration rates and the regression relationship defined as in equation (36).

Region	Projected Flows									Observed Flows
	$\tilde{O}_i(1)$	$\tilde{O}_i(2)$	$\tilde{O}_i(3)$	$\tilde{O}_i(4)$	$\tilde{O}_i(5)$	$\tilde{O}_i(6)$	$\tilde{O}_i(7)$	$\tilde{O}_i(8)$	$\tilde{O}_i(9)$	O_i
1. North	167788	134498	108880	116492	136289	136064	138314	138089	139777	166341
2. North West	234527	261860	243828	237849	263215	263441	266830	266830	269992	275809
3. Yorkshire and Humberside	192403	179676	148695	151266	180405	181425	183173	184193	186524	229715
4. E. Midlands	183689	190653	174102	165665	200954	194695	205126	198737	202257	223135
5. W. Midlands	225261	213089	192920	171983	217496	216309	221225	220038	223089	249740
6. East Anglia	95901	95079	93721	98392	99790	101378	102436	104130	106406	105302
7. South East	583242	669333	585772	759047	675620	680191	684763	687906	697907	673952
8. South West	216789	217989	200539	185397	245263	227558	251918	233334	238231	239909
9. Wales	113446	115462	111829	112368	116744	117561	118668	119585	121234	123181
10. Scotland	169125	179314	164412	158129	180026	180204	181629	182341	183053	166591
Total	2182191	2256993	2124698	2155588	2315802	2298926	2354082	2337183	2368470	2453675
O_i v $\tilde{O}_i$ Correlation coefficient	0.997	0.994	0.987	0.982	0.993	0.994	0.993	0.994	0.994	1.00

* All coefficients are significant at the 99% level of confidence

Table 13: <u>Observed and Projected Immigration Flow Totals, 1966-71 (with Correlation Coefficients)</u>

Region	Projected Flows									Observed Flows
	$\tilde{D}_j(1)$	$\tilde{D}_j(2)$	$\tilde{D}_j(3)$	$\tilde{D}_j(4)$	$\tilde{D}_j(5)$	$\tilde{D}_j(6)$	$\tilde{D}_j(7)$	$\tilde{D}_j(8)$	$\tilde{D}_j(9)$	D_j
1. North	118146	204009	42692	64865	123441	122780	129398	128737	133701	140599
2. North West	213267	169119	160116	142842	217253	217917	227883	227883	237185	233932
3. Yorks and Humberside	169263	205751	77561	86131	171406	174405	179548	182547	189403	180779
4. E. Midlands	239294	211166	190592	165665	269590	251182	281861	263070	273425	271052
5. W. Midlands	208872	244041	150049	87238	221833	218343	232800	229310	238283	210809
6. East Anglia	149144	135849	145242	158952	163000	167670	170784	175766	182460	178985
7. South East	559711	404151	608454	820237	578200	591646	605093	620220	643751	619505
8. South West	331277	297096	230311	235993	411767	359714	431361	376702	391106	364079
9. Wales	115010	111654	102937	104554	117478	120183	123147	125842	130692	129600
10. Scotland	88489	96370	44507	26180	90584	91108	95296	97391	99485	124340
Total	2192473	2079205	1802461	1892657	2364582	2314948	2477171	2427468	2519491	2453675
D_j v $\tilde{D}_j$ Correlation Coefficient	0.997	0.909	0.986	0.953	0.987	0.997	0.997	0.997	0.997	1.00

* All coefficients are significant at the 99% level of confidence

Table 14: Observed and Projected Outmigration Flow Totals, 1966-71 (with Correlation Coefficients)

Tables 13 and 14 also contain the Pearson's coefficients which represent the degree of correlation between each set of projected totals and the set of observed in- or out-migration flow totals. The highest coefficients for the inmigration flow totals are associated with $\tilde{D}_j(1)$, $\tilde{D}_j(6)$, $\tilde{D}_j(8)$ and $\tilde{D}_j(9)$, whereas the projections of outmigration based upon the historical rates, $\tilde{O}_i(1)$, appear to be superior to the other projections, on the basis of this goodness-of-fit statistic. However, evaluation of alternative projections can be assessed more accurately by comparison of the deviation of each projected flow total from its observed value, and by comparison of the sums of absolute deviations for the projection sets.

Table 15 illustrates the 'projected-minus-observed' residuals associated with each set of inmigration flows for 1966-71. The absolute values of the derivations summed for each projection set indicate that inmigration based on estimated or projected net migration figures is extremely inaccurate. The regional residuals show severe underprediction in almost all cases. The most accurate projections on the other hand, are those produced on the assumption of a 10% increase in the historical inmigration rates for the previous period, and adjustment using the 1966-71 employment shift multipliers. The projections obtained using the best-fit growth factor and the 1966-71 employment shift multipliers appear to be less accurate and the residuals indicate overprediction in eight out of ten regions.

The projected totals that have been presented in Tables 13 and 14 require some adjustment to ensure that:

$$\sum_j \tilde{D}_j = \sum_i \tilde{O}_i \qquad (66)$$

for each set of projections. This condition is essential for balancing factor convergence in constrained spatial interaction models. It was decided to adjust the outmigration projections rather than the inmigration totals and therefore the former flows were multiplied by a constant defined as $\sum_j \tilde{D}_j / \sum_i \tilde{O}_i$, where $\sum_i \tilde{O}_i$ refers to the summation of the unadjusted outmigration totals.

Region	$\tilde{D}_j - D_j$								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
1. North	-22,453	63,410	-97,907	-75,734	-17,158	-17,819	-11,201	-11,862	-6,898
2. North West	-20,465	-64,813	-73,816	-91,090	-16,679	-16,015	-6,049	-6,049	3,253
3. Yorks. and Humberside	-21,516	24,972	-103,218	-94,648	-9,373	-6,374	-1,231	1,768	8,624
4. E. Midlands	-31,758	-59,886	-80,460	-105,387	-1,462	-19,870	10,809	-7,982	2,373
5. W. Midlands	-1,937	33,232	-60,760	-123,571	11,024	7,534	21,991	18,501	27,474
6. East Anglia	-23,836	-43,131	-33,738	-20,028	-15,980	-11,310	-8,196	-3,214	3,480
7. South East	-59,794	-215,354	-11,051	200,732	-41,305	-27,859	-14,412	715	24,246
8. South West	-32,802	-66,983	-83,768	-128,086	47,708	-4,365	67,282	12,623	27,027
9. Wales	-14,590	-17,946	-26,663	-25,046	-12,112	-9,417	-6,453	-7,758	1,092
10. Scotland	-35,851	-27,770	-79,873	[illegible]	-33,756	-33,232	-29,044	-26,949	-24,855
Total	-261,202	-374,469	-651,214	-561,018	-89,093	-138,701	28,496	-26,207	65,816
Absolute Total	261,202	617,697	651,214	962,482	206,557	153,795	176,678	93,421	129,322

Table 15: Projected-minus-Observed Immigration Flow Residuals 1966-71

5.1.2 Inter-regional Migration Projection Models

The distribution models defined in Section 3 can be applied to produce interregional migration flow matrices based on either the observed out- and in-migration totals or any one set of projected out- and in-migration totals. It is only the historical rates model that does not require independent projections or estimates of total regional out- and in-migration. In Table 16 goodness-of-fit statistics enable the results of alternative distribution models to be compared. All the models except the historical rates model have been tested using the observed out- and in-migration totals for 1966-71. Most of the models have also been tested with one set of projected in- and out-migration totals from Tables 13 and 14. The $\tilde{D}_j(6)$ and $\tilde{O}_i(6)$ projections were selected for this illustrative purpose and the outmigration totals were adjusted to ensure that $\sum_j \tilde{D}_j(6) = \sum_i \tilde{O}_i(6)$. The contiguity models were tested with observed flows because of the problems associated with dis-aggregating projected total out- and in-migration into flows between contiguous and non-contiguous regions, and the singly constrained spatial interaction models were tested using population stocks rather than employment totals as proxy attractiveness factors. Three summary conclusions can be drawn from the results.

Firstly, the growth factor model using observed out- and in-migration totals appears to give the 'best-fit' distribution of interregional migrants according to the index of dissimilarity and the sum of squared deviations. Hagerstrand provides the argument for supposing migration streams to be historically determined and the conclusion is that a model for a projection period (θ) which incorporates $M_{ij}(\theta-1)$ will always tend to be better than a model that uses $d_{ij}^{-\beta}(\theta-1)$. Since the d_{ij} matrix in projective models is generally assumed to remain fixed, much of the rationale of the transport model applied to migration flows is no longer appropriate. Transport models are normally formulated with the intention of testing alternative plans either through changes in attractiveness factors in singly constrained models, or through changes in the elements contained in the trip cost matrix. The situation with migration is different because there does not exist an operational method for measuring the costs associated with moving. Physical distance is therefore employed in the absence of a satisfactory measure, and this is certainly one of the areas of limitation in this style of modelling.

Model Specification		Model Equation	Coefficient of Correlation	Coefficient of Determination	Index of Dissimilarity (I.D.)	Sum of Squared Deviations (SS)
Historical Rates Model		(2)	0.99964	0.99927	3.9	2,642,535,958
Growth Factor Model		(14)	0.99402 0.99827	0.98807 0.99654	4.9 2.6	1,239,451,865 313,610,939
Doubly Constrained Spatial Interaction Models	Power Function, Generalized β_*	(15)	0.95345 0.95126	0.90907 0.9049	11.1 11.3	8,312,601,092 8,530,822,547
	Exponential Function, Generalized β_*	(16)	0.94795 0.94[illegible]	0.89861 0.89598	12.7 12.3	9,200,702,148 9,47[illegible],40[illegible],840
	Power Function Origin-specific β_i	(17)	0.95[illegible]4 0.95487	0.91516 0.91177	11.[illegible] 10.8	7,757,585,513 7,895,843,444
	Exponential Function, Origin-specific β_i	(18)	0.95659 0.95574	0.91506 0.91344	11.6 11.3	7,764,330,924 7,748,692,631
	Power Function, Destination-Specific β_j	(19)	0.95[illegible] 0.95197	0.90831 0.90624	11.2 11.2	8,327,446,954 8,393,367,979
	Exponential Function, Destination-Specific β_j	(20)	0.96094 0.96028	0.92341 0.92214	10.6 10.1	7,023,774,886 7,004,197,117
	Power Function, Contiguous/non-contiguous, β_*	(21)	– 0.98728	– 0.97472	– 5.9	– 2,281,413,746
	Exponential Function, Contiguous/non-contiguous β_*	(22)	– 0.98757	– 0.97530	– 5.9	– 2,226,567,923
Singly Constrained Spatial Interaction Models	Production constrained, power function Generalized β_*	(23)	0.80798 0.80724	0.65284 0.65164	18.4 18.1	31,118,460,999 31,206,368,845
	Production constrained, exponential function, Generalized β_*	(24)	0.80077 0.80[illegible]	0.64123 0.64575	19.0 18.8	32,395,398,950 31,940,360,242
	Attraction constrained, Power function Generalized β_*	(25)	0.91186 0.91162	0.83094 0.83142	14.1 13.8	15,765,014,060 15,191,916,325
	Attraction constrained Exponential function Generalized β_*	(26)	0.85441 0.85[illegible]	0.72995 0.72834	18.8 18.4	24,307,703,116 24,105,585,[illegible]

Notes: (1) The top figure in each case is the goodness-of-fit statistic associated with the distribution based on projected row and column totals, and the bottom figure, the statistic associated with observed row and column totals.

(2) Index of Dissimilarity is defined as $\frac{1}{2}(100) \sum_{\substack{i \neq j \\ j \neq i}} \left| \frac{M_{ij}}{\sum_{\substack{i \neq j \\ j \neq i}} M_{ij}} - \frac{\hat{M}_{ij}}{\sum_{\substack{i \neq j \\ j \neq i}} \hat{M}_{ij}} \right|$

(3) SS is defined as $\sum_{\substack{i \neq j \\ j \neq i}} (M_{ij} - \hat{M}_{ij})^2$

(4) The singly constrained models have been computed with population stocks as proxy attractiveness factors.

Table 16: Goodness-of-Fit Statistics Associated with Models for Projecting Inter-regional Migration, 1966-71

A second conclusion is that there are significant differences between the goodness-of-fit statistics for alternative constrained spatial interaction models, but none is as competent in projecting the distribution of inter-regional migrants as the growth factor model. The generalised decay parameter model with a negative power function generates a more accurate distribution than the model incorporating the negative exponential function. This observation is consistent with the model results for the historical period, 1961-66. The zone-specific parameter models generally improve on the distribution generated by the models using generalised parameters, and it is the destination-specific model with a negative exponential distance function that is most accurate. Both the continguity/non-continguity split models produce relatively good distributions, but the singly constrained models are much less efficient. The attraction constrained model with a negative power decay function is the best of the four.

One question that arises from the summary of results in Table 16 is concerned with which of the statistics is most reliable for comparative analysis. Each reflects the goodness-of-fit in a rather different way, and it is not inconsistent to observe that the historical rates model produces a distribution which, when compared with the observed distribution, has a coefficient of determination of 0.99927 and a dissimilarity index of 3.9, whereas the growth factor model generates a distribution based on observed totals which has a lower coefficient of determination of 0.99654 but also a lower index of dissimilarity of 2.6. Probably the most reliable statistic that can be used for comparing different models is therefore the sum of squared deviations, although the significance of accurate projections of flows between certain regions using any particular model can be obscured by a very inaccurate projection of flows between other regions in the same model run. The sum of squares statistic suggests that although the growth factor model is clearly the 'optimum' model, the continguity/non-contiguity split model produces a more significant set of results than does the historical rates model.

One implication that might be drawn from the results is that while historical rates models or growth factor models are perhaps more applicable for <u>short term</u> projection because of the significance of the 'historical effect' in migration behaviour, it may be the case that spatial interaction models are more useful in the <u>long term</u>, when rates of migration are more likely to experience dramatic fluctuation or when policy changes are more likely to be influential. The conclusion that can be drawn in this context is that it is more important in the short term, to forecast total outmigration and inmigration than it is to forecast other exogenous variables that might be tied into spatial interaction models. In other words, the projection of system variables rather than control variables seems to be more important in the short term. It has been shown that significant relationships between policy variables and migration are difficult to identify. Nevertheless, the construction of spatial interaction models with policy sensitive out- and in-migration variables becomes more appropriate for projection in the long term.

5.2 Total and Inter-regional Migration Flow Forecasts for 1971-76

The next step in the analysis involves the preparation of sets of migration projections in the absence of observed data, and the tentative selection of suitableprojections as forecasts.

5.2.1 A Set of Inmigration and Outmigration Forecasts

One set of projected inmigration rates and flows for 1971-76 can be produced by assuming a further increase in the general migration rate equivalent to that increase between the two previous five year periods, and by applying the 1966-71 employment shift multipliers. The outmigration rates for the period can then be estimated using the 1966-71 observed rate regression relationship calculated as:

$$o_i(1966\text{-}71) \;=\; 31.35 + 0.34\, d_i(1966\text{-}71) \qquad (69)$$

and the outmigration totals would have to be adjusted to ensure that $\sum_j \hat{D}_j = \sum_i \hat{O}_i$. Inmigration and outmigration projections produced in this particular way are presented as forecasts in Table 17.

The regional net migration rate balances that can be derived from these very tentative 'forecasts' for 1971-76 may be compared with a set of net migration rate 'estimates' for the same period. Rees (1978) uses totals of population, births, deaths and external migration data to generate estimates of net migration for 1975-76. The rates for this single year, together with the observed rates for 1970-71 can be averaged and multiplied by five to provide an alternative set of rates for 1971-76 (Table 18) for each region. The two sets of net migration rates are not directly comparable because Rees uses post-1974 Standard Regions as the zones in his system of interest. Nevertheless, certain crude comparisons can be made between them. Both sets of values are consistent in that relatively high rates of netmigration gain are associated with the East Midlands, East Anglia and the South West. The major inconsistencies in

Region	Total Inmigration		Total Outmigration		
	Rate*	Flow	Rate*	Unadjusted Flow	Adjusted Flow**
1. North	48.0	157,996	47.7	157,009	169,816
2. North West	39.1	261,258	44.6	298,007	322,315
3. Yorkshire & Humberside	47.2	202,849	47.4	203,708	220,324
4. East Midlands	80.8	321,072	58.8	233,652	252,711
5. West Midlands	48.2	246,579	47.7	244,022	263,927
6. East Anglia	140.7	233,805	79.2	131,609	142,344
7. South East	42.6	729,695	45.8	784,509	848,500
8. South West	116.3	449,871	70.9	274,255	296,626
9. Wales	55.3	150,713	50.2	136,813	147,973
10. Scotland	26.6	138,951	40.4	211,038	228,253
		2,892,789		2,674,621	2,892,789

*Rate per thousand, rounded to one decimal place

**The outmigration flows are adjusted so that $\sum_j \tilde{D}_j = \sum_i \tilde{O}_i$

Table 17. Total Regional Inmigration and Outmigration Forecasts, 1971-76.

Region	1971-76 Net migration Rate Balances (in thousands) based on estimates for 1970-71 and 1975-76 from Rees (1978, Table 12)				1971-76 Net Migration Rate Projections derived from Table 17	Difference
	1970-71 (a)	1975-76 (b)	Average $(c)=\frac{(a)+(b)}{2}$	Average x 5 (d)=(c) x 5	(e)	(d) - (e)
1. North	-2.9	4.9	1.00	5.00	-0.3	5.3
2. North West	-3.1	-16.1	-9.60	-48.00	-5.5	-42.5
3. Yorkshire and Humberside	-1.1	-1.3	-1.20	-6.00	-0.2	-5.8
4. East Midlands	16.5	6.7	11.60	58.00	22.0	36.0
5. West Midlands	-5.2	-15.8	-10.50	-52.50	0.5	-53.0
6. East Anglia	22.3	19.1	20.70	103.50	61.5	42.0
7. South East	-48.9	-53.6	-51.25	-256.25	-3.2	-253.05
8. South West	22.4	40.7	31.55	157.75	45.4	112.35
9. Wales	3.7	9.9	6.80	34.00	5.1	28.9
10. Scotland	-3.7	5.5	0.90	4.50	-13.8	18.3

Table 18 : Estimated and Projected Regional Net Migration Rate Balances for 1971-76

the two sets of figures are associated with net migration rates in the South East, the West Midlands and the North West. The estimates suggest very much higher rates of net migration loss during 1971-76 than do the forecasts. The net migration estimate for Scotland, on the other hand, suggests that the five year rate will be positive, while the forecast rate indicates that a net out-migration rate will prevail.

If the estimated figures are assumed to be the 'observed' rates, the forecasted rates are not very accurate. This forecasting methodology has seriously underestimated losses in the South East, West Midlands and the North West, overestimated losses from Scotland, and underestimated net gains to the East Midlands, East Anglia, the South West and Wales. Shifts in regional attractiveness have taken place however. Metropolitan decentralisation now extends well beyond the South East, and in Scotland, North Sea oil can be held as partly responsible for stemming the continuation of net migration loss.

5.2.2 Two Sets of Inter-regional Migration Flow Forecasts

Comparison of the results of the distribution models for 1966-71 provides some guidance when projecting the inter-regional migration matrix for 1971-76. Two sets of forecasted inter-regional migration flows have been prepared for this period. The first distribution (top set figures in Table 19) is based on the growth factor model using the observed matrix for 1966-71 and the forecasts of total outmigration and inmigration that were presented in Table 17. Rounding errors prevent equivalence of the row and column totals for the flows in Table 19 and the respective Table 17 totals. The second distribution (bottom set of figures in Table 19) is based on the same out- and in-migration totals but here they are distributed using a contiguity/non-contiguity split model. Comparison of the parameters calibrated for models of the historical periods 1961-66 and 1966-71 indicates that the propensities to migrate over distance between contiguous or non-contiguous regions have changed. The non-contiguous region parameter associated with a negative power function fell from -0.743 to -0.663 while the contiguous region parameter calibrated as +0.1948 for 1961-66 was recalibrated as +0.2254 for 1966-71. The parameters for 1966-71 were used in the forecasting run of the model. The other problem associated with using the split model in a forecasting context was the problem of how to disaggregate the out- and in-migration to and from contiguous and non-contiguous regions. Table 20 illustrates the observed flow totals and proportions of migration between contiguous and non-contiguous regions in

Region of Origin	Method	N	NW	YH	EM	WM	EA	SE	SW	W	S	Out-migration Total O_i
N	(a)		21549	29610	18850	13414	7265	48789	11542	3932	14653	169804
	(b)		23596	28987	12728	11586	6353	52052	13735	4469	16535	169541
NW	(a)	21247		35409	35940	33541	10800	96639	33524	38750	16445	322295
	(b)	27040		44486	39978	29558	11854	97321	29045	26610	16236	322128
YH	(a)	32610	34105		47853	15173	8185	52008	15447	5959	9070	220410
	(b)	29029	39560		47340	14340	7064	54352	14526	4673	9310	220194
EM	(a)	13154	23469	39858		32398	23806	74500	26664	7959	10891	252699
	(b)	9422	27917	37174		17245	34363	78805	29198	8802	10468	253394
WM	(a)	10481	32158	12581	42124		9791	79230	45615	22050	9891	263921
	(b)	8621	30255	11198	25278		12568	89257	59159	17921	10528	264785
EA	(a)	4457	6814	6851	21295	7319		71523	14654	4180	5230	142323
	(b)	4273	7068	4986	25393	11360		67390	13079	3957	4965	142471
SE	(a)	40460	76243	41516	103622	75902	148097		257372	44759	54697	848668
	(b)	44946	74490	49252	115979	89617	134210		234614	50954	54552	848614
SW	(a)	9843	17944	10295	18614	28368	12542	168204		17692	13065	296568
	(b)	8532	15995	9471	24505	38486	12080	152016		24947	11184	297216
W	(a)	4222	20443	5463	9633	21803	4541	51000	25849		5010	147964
	(b)	3836	22363	4210	10207	14714	5049	50651	31485		5172	147687
S	(a)	21521	28624	15066	23141	18661	8778	87803	19203	5434		228231
	(b)	22297	20013	13586	19664	19674	10265	87850	25032	8378		226759
In-migration Total D_j	(a)	157995	261349	202850	321072	246579	233805	729696	449870	150715	138952	2892883
	(b)	157996	261257	202850	321072	246580	233805	729694	449873	150711	138950	2892789

Table 19: <u>Inter-regional Migration Flow Forecasts for 1971-76 Based on (a) Growth Factor Methodology, (b) Contiguity/Non-contiguity Split Methodology</u>

Region of Origin	Time Period	Contiguous Flow Total		Non-contiguous Flow Total		Contiguous Flow Proportion		Non-contiguous Flow Proportion	
		O_i	D_j	O_i	D_j	O_i	D_j	O_i	D_j
1. North	1961-66	63470	60203	101635	55894	0.384	0.519	0.616	0.481
	1966-71	66656	69699	99685	70900	0.401	0.496	0.599	0.504
2. North West	1961-66	121624	112943	107811	95853	0.530	0.540	0.470	0.460
	1966-71	142580	128662	133229	105270	0.517	0.550	0.483	0.450
3. Yorkshire & Humberside	1961-66	101522	88947	86067	75894	0.541	0.540	0.459	0.460
	1966-71	120065	98240	109650	82539	0.523	0.543	0.477	0.457
4. East Midlands	1961-66	138674	169820	35284	56928	0.797	0.749	0.203	0.251
	1966-71	171548	214336	51587	56716	0.769	0.791	0.231	0.209
5. West Midlands	1961-66	183837	146868	32121	52799	0.851	0.736	0.149	0.264
	1966-71	208729	162150	41011	48659	0.836	0.769	0.164	0.231
6. East Anglia	1961-66	58166	99596	31946	40569	0.645	0.711	0.355	0.289
	1966-71	68274	129028	37028	49952	0.648	0.721	0.352	0.279
7. South East	1961-66	374935	282135	185835	256090	0.669	0.524	0.331	0.476
	1966-71	454392	329199	219560	290306	0.674	0.531	0.326	0.469
8. South West	1961-66	149151	232519	55497	80011	0.729	0.744	0.271	0.256
	1966-71	173250	263391	66659	100688	0.722	0.723	0.278	0.277
9. Wales	1961-66	56198	57506	54940	53875	0.506	0.516	0.494	0.484
	1966-71	56800	68880	66381	60720	0.461	0.531	0.539	0.469
10. Scotland	1961-66	14513	11603	152954	76177	0.087	0.132	0.913	0.868
	1966-71	16117	14826	150474	109514	0.097	0.119	0.903	0.881

Table 20: **Flow Totals and Proporations of Outmigration and Immigration Between Contiguous and Non-Contiguous Regions, 1961-66 and 1966-71**

1961-66 and 1966-71. Since there do not appear to be any significant changes in the proportions for these periods, the 1966-71 proportions were applied to disaggregate the forecasted 1971-76 totals.

One important feature of the type of research reported in this paper is that projection and forecasting of aggregate migration totals or inter-regional migration distributions can be achieved through a combination of different methods. The approach that is selected for forecasting will depend upon the time period for which projection is made, the availability of data, and the particular purpose for which the projection is required. Several alternative models for projecting the inter-regional distribution of migrants for 1966-71 have been tested against the observed distribution so far, and two matrices of 'forecasted' inter-regional flows have been generated for the subsequent five year period, 1971-76. In the next section, observed, estimated and projected data sets have been used to construct regional population accounts for 1966-71 and 1971-76.

5.3 Regional Accounts

5.3.1 The Construction of Regional Population Accounts for 1966-71 and 1971-76

Population accounts for a system of ten regions plus a rest-of-the-world zone can be estimated using the accounts-based model programme developed by Jenkins and Rees (1977).

The data required by the programme to produce accounts for 1966-71 is as follows:

(a) Initial populations: Regional population totals on census data, 1966, derived from an aggregation of county population totals (Table A1).

(b) Exist/survive inter-regional migration: An observed migration matrix of internal flows, obtained by aggregation of inter-county flows available in unpublished form from O.P.C.S. Immigration flows were available from the same source, and 'unknown' emigration flows estimated by Stillwell (1979) were incorporated. Internal and external flows were adjusted for boundary changes (Appendix).

(c) Birth/survive inter-regional migration: A matrix of internal and external birth and survival flows obtained from an auto-projection run of the accounts-based model used initially to produce 1961-66 accounts. A birth and survival transition rate matrix was defined for 1961-66 and it was assumed that the rates would remain constant for 1966-71.

(d) Births totals: Regional totals of births for 1966-71 were obtained by an aggregation of births totals in respective counties. The annual births statistics published by the Registrar General have been amalgamated and adjusted to provide a measure of the number of births in the five year period prior to the 1971 census.

(e) Deaths totals: Regional totals of deaths were estimated on the basis of Registrar General annual deaths statistics for counties, aggregated using techniques similar to those employed in estimating the five year birth totals.

Population accounts for subsequent five year periods can be projected using the same model programme, and where additional data is available (observed birth rates or death rates for example), this can be incorporated into these projections. Alternatively, auto-projection runs of the model can be undertaken in which the results are totally dependent upon a base period data set. The accounts for 1971-76 have been estimated from data, some of which is observed and some of which has itself been projected. The data required to produce accounts for 1971-76 is as follows:

(a) Initial populations: Regional population totals on census data, 1971, derived from an aggregation of county population totals.

(b) Exist/Survive inter-regional migration: The matrix of internal inter-regional migration flows, obtained by distributing the 1971-76 total out- and in-migration forecasts (Table 17) using a contiguity/non-contiguity split model (Table 19). External flows were assumed to be the same as in the previous period. This is the assumption made in the auto-projection runs of the accounts-based model.

(c) Birth/Survive inter-regional migration: The complete birth and survival matrix was taken from the auto-projection run based on 1961-66 and this means that the 1961-66 birth and survival rates were assumed to remain the same throughout 1966-71 and 1971-76.

(d) Total births and total deaths: Once again, the totals of regional births and deaths for 1971-76 projected automatically by the 1966-71 auto-projection model were incorporated.

The nature of this particular data set serves to illustrate that projected or forecasted accounts can be generated that allow the consequences of making certain assumptions (perhaps conditioned by policy measures) or the implications of adopting certain migration distributions defined with different techniques, to be investigated. Undoubtedly the accuracy of both sets of accounts could be improved substantially ('observed' births and deaths totals could be estimated for 1971-76 for example).

Although the complete regional accounts matrices constructed for the periods 1966-71 and 1971-76 are presented elsewhere (Stillwell, 1979), some brief comments on the characteristics of regional population illustrated by the accounts are included here.

5.3.2 Analysis of Regional Population Characteristics, 1966-71 and 1971-76.

The main features of the accounts tables can be summarised by classifying the regions according to their population change characteristics, and by analysing changes in the distribution of population in the main system. Summary tabulations are presented for 1966-71 (Table 21) and 1971-76 (Table 22). The main feature is that the classified regional pattern for 1966-71 was similar to that experienced in the first half of the decade, although there were obviously differences in the actual components of change. East Anglia and the South West continued to experience higher net inmigration than natural increase, while natural increase exceeded net inmigration in the East Midlands. However, unlike the previous period, outmigration from the South East was sufficient to exceed inmigration, but the total net migration balance was still offset by natural increase. This was also typical of the situation in the other six regions. As a result, the distribution of population indicates that East Anglia, the South West and the Midlands regions maintained their growth trends through 1966-71 while the population in the remaining parts of the country, as a percentage of total population, was declining. The main difference between this pattern and the general trends provided by the forecasted accounts for 1971-76 is that Scotland experienced an overall decline in population due to increased net outmigration. As in the previous period, the percentage population increase in East Anglia, the South West and the Midlands continued at the expense of higher growth rates in the rest of the country. The sets of historical and projected accounts illustrate how independent estimates or projections of inter-regional migration can be integrated into a model for describing a more comprehensive picture of regional population change. The input data can of course be refined and improved to give more accurate forecasts.

Region	Births	Deaths	Immigration	Outmigration	Nat.Incr.	Net Mig.	Total Change
Group 2. Net Immigration exceeds Natural Increase							
EA	131879.0	92512.0	246945.0	150593.0	39367.0	96352.0	135719.0
SW	299950.0	235343.0	471213.0	334615.0	64607.0	136598.0	201205.0
Group 3. Natural Increase exceeds Net Immigration							
EM	338821.0	217178.0	356130.0	311553.0	121643.0	44577.0	166220.0
Group 4. Natural Increase exceeds Net Outmigration							
N	269621.0	197011.0	182892.0	234347.0	72610.0	-51455.0	21155.0
NW	576306.0	421354.0	345263.0	474806.0	154952.0	-129543.0	25409.0
YH	370535.0	259589.0	253147.0	344180.0	110946.0	-91033.0	19913.0
WM	461263.0	268335.0	313730.0	369739.0	192928.0	-56009.0	136919.0
SE	1403069.0	963279.0	1248352.0	1459655.0	439790.0	-211303.0	228487.0
W	218236.0	172954.0	163583.0	167375.0	45282.0	-3792.0	41490.0
S	462961.0	311679.0	201296.0	314781.0	151282.0	-113485.0	37797.0

Distribution of Population in the Main System					
Region	Old Pop.	%	New Pop.	%	% Change in Pop.
N	3309422.0	6.239	3330574.8	6.161	0.639
NW	6643824.0	12.524	6669232.1	12.336	0.382
YH	4285147.0	8.078	4305062.2	7.963	0.465
EM	3834846.0	7.229	4001067.3	7.401	4.334
WM	4985006.0	9.397	5121923.9	9.474	2.747
EA	1556829.0	2.935	1692547.1	3.131	8.718
SE	16808129.0	31.685	17036616.3	31.513	1.359
SW	3693161.0	6.962	3894371.2	7.204	5.448
W	2694684.0	5.080	2736172.4	5.061	1.540
S	5236064.0	9.871	5273858.7	9.755	0.722
	53047112.0		54061425.9		

Table 21 : Analysis of Regional Population Characteristics, 1966-71

Region	Births	Deaths	Immigration	Outmigration	Nat.Incr.	Net Mig.	Total Change
Group 2. Net Immigration exceeds Natural Increase							
EA	142788.0	100157.0	303577.0	189211.0	42631.0	114366.0	156997.0
SW	315731.0	247717.0	560019.0	394102.0	68014.0	165917.0	233931.0
Group 3. Natural Increase exceeds Net Immigration							
EM	353215.0	226398.0	407800.0	343141.0	126817.0	64659.0	191476.0
Group 4. Natural Increase exceeds Net Outmigration							
N	271415.0	198324.0	200963.0	237735.0	73091.0	-36772.0	36319.0
NW	578697.0	423104.0	373661.0	522531.0	155593.0	-148870.0	6723.0
YH	372404.0	260898.0	276038.0	334437.0	111506.0	-58399.0	53107.0
WM	473892.0	275681.0	350656.0	385571.0	198211.0	-34915.0	163296.0
SE	1422708.0	976761.0	1362376.0	1639931.0	445947.0	-277555.0	168392.0
W	221630.0	175644.0	185519.0	192728.0	45986.0	-7209.0	38777.0
Group 5. Net Outmigration exceeds Natural Increase							
S	466400.0	313995.0	216523.0	376851.0	152405.0	-160328.0	-7923.0

Distribution of Population in the Main System					
Region	Old Pop.	%	New Pop.	%	% Change in Pop.
N	3291589.0	6.099	3327906.1	6.050	1.103
NW	6681780.0	12.381	6688504.9	12.159	0.101
YH	1297641.0	7.963	4350747.2	7.909	1.236
EM	3973665.0	7.303	4165139.0	7.572	4.819
WM	5115755.0	9.479	5279049.9	9.597	3.192
EA	1661725.0	3.079	1818722.7	3.306	9.448
SE	17128990.0	31.739	17297382.8	31.444	0.983
SW	3868195.0	7.168	4102127.4	7.457	6.048
W	2725365.0	5.050	2764141.1	5.025	1.423
S	5223725.0	9.679	5215803.1	9.482	-0.152
	53968430.0		55009524.2		

Table 22 : Analysis of Regional Population Characteristics, 1971-76

6. SUMMARY

A series of statements can be used to indicate some of the important features and conclusions of the paper:

(1) It has been demonstrated that migration is the most important component of regional demographic change, but it has also become clear that the projection or forecasting of inter-regional migration can be investigated through a variety of approaches.

(2) A three-stage process for migration projection has been described:

   (a) The projection of regional out- and in-migration totals,
   (b) The distribution of inter-regional migration flows, using the projected out- and in-migration totals.
   (c) The construction of population accounts using the projected distribution of inter-regional exist/survive migrants.

(3) Alternative methods of projecting total regional outmigration and inmigration have been considered, which make use of historical out- and in-migration rates, time series net migration information and estimated values for socio-economic indicators. The most accurate total out- and total in-migration projections for the pseudo-projection period, 1966-71, were those produced on the basis of assuming a 10% growth in the general rate of inter-regional in-migration associated with 1961-66, and regional adjustment using the 1966-71 employment shift multipliers. Severe problems were encountered when relating net migration time series data with regional indicators.

(4) The observed distribution of inter-regional migrants, 1966-71, adjusted for boundary changes, has been employed as test data in order to compare and evaluate a range of different distribution models used in projection mode, and containing different assumptions relating to unchanged rates, parameters or variable values. The growth factor model appears to generate the best-fit distribution and the continguity/non-contiguity split model is also relatively good. As a general rule, the singly constrained spatial interaction models are less accurate than either the generalised parameter or zone-specific parameter, doubly constrained models.

(5) The population accounts for 1966-71 indicate that the trends in regional population change evident for 1961-66 have continued, with East Anglia and the South West being the regions of major population

growth. The accounts projected for 1971-76 reveal that, according to the assumptions made, Scotland has experienced a net decline in population due to outmigration.

The methods discussed in this paper represent a preliminary investigation of the various ways in which migration projection models can be developed, tested and integrated into a framework for overall population projection. Some of the results have been disappointing but they have contributed, nevertheless, to an appraisal of the problems involved in migration projection, particularly those associated with data, and to an improved understanding of the relationship between migration in the past and migration in the future.

Acknowledgement

The author wishes to acknowledge the valuable comments made by Dr Philip Rees, on earlier versions of this paper.

References

Baxter, R., and Williams, I. (1978) Population forecasting and uncertainty at the national and local level. Progress in Planning, 9(1), Pergamon, Oxford.

Blanco, C. (1963) The determinants of interstate population movements. Journal of Regional Science, 5, 74-84.

Brass, W. (1974) Perspectives in population prediction: illustrated by the statistics of England and Wales. Journal of the Royal Statistical Society, Series A, 137(4), 532-57.

Central Statistical Office (1972) Abstract of Regional Statistics. No. 8, H.M.S.O., London.

Cordey-Hayes, M., and Gleave, D. (1974) Migration movements and the differential growth of city regions in England and Wales. Papers of the Regional Science Association, 33, 99-123.

Cordey-Hayes, M., and Gleave, D. (1975) Dynamic models of the interaction between migration and the differential growth of cities. In Baxter, R., Echenique, M., and Divers, J. (Ed.) Urban development models. L.U.B.F.S. Conference Proceedings No. 3, The Construction Press, London.

Department of Employment and Productivity (1971) British labour statistics Historical Abstract, 1886-1968. H.M.S.O., London.

Donovan, S. (1971) The calibration of a regional migration flow model. M.Sc. dissertation. Department of Computational Science, University of Leeds.

Drewe, P. (1971) Steps towards action-oriented migration research. Papers, Regional Science Association, 26, 145-64.

Ginsberg, R.B. (1973) Stochastic models of residential and geographic mobility for heterogeneous populations. Environment and Planning, 5, 113-24.

General Register Office, Edinburgh (1965) Census 1961 Scotland, Vol. 2. Usual Residence. H.M.S.O., Edinburgh.

General Register Office, Edinburgh (1965) Census 1961 Scotland, Occupation and Industry National Summary Tables. H.M.S.O., Edinburgh.

General Register Office, Edinburgh (1968) Sample Census 1966 Scotland, Migration Tables Part 1. H.M.S.O., Edinburgh.

General Register Office, Edinburgh (1968) Sample Census 1966 Scotland, Economic Activity County Tables. H.M.S.O., Edinburgh.

General Register Office, Edinburgh (1969) Sample Census 1966 Scotland, Usual Residence and Birthplace Tables. H.M.S.O., Edinburgh.

General Register Office, Edinburgh (1974) Census 1971 Scotland, Usual Residence and Birthplace Tables. H.M.S.O., Edinburgh.

General Register Office, London (annual to 1968) The Registrar General's Statistical Review of England and Wales for the year 19--, Part II, Tables, Population. H.M.S.O., London.

General Register Office, London (1964) Census 1961, England and Wales, Usual Residence Tables. H.M.S.O., London.

General Register Office, London (1965) Census 1961, England and Wales, Occupation, Industry, Socio-economic Groups. H.M.S.O., London.

General Register Office, London (1968) Sample Census 1966 England and Wales, Economic Activity Leaflets. H.M.S.O., London.

General Register Office, London (1969) Sample Census 1966 England and Wales, Usual Residence Tables. H.M.S.O., London.

General Register Office, London (1972) Census 1971 England and Wales, County Reports (Table 4: Intercensal changes of boundary). H.M.S.O., London.

General Register Office, London (1974) Census 1971 England and Wales, Usual Residence Tables. H.M.S.O., London.

Greenwood, M.J. (1975a) Simultaneity bias in migration models: an empirical examination. Demography, 12, 519-36.

Greenwood, M.J. (1975b) A simultaneous equations model of urban growth and migration. Journal of the American Statistical Association, 70, 797-807.

Greenwood, M.J. (1975c) Research on internal migration in the United States; a survey. Journal of Economic Literature, 13(1), 397-433.

Hagerstrand, T. (1957) Migration and area: survey of Swedish migration fields and hypothetical considerations on their genesis. In Hannerberg, D., Hagerstrand, T., and Odeving, B. (Eds.) Migration in Sweden. A Symposium. Lund Studies in Geography, Series B, Human Geography, No. 3, 27-158, Gleerup, Lund.

Hart, R.A. (1970) A model of inter-regional migration in England and Wales. Regional Studies, 4, 279-96.

Jack, A.B. (1970) A short-run model of inter-regional migration. Manchester School of Economic and Social Studies, 38(1), 15-28.

Jenkins, J.C., and Rees, P.H. (1977) Computer programmes for aggregate accounts-based and associated forecasting models of population. Working Paper 205/Computer Manual 2, School of Geography, University of Leeds.

Joseph, G. (1977) Inter-regional population distribution and growth in Britain - a projection exercise. Scottish Journal of Political Economy, XXI(2), 159-69.

Joseph, G. (1975) A Markov analysis of age/sex differences in inter-regional migration in Great Britain. Regional Studies, 9, 69-78.

Ledent, J. (1978) Regional multiplier analysis: a demometric approach. Environment and Planning A, 10, 537-60.

Lianos, T.P. (1972) The migration process and time lags. Journal of Regional Science, 12(3), 425-33.

Lowry, I.S. (1966) Migration and metropolitan growth: two analytical models. Chandler, San Fransisco.

Mackinnon, R.D. (1975) Controlling interregional migration processes of a Markovian type. Environment and Planning A, 7, 781-92.

Masser, I. (1970) A test of some models for predicting inter-metropolitan movement of population in England and Wales. Working Paper No. 9, Centre for Environmental Studies, London.

Miller, A.R. (1967) The migration of employed persons to and from Metropolitan areas in the United States. Journal of the American Statistical Society, 62, 1418-32.

Miller, E. (1973) Is outmigration affected by economic conditions? Southern Economic Journal, 9(3), 396-405.

Morrison, P.A. (1971) Chronic movers and the future redistribution of population: a longitudinal analysis. Demography, 2, 171-84.

Morrison, P.A. (1973) Theoretical issues in design of population mobility models. Environment and Planning A, 5, 125-34.

Morrison, P.A., and Relles, D.A. (1974) A method of estimating and projecting migration flows for U.S. Cities. Paper presented at the British Section Meeting of the Regional Science Association, London.

Office of Population Censuses and Surveys (annual from 1969) The Registrar General's Statistical Review of England and Wales for the year 19--. Part II Tables, Population. H.M.S.O., London.

Office of Population Censuses and Surveys (1975) Census 1971 England and Wales, Economic Activity County Leaflets. H.M.S.O., London.

Okun, B. (1968) Interstate population migration and state income inequalities. A simultaneous equations approach. Economic Development and Cultural Change, 16, 297-311.

Oliver, F.R. (1964) Inter-regional migration and unemployment, 1951-61. Journal of the Royal Statistical Society Series A, 127, 42-69.

Pittenger, D.B. (1976) Projecting State and Local Populations. Ballinger, Cambridge, Mass.

Ramprakash, D. (1973) Official sub-national population projections. Office of Population Censuses and Surveys, H.M.S.O., London.

Rees, P.H. (1977a) The measurement of migration from census data and other sources. Environment and Planning A, 9, 247-72.

Rees, P.H. (1977b) Population forecasts for British regions: a comparison. Paper presented at the Joint Meetings of the British Society for Population Studies and the Population Geography Study Group at the University of Liverpool (21-23 September). Published in Hobcraft, J., and Rees, P.H. (Eds.) Spatial demographic development. Croom Helm, London, 1980.

Rees, P.H. (1978) Migration and settlement in the United Kingdom, dynamics and policy. Working Paper WP-78-36. International Institute for Applied Systems Analysis, Laxenburg, Austria.

Rees, P.H., and Wilson, A.G. (1973) Accounts and models for spatial demographic analysis 1: aggregate population. Environment and Planning, 5, 61-90.

Rees, P.H., and Wilson, A.G. (1977) Spatial population analysis. Arnold, London.

Rogers, A. (1966) A Markovian policy model of inter-regional migration. Papers, Regional Science Association, 17, 205-24.

Rogers, A. (1967) Regression analysis of inter-regional migration in California. Review of Economics and Statistics, 49, 262-67.

Rogers, A. (1978) Demometrics of migration and settlement. In Batey, P.J. (Ed.) Theory and Method in Urban and Regional Analysis, London Papers in Regional Science 8. Pion, London, 1-30.

Saboia, J.L.M. (1974) Modelling and forecasting populations by time series: the Swedish case. Demography, 11(3), 483-92.

Senior, M.L. (1979) From gravity modelling to entropy maximising: a pedagogic guide. Progress in Human Geography, 3(2), 175-210.

Shaw, R.P. (1975) Migration theory and fact: a review and bibliography of current literature. Bibliography Series No. 5, Regional Science Research Institute, Pennsylvania.

Stillwell, J.C.H. (1975) Models of inter-regional migration: a review. Working Paper 100, School of Geography, University of Leeds.

Stillwell, J.C.H. (1976) A user's guide to a simple matrix aggregation programme. Working Paper 141, School of Geography, University of Leeds.

Stillwell, J.C.H. (1977) Some historical tests of spatial interaction models for inter-area migration: Part 1: inter-county migration. Working Paper 185, School of Geography, University of Leeds.

Stillwell, J.C.H. (1977) Some historical tests of spatial interaction models for inter-area migration: Part 2: inter-region migration. Working Paper 199, School of Geography, University of Leeds.

Stillwell, J.C.H. (1978) Inter-zonal migration: some historical tests of spatial interaction models. Environment and Planning A, 10, 1187-1200.

Stillwell, J.C.H. (1979) Migration in England and Wales: a study of inter-area patterns, accounts, models and projections. Unpublished Ph.D. thesis. School of Geography, University of Leeds.

Weeden, R. (1973) Interregional migration models and their application to Great Britain. National Institute of Economic and Social Research Papers II. Cambridge University Press, London, 43-105.

Wilson, A.G. (1974) Urban and regional models in Geography and Planning. Wiley, London.

APPENDIX. PREPARATION OF THE OBSERVED MIGRATION DATA BASE FOR 1966-71, REQUIRED FOR COMPARING THE PROJECTED INTER-REGIONAL MIGRATION DISTRIBUTIONS.

Analyses of age/sex aggregated and disaggregated inter-county and inter-region migration and survival flows have been undertaken for the periods 1961-66 and 1965-66 (Stillwell, 1979), based on observed data (G.R.O., 1968) for a set of aggregated-counties which has been consolidated to produce matrices of inter-region migration flows where the regions are those illustrated in Figure 2.

Age/sex aggregated inter-county migration flows data for the periods 1966-71 and 1970-71 has also been collected from microfilm tabulations at the Office of Population Censuses and Surveys and some aggregation of this information is required to produce an inter-region matrix for 1966-71 which can be used as an observed data set against which flows projected in various ways can be compared and evaluated. Direct compatibility between the two five year period observed information sets is necessary because some of the model tests reported in the paper use 1961-66 migration flows or rates between regions as primary ingredients for projecting movement between the same spatial units in 1966-71. However, further adjustment to certain elements of the 1966-71 matrix is necessary because several counties have experienced boundary changes between the respective census dates in 1966 and 1971. The boundary change problem implies an additional complication associated with inter-area migrants becoming intra-area migrants and vice versa upon readjustment, which is in itself a potentially extensive research area. This appendix contains a short outline of how the complication arises and the type of procedure that is required to re-estimate the migration flows involved.

The ten aggregated-counties that experienced boundary changes during 1966-71 were Durham, Yorkshire; North Riding, Yorkshire; East Riding, Yorkshire; West Riding, Derbyshire, Hereford and Worcester, Warwickshire, Greater London, Kent and Surrey (G.R.O., 1972). In some cases the change involved a regional/county boundary; in other cases, the change involved a county boundary within one particular region. The 1971 usually resident census population totals have been adjusted to be consistent with the 1966 population totals using the following estimating equation, expressed verbally as:

$$\text{Adjusted Usually-resident population of county i} = \text{Usually-resident 1971 population of county i} \pm \frac{\text{1971 Census estimate of 1961 enumerated population of zone x that is involved in boundary change.}}{\text{1971 Census estimate of 1961 enumerated population of county i.}} \text{Enumerated population of county i, 1971} \quad (A.1)$$

and the adjusted figures for both aggregated counties and regions are presented in Table A.1.

The readjustment procedure for the 1966-71 inter-county flow matrix is undertaken in three stages. Initially, the flows between the pairs of counties involved in each boundary change are set equal to zero and the row and column totals of the matrix are altered accordingly. The matrix is then multiplied using a proportional aggregation matrix which serves to reduce all the outflows from and inflows to each county that gained population with the change in boundary during the period, and to increase all the outflows from and inflows to each county that lost population through boundary change. Yorkshire, West Riding, for example, gained population from Derbyshire in 1967 because of boundary reorganisation and therefore all the flows into and out of Yorkshire, West Riding are reduced by a factor proportional to this gain:

$$\frac{\text{Gain in population by Yorkshire, West Riding}}{\left[\begin{matrix}\text{1961 Census population of}\\ \text{Yorkshire, West Riding}\end{matrix}\right]\left[\begin{matrix}\text{1961 Census population}\\ \text{of Derbyshire}\end{matrix}\right]}$$

which in numerical terms is:

$$\frac{32{,}224}{[3{,}644{,}582][877{,}620]} = 0.00713$$

This proportion of each directional flow out of or in to Yorkshire, West Riding from other counties except Derbyshire is subtracted and added to the flows out of and into Derbyshire respectively. Proportional factors are defined for the other counties in a similar way.

The third stage of the adjustment procedure requires estimation of the flows between the pairs of counties directly concerned in the boundary change (set to zero in the matrix). The problem is outlined in Figure A.1 using a set of four hypothetical areas A,B,C,D as an example system of interest, where x is the zone actually changing from C's ownership to B's ownership during the period. The necessary readjustment is to transform the system described by the areas A,B+x, C-x and D (identified on the horizontal) back to the system described by the areas A,B,C,D

Aggregated-Counties	1961 Populations		1966 Populations+		1971 Populations		Regions
	Counties	Regions	Counties	Regions	Counties	Regions	
1. Cumbria	358269	3255412	366730	3309422	362335	3291589	1 North
2. Durham	1520688		1538289		1528410*		
3. Northumberland	821820		821767		794645		
4. Yorkshire, N.R.	554635		582636		606199*		
7. Cheshire	1373087*	6507572	1473035	6643824	1554925	6681780	2 North West
8. Lancashire	5134485		5170789		5126855		
5. Yorkshire, E.R.	527236	4177822	543994	4285147	543427*	4297641	3 Yorks. & Humberside
6. Yorkshire, W.R.	3650586		3741153		3754214*		
9. Derbyshire	879171*	3634199	913848	3834846	920745*	3973665	4 East Midlands
10. Leicestershire	706373*		755863		796060		
11. Lincolnshire	744397*		781980		811840		
12. Northamptonshire	399066*		430777		469785		
13. Nottinghamshire	905192*		952378		975235		
14. Hereford & Worcester	754204*	4762052	796493	4985006	832473*	5115755	5 West Midlands
15. Salop	298130*		314356		337550		
16. Staffordshire	1690854*		1803182		1861620		
17. Warwickshire	2019664*		2070975		2084112*		
18. Cambridgeshire	430500*	1463599	463790	1556829	499815	1661725	6 East Anglia
19. Norfolk	560933		581084		615760		
20. Suffolk	472166		511955		546150		
21. Bedfordshire	383399*	16182789	432560	16808129	464925	17128990	7 South East
22. Berkshire	505150		575283		634475		
23. Buckinghamshire	484148*		538487		589445		
24. Essex	1105689*		1253762		1362080		
25. Greater London	8018519*		7809105		7479528*		
26. Hampshire	1330822		1455469		1560005		
27. Hertfordshire	788797*		867559		927150		
28. Kent	1197658*		1319351		1397809*		
29. Oxfordshire	301897		336648		373455		
30. Surrey	904291*		967246		999698*		
31. Sussex East	660798		701481		741700		
32. Sussex West	408772		453729		491885		
33. Isle of Wight	92849		97649		106895		
34. Cornwall & Scilly Is.	339101*	3485887	355371	3693161	378505	3868195	8 South West
35. Devon	813521*		850429		889640		
36. Dorset	311673*		329721		358695		
37. Gloucestershire	999530*		1063201		1074715		
38. Somerset	596253*		639148		679010		
39. Wiltshire	423809*		455291		487630		
40. N. Wales	529959	2640175	550027	2694684	573500	2725365	9 Wales
41. Mid Wales	117950		116249		114560		
42. West Wales	314782		315096		313815		
43. South Wales	1677484		1713312		1723490		
44. Scotland	5179344	5179344	5236064	5236064	5223725	5223725	10 Scotland

+ 1966 Census totals have been upgraded for underenumeration

* Figure adjusted for boundary change

Table A.1: Usually-Resident (Census Definition) Populations of Aggregated-Counties and Regions, 1961, 1966, 1971

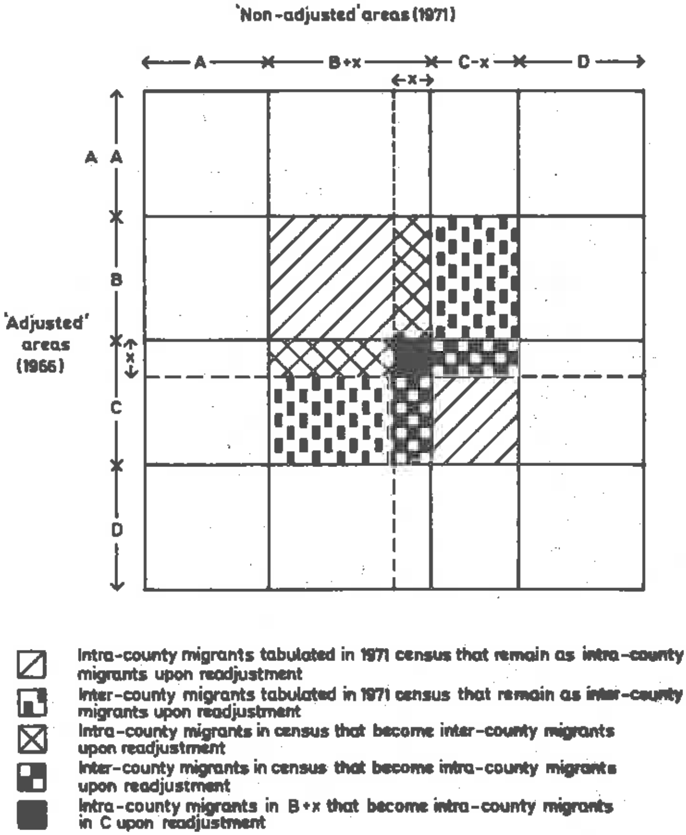


Figure A.1: <u>The Inter/Intra-area Migration Estimation Problem when Boundary Change has Occurred</u>

(identified on the vertical) where B is the area that loses when readjustment takes place, and C is the area that gains.

An overall constraint equation with respect to the flows within and between areas B and C can be defined either using unadjusted elements:

$$\sum_{B,C} M_{BC} = M_{B+x,B+x}+M_{C-x,C-x}+M_{B+x,C-x}+M_{C-x,B+x} \quad (A.2)$$

or using adjusted flows:

$$\sum_{B,C} M_{BC} = M_{BB}+M_{CC}+M_{BC}+M_{CB} \quad (A.3)$$

The total migration within the system will obviously be the same in both cases. The terms on the right hand side of equation (A.2) are the flows available from the 1971 Census, and the terms on the right hand side of equation (A.3) are the adjusted flows. The adjusted flows can be defined individually as:

$$M_{BB} = M_{B+x,B+x}-M_{B,x}-M_{x,B}-M_{xx} \quad (A.4)$$

$$M_{CC} = M_{C-x,C-x}+M_{C-x,x}+M_{x,C-x}+M_{xx} \quad (A.5)$$

$$M_{BC} = M_{B+x,C-x}-M_{x,C-x}+M_{B,x} \quad (A.6)$$

and

$$M_{CB} = M_{C-x,B+x}-M_{C-x,x}+M_{x,B} \quad (A.7)$$

where some of the terms on the right hand side of the equations are known and others are unknown. The flows that are unknown are as follows:

(1) $M_{C-x,x}$
(2) $M_{x,C-x}$
(3) $M_{x,B}$
(4) $M_{B,x}$
(5) $M_{x,x}$

and each of these can be estimated independently on the basis of the probability of being located in respective areas at the start and/or the end of the period. The unknown terms in the eqation set (A.4)-(A.7) cancel out to ensure that the constraint equation (A.3) is satisfied. The unknown flows are estimated as proportions of the observed flows as follows:

$$(1) \quad M_{C-x,x} = M_{C-x,B+x} \frac{P_x(t)}{P_B(t)P_x(t)} \qquad (A.8)$$

$$(2) \quad M_{x,C-x} = M_{B+x,C-x} \frac{P_x(t)}{P_B(t)P_x(t)} \qquad (A.9)$$

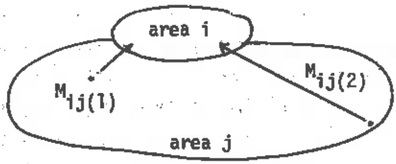
$$(3) \quad M_{x,B} = M_{B+x,B+x} \frac{P_x(t)}{P_B(t)P_x(t)} \frac{P_B(t+T)}{P_B(t+T)P_x(t+T)} \qquad (A.10)$$

$$(4) \quad M_{B,x} = M_{B+x,B+x} \frac{P_B(t)}{P_B(t)P_x(t)} \frac{P_x(t+T)}{P_B(t+T)P_x(t+T)} \qquad (A.11)$$

$$(5) \quad M_{x,x} = M_{B+x,B+x} \frac{P_x(t)}{P_B(t)P_x(t)} \frac{P_x(t+T)}{P_B(t+T)P_x(t+T)} \qquad (A.12)$$

where P(t) and P(t+T) represent populations of the areas subscripted in 1966 and 1971.

When the estimating equations are applied, the adjusted flows presented in Table A.2 are produced. It appears that most of the flow adjustments are reasonable in relation to the 1961-66 observed figures, apart from those between Durham and Yorkhisre, North Riding, where inter-area flows are <u>over</u>-estimated and intra-area flows are <u>under</u>-estimated. There are several reasons why the inter-area flows are overestimated. Ideally, the probability of an internal migration crossing a boundary when boundary change adjustment is undertaken, depends upon where the migrant comes from and how close his location is to the boundary. The closer that the migrant is located to the boundary, the higher will be the probability (p) that the migration will actually involve a boundary crossing. In general terms, if p_1 is the probability associated with $M_{ij(1)}$ and p_2 is the probability associated with $M_{ij(2)}$



then $p_1 > p_2$ rather than $p_1 = p_2$, which is the implicit assumption. Relaxation of the assumption of equal probabilities implies some measurement of the distance involved in migration.

Counties Involved	Flow	Unadjusted 1966-71 Flow	Adjusted 1966-71 Flow	1961-66 Flow
B* Yorkshire W.R.	M_{BB}	1060310	1041807	1011920
	M_{CC}	185390	185629	195420
C** Derbyshire	M_{BC}	10880	19996	15020
	M_{CB}	7220	16368	9120
	Total	1263800	1263800	
B Yorkshire N.R.	M_{BB}	183740	126606	135070
	M_{CC}	396350	404932	416940
C Durham	M_{BC}	7910	32481	9080
	M_{CB}	11380	35361	10770
	Total	599380	599380	
B Yorkshire W.R.	M_{BB}	1060310	1058591	1011920
	M_{CC}	183740	183763	135070
C Yorkshire N.R.	M_{BC}	17660	18505	14651
	M_{CB}	9500	10351	10000
	Total	1271210	1271210	
B Yorkshire W.R.	M_{BB}	1060310	1059856	1011920
	M_{CC}	153320	153325	138700
C Yorkshire E.R.	M_{BC}	12580	12804	11500
	M_{CB}	8660	8885	8050
	Total	1234870	1234870	
B Hereford & Worcester	M_{BB}	185750	185712	165520
	M_{CC}	556160	556164	500900
C Warwickshire	M_{BC}	14890	14908	11660
	M_{CB}	31730	31746	25310
	Total	788530	788530	
B Kent	M_{BB}	309850	309470	281770
	M_{CC}	1865990	1866066	1882500
C Greater London	M_{BC}	31120	31291	26810
	M_{CB}	92540	92673	94990
	Total	2299500	2299500	
B Surrey	M_{BB}	166940	166159	155640
	M_{CC}	1865990	1866259	1882560
C Greater London	M_{BC}	34190	34500	32600
	M_{CB}	80340	80542	84110
	Total	2147460	2147460	

* B is the county of loss on readjustment

** C is the county of gain on readjustment

Table A.2: Adjusted Flows Within and Between Counties Involved in Boundary Changes, 1966-71

Two methods of re-estimating migration flows can therefore be suggested. The first method incorporates a measure of the distance migrated (d_{ij}) in the calculation of the probability of migration between any two areas i and j as follows:

$$P_{ij} = \frac{P_i(t)P_j(t+T)d_{ij}^{-\beta}}{\sum_{ij} P_i(t)P_j(t+T)d_{ij}^{-\beta}} \qquad \text{(A.13)}$$

but problems arise in the measurement of intra-area distances. A second method is therefore preferred which simply allocates the known flow total of intra- and inter-area moves for the two areas concerned, according to the 1961-66 allocation. Thus, for Yorkshire, North Riding (B) and Durham (C), the 1966-71 adjusted flows are defined as follows:

1961-66 Flows

	B	C	
B	135070	9080	
C	10770	416940	
			571860

1961-66 Proportions

	B	C	
B	0.236	0.016	
C	0.019	0.729	
			1.0

1966-71 Flows

	B	C	
B	141454	5950	
C	11388	436948	
			599380

The adjusted inter-county migration matrix is finally aggregated to provide an observed data set of inter-region migrations for 1966-71.